CHAPTER 3

Characterization

After testing the quantification method, we still had many questions to answer. What can we do with these results? Do light element concentrations really give us new insights about environmental particles? Will our efforts result in a more detailed characterization of the analyzed particles? We found our first answers by applying the classical, routine data processing software on the data obtained by our new method. After that, we developed and tested other techniques to extend our capabilities for characterizing single particles.

Table of Contents

1. Introduction

In this chapter we will discuss particle characterization on two different levels. The first will deal with *data processing*, while the second is about methods for *extended particle analysis*.

When TW-EPMA is applied for the analysis of environmental samples, the analyzed set of particles has of course been characterized to a certain degree. For each particle, we have obtained morphological and spectra data like the diameter or the elemental intensities in the spectrum. The subsequent quantification of the particle compositions through iterative Monte Carlo simulations adds some value to this initial characterization, since the interpretation can be done more exactly. However, large numbers of particles have to be analyzed in order to characterize the overall composition of a sampled aerosol, which results in high amounts of data. For example, if we would analyze about 1000 particles, it is not inconceivable that the data matrix could consist of around 20000 values, if we take into account that the particle size and the elemental concentrations of about 15-20 elements are obtained for all particles. Since a huge data matrix on its own doesn't tell much more about the analyzed aerosol, **data processing tools** have to be applied. Their main goal is to recognize the differences and similarities between the analyzed particles in order to find out which particle types are present, and to determine their relative or absolute abundance. The involved algorithms roughly consist of mainly two steps, namely 'classification' and 'identification'. Although these terms may appear to be just synonyms, they are not; however, there is indeed a close relation between both that could be called 'recognition'. *Classification* is the investigation of objects in order to establish if they can validly be summarized in terms of a small number of classes of similar objects. *Identification* corresponds to the process of assigning names to the objects, or to the classes they belong to. The latter definition indicates that identification could be performed on the level of either single or classified objects, and therefore, there is no need to fix the order in which classification and identification are applied. *Recognition* on the other hand is the ever-present, mutual process of finding patterns based on the available characteristic variables. In case of single particle analysis, the more classical way of data processing starts with the grouping of similar particles using mathematical algorithms (classification) followed by the designation of the obtained particle classes using the average concentrations or intensities of these classes (identification). Another option would be to first assign a specific particle type to each particle by the individual concentration or intensity values (identification), followed by a bookkeeping procedure to count the number of particles of every type (classification). Experiments with both methods will be discussed in section 2.

If the goal of a researcher is to study the mechanisms of particle generation and transformation, his focus will not only be on the overall composition of a sampled aerosol. For example, if he is interested in the reaction mechanisms of sea salt to other compounds like nitrates and sulfates, it is less important to know how large the fraction of soil dust particles is. He needs to find out for each particle if it consists of only pure or aged sea salt, or if it also contains other compounds. He probably would also like to know where exactly all these species occur in the particle (at the surface, inside, everywhere). For this kind of research, we need to limit or extend our analysis to a profound, in-depth study of a representative selection of particles. Experiments with techniques that were developed and optimized for this reason, **Grazing-exit** and **Multiple-Voltage TW-EPMA**, will be discussed in section 3 this chapter.

2. Data processing tools

2.1 Mathematical classification prior to identification

- **Introduction**

Soon after showing the capabilities of our iterative Monte Carlo quantification method, we took our study one step further. The first logical step was of course to subject our data to the classical data processing techniques that were already used since long for the processing of data obtained by conventional EPMA, namely clustering analysis and principal component analysis.[1,2,3,4,5,6,7,8]

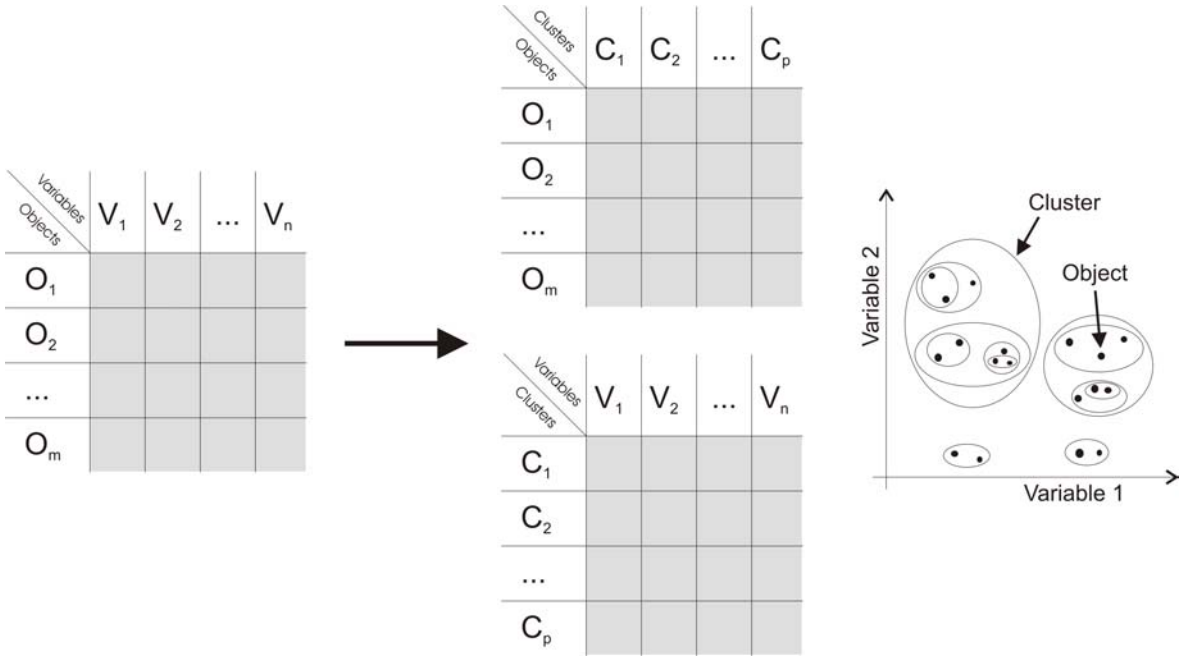

Figure 1: Clustering analysis

A data matrix consisting of n variables (e.g. the measured concentrations of elements V_n) for a set of m objects (e.g. particles O_m) could be split-up into p classes or clusters C_p by determining which objects have similar values for the observed variables. If two particles have similar properties, their concentration values will be close to each other, so the distance between two objects in an n-dimensional coordinate system is taken as a purely mathematical measure of their similarity. The two most similar or closest objects are then combined into a new cluster object, after which the cluster algorithm is repeated again and again until all objects of the data matrix are classified into a smaller set of identifiable clusters that allows for an easier interpretation. In the end, each particle is assigned to a cluster, and the average elemental concentrations are known for each cluster.

The most important factors for the classification process are 1) the pre-processing of the raw data, 2) the calculation of the distance between the objects, 3) the method for combining different objects into new cluster objects, and 4) the decision about when to stop the clustering. In the sections below, we will discuss them in more detail in view of the application of clustering analysis to datasets obtained with TW-EPMA. However, it is not in the scope of this thesis to provide an extensive discussion about the different aspects of clustering analysis; these can be found in literature.[9,10]

Principal component or factor analysis is also a purely mathematical method to distinguish different objects and variables from each other (see Figure 2).[11] By performing a transformation of the abovementioned *n*-dimensional coordinate system, it is possible to split the data matrix into two new matrices that reflect the hidden relationships between either the variables or the objects. The matrix of the so-called *loadings* then contains independent components which describe the relationships between the variables, while the matrix of the *scores* contains similar components to represent the similarities between the objects. These components can also be graphically represented in a coordinate system of certain dimensions to show the similarities between the investigated objects or variables. Therefore, it should also possible to obtain particle classification in an indirect way, based on the interpretation of the loading and score plots (often combined in a bi-plot). Like in the clustering procedure, the pre-processing of the dataset is important, as are other factors.

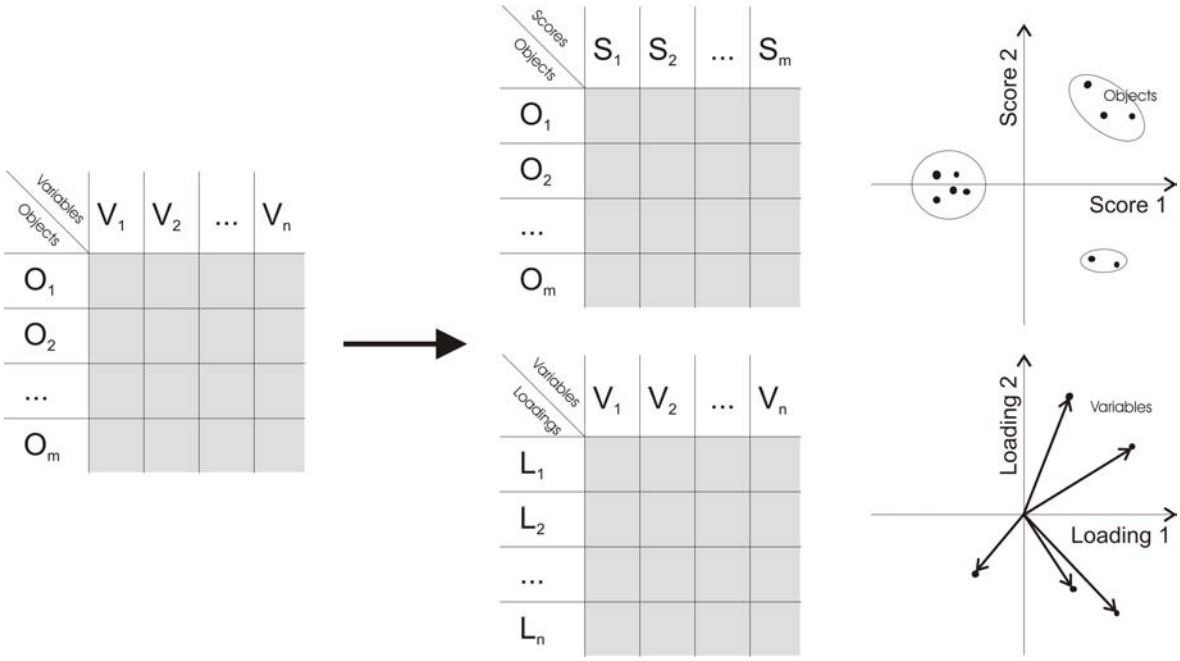

Figure 2: Principal component analysis

- **Experiments**

The aim of our experiments was 1) to demonstrate the applicability of cluster and principal component analyses for extracting the most relevant information from low-*Z* EPMA data sets, 2) to test the classification methods using a mixture of standard particles, 3) to investigate the influence of different input data (X-ray intensities or elemental concentrations) and scaling functions on the classification results.[12]

Before the arrival of TW-EPMA, clustering and principal component analysis were already applied on datasets of particles that were analyzed by conventional EPMA. Of course, these datasets mostly consisted of intensities, because of the lack of adequate quantification procedures. TW-EPMA now offers concentration values which are expected to provide a better particle classification and identification. Therefore, our experiments were set up to compare the differences between characteristic X-ray intensities (I_i), square root scaled intensities ($\sqrt{I_i}$) and calculated semi-quantitative concentration values (c_i) as input data ($i = 1, \ldots, n$; n denotes the total number of determined elements). In addition to the different raw data, the effect of normalization to the total sum or to the maximum, as well as standardization using the *z*-transform was tested.[9,11,12]

The data set used for testing the different clustering techniques, scaling, normalization and standardization methods, consists of thin-window EPMA data of 14 different particle types, each with known abundance. The "expected" composition of the standard mixture is shown in Table 1. Using this *a priori* knowledge about the structure of the data set, the number of clusters was set to 14 for all of the tested combinations. The standard particle types were selected in such a way, that the capabilities of the thin-window EPMA could be shown maximally in the data set. Particle types that are generally indistinguishable with conventional EPMA were also included, e.g. two kinds of biogenic particles, five types of particles containing C, O and Na, and also ammonium salts, which are very common in the atmosphere.

Table 1: Composition of the test particulate standard mixture

Particle	Number	Average concentration (wt%)									Diam.
type	analysed	C	N	O	Na	Si	S	Cl	Ca	Fe	(μm)
Algae	10	57.3	13.3	25.3	0.0	0.2	1.0	0.1	0.1	0.0	5.9
Lichen	11	67.7	0.0	29.9	0.0	1.6	0.0	0.4	0.1	0.0	3.7
$CaCO_3$	10	11.6	2.8	52.1	0.0	0.0	0.0	0.0	33.4	0.0	2.4
CaO	15	7.4	3.4	49.2	0.0	0.3	0.0	0.1	39.2	0.0	3.1
Na-acetate	6	29.6	1.1	37.4	31.2	0.0	0.1	0.4	0.0	0.0	37.3
Na-formate	5	14.4	0.0	45.3	40.3	0.0	0.0	0.0	0.0	0.0	46.8
Na-oxalate	10	12.4	0.1	42.3	42.8	2.4	0.0	0.0	0.0	0.0	3.1
Na_2CO_3	10	10.2	0.2	49.6	39.4	0.6	0.0	0.0	0.0	0.0	5.0
$NaHCO_3$	7	10.6	0.0	54.2	35.1	0.0	0.2	0.0	0.0	0.0	30.2
Fe_2O_3	10	1.9	1.0	33.2	0.0	0.2	0.0	0.0	0.0	63.7	4.0
NaCl	10	0.0	0.0	0.7	43.9	0.0	0.2	54.9	0.0	0.0	4.4
NH_4NO_3	9	1.1	34.8	64.0	0.0	0.1	0.0	0.0	0.0	0.0	14.2
$(NH_4)_2SO_4$	10	0.4	20.8	57.5	0.0	0.0	21.3	0.0	0.0	0.0	12.5
NH_4HSO_4	3	1.0	20.1	58.7	0.0	0.0	20.3	0.0	0.0	0.0	13.5

In the past, two types of clustering analysis were developed, namely hierarchical (HCA) and non-hierarchical (NHCA). The difference between both is the way in which objects are assigned to clusters using a certain measure of similarity. In classical, *hierarchical clustering*, the clustering consists of a one-way cascade of consecutive assignments of objects to clusters as described before. This type of clustering is said to impose a pattern on the dataset, since each object is irrevocably fixed to a cluster. After each assignment in *non-hierarchical clustering*, it is possible to re-evaluate all of the previous assignments, for it could be that a clustered object better fits to another cluster.

Therefore, we tested both HCA and NHCA to check their capability to distinguish the different particle types. The cluster analysis calculations were carried out using the statistical software package DPP developed at the University of Antwerp.[13] The Ward's method was chosen for the strategy of HCA, since it was proven to be the most suitable for clustering of aerosol particle data.[14] Both for HCA and NHCA the Euclidean distance was used as a measure of the dissimilarity of the particles. The NHCA calculations were done using the *K*-median method by the MASLOC algorithm.[9] The calculations were carried out using the SIMCA-P software package (Umetri AB and Ericsson Erisoft AB, Sweden), based on the covariance matrix of the different raw data.

- **Hierarchical clustering analysis**

Tables 2a-e show the results of HCA on the standard test data set, using different input data, normalization and scaling, with and without standardization using *z*-transform. The sampling errors shown in the last row correspond to the standard deviation of the expected binomial distribution for the particle types in the sample. Since HCA is a so-called unsupervised classification method, the procedure is purely mathematical without interference by the analyst. The obtained 14 clusters are not necessarily the same as the expected particle types in Table 1; therefore, the obtained clusters are denoted only by numbers.

The particle types that were the least similar to the others, *iron oxide* and *ammonium nitrate*, were classified as separate groups, independently on the input data and the data treatment used. The *NaCl* particles also formed a separate group when the data were not subjected to standardization using the *z*-transform. The standardization resulted in equal weight for every variable, but according to the measuring conditions and the capabilities of the EPMA technique, elements with concentrations of less than a few percent in the particles are subjected to larger measurement errors. In other words, the technique is only capable of detecting major and minor elements and not trace elements. Therefore, as the *z*-transform increases the weight of trace elements compared to major elements, it is not recommended to use for classification of particles using thin-window EPMA data. Table 2b shows the effect of the *z*-transform, showing that the diversity of particles in the unimportant variables is more important than great differences in the variables connected to major elements. Therefore, all the particles containing C, O and Na were classified to one large group, and a relatively large number of singlets were observed, according to the Mg, Si or Ti content of the particles. This is the reason of the formation of two groups from the NaCl particles, because of the higher Si content of one particle.

If the clustering was carried out without *z*-transform, the *ammonium sulfate* and *ammonium bisulphate* particles usually formed one group. This is most probably due to the fact that these compounds are very sensitive to humidity and electron bombardment. The composition after bombardment is most probably $(NH_4)_2SO_4$, because this compound is the more stable. The N/S concentration ratios in Table 1 also indicate the presence of ammonium sulfate in both compounds. The difference between the average concentrations is less than 5 % in these two compounds, so the classification of them to one group is natural.

Biogenic particles are very common in atmospheric aerosols, but they can consist of diverse types of micro-organisms, fragments of plants and insects. The most important difference between the two standard biogenic particle types used for this study is their nitrogen content, as can be seen in Table 1. Using HCA, these particle types can also be distinguished, when the clustering is based on concentration or square root scaled intensity values (see Table 2a and 2d). Using the intensity values as input for HCA, all of the biogenic particles form one group (Class 1 in Table 2e). Even if the number of clusters is increased, the resulting two groups cannot be connected to algae and lichen. From Table 2, one can remark that one algae particle is always classified as lichen, and one Na-acetate particle is usually regarded as biogenic, too. If we look at the X-ray spectra of these particles, the low nitrogen intensity of the alga particle and the high carbon peak found in the spectrum of the Na-acetate particle indicates that the "false" classification of these particles is not caused by the clustering.

Table 2: Classification of the test particulate standard mixture using HCA, SE denotes the sampling error connected to each class

(a) Based on concentration values, normalized to the maximum, not standardized

	Classified to group No:													
Particle type	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
Algae	-	-	1	-	-	-	-	-	9	-	-	-	-	-
Lichen	-	-	11	-	-	-	-	-	-	-	-	-	-	-
$CaCO_3$	-	-	-	4	-	-	-	-	-	-	6	-	-	-
CaO	-	-	-	9	-	-	-	-	-	-	3	3	-	-
Na-acetate	2	-	1	-	-	-	-	-	-	-	-	-	3	-
Na-formate	4	-	-	-	-	-	-	-	-	1	-	-	-	-
Na-oxalate	8	-	-	-	-	-	-	-	-	1	-	-	-	1
Na_2CO_3	-	-	-	-	-	-	3	-	-	7	-	-	-	-
$NaHCO_3$	-	-	-	-	-	-	7	-	-	-	-	-	-	-
Fe_2O_3	-	-	-	-	-	10	-	-	-	-	-	-	-	-
NaCl	-	-	-	-	10	-	-	-	-	-	-	-	-	-
NH_4NO_3	-	-	-	-	-	-	-	9	-	-	-	-	-	-
$(NH_4)_2SO_4$	-	10	-	-	-	-	-	-	-	-	-	-	-	-
NH_4HSO_4	-	3	-	-	-	-	-	-	-	-	-	-	-	-
Total	14	13	13	13	10	10	10	9	9	9	9	3	3	1
SE	4	4	4	4	3	3	3	3	3	3	3	2	2	1

(b) Based on concentration values, normalized to the maximum, standardized

	Classified to group No:													
Particle type	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
Algae	-	-	-	-	2	-	-	-	6	-	1	-	-	1
Lichen	-	-	-	-	8	-	-	-	-	2	1	-	-	-
$CaCO_3$	-	10	-	-	-	-	-	-	-	-	-	-	-	-
CaO	-	8	-	-	-	-	-	7	-	-	-	-	-	-
Na-acetate	5	-	-	-	-	-	-	-	-	1	-	-	-	-
Na-formate	5	-	-	-	-	-	-	-	-	-	-	-	-	-
Na-oxalate	9	-	-	-	-	-	-	-	-	-	-	1	-	-
Na_2CO_3	10	-	-	-	-	-	-	-	-	-	-	-	-	-
$NaHCO_3$	7	-	-	-	-	-	-	-	-	-	-	-	-	-
Fe_2O_3	-	-	-	10	-		-	-	-	-	-	-	-	-
NaCl	-	-	-	-	-	9	-	-	-	-	-	-	1	-
NH_4NO_3	-	-	-	-	-	-	9	-	-	-	-	-	-	-
$(NH_4)_2SO_4$	-	-	10	-	-	-	-	-	-	-	-	-	-	-
NH_4HSO_4	-	-	3	-	-	-	-	-	-	-	-	-	-	-
Total	36	18	13	10	10	9	9	7	6	3	2	1	1	1
SE	6	4	4	3	3	3	3	3	2	2	1	1	1	1

Table 2: Classification of the test particulate standard mixture using HCA, SE denotes the sampling error connected to each class

(c) Based on concentration values, normalized to the total sum, not standardized

	Classified to group No:													
Particle type	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
Algae	-	-	-	-	1	-	-	-	9	-	-	-	-	-
Lichen	-	-	-	-	10	-	-	-	-	-	-	-	1	-
$CaCO_3$	-	-	-	4	-	-	-	-	-	6	-	-	-	-
CaO	-	-	-	9	-	-	-	-	-	3	3	-	-	-
Na-acetate	-	2	-	-	-	-	-	-	-	-	-	3	1	-
Na-formate	1	4	-	-	-	-	-	-	-	-	-	-	-	-
Na-oxalate	1	8	-	-	-	-	-	-	-	-	-	-	-	1
Na_2CO_3	10	-	-	-	-	-	-	-	-	-	-	-	-	-
$NaHCO_3$	7	-	-	-	-	-	-	-	-	-	-	-	-	-
Fe_2O_3	-	-	-	-	-	10	-	-	-	-	-	-	-	-
NaCl	-	-	-	-	-	-	10	-	-	-	-	-	-	-
NH_4NO_3	-	-	-	-	-	-	-	9	-	-	-	-	-	-
$(NH_4)_2SO_4$	-	-	10	-	-	-	-	-	-	-	-	-	-	-
NH_4HSO_4	-	-	3	-	-	-	-	-	-	-	-	-	-	-
Total	18	14	13	13	11	10	10	9	9	9	3	3	2	1
SE	4	4	4	4	3	3	3	3	3	3	2	2	1	1

(d) Based on square root intensity values, normalized to the maximum, not standardized

	Classified to group No:													
Particle type	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
Algae	-	-	-	1	-	-	-	9	-	-	-	-	-	-
Lichen	-	-	-	11	-	-	-	-	-	-	-	-	-	-
$CaCO_3$	-	10	-	-	-	-	-	-	-	-	-	-	-	-
CaO	-	5	-	-	-	-	-	-	9	-	-	-	-	1
Na-acetate	2	-	-	-	-	-	-	-	-	-	3	-	1	-
Na-formate	5	-	-	-	-	-	-	-	-	-	-	-	-	-
Na-oxalate	9	-	-	-	-	-	-	-	-	-	-	1	-	-
Na_2CO_3	8	-	-	-	-	-	-	-	-	2	-	-	-	-
$NaHCO_3$	-	-	-	-	-	-	-	-	-	7	-	-	-	-
Fe_2O_3	-	-	-	-	-	10	-	-	-	-	-	-	-	-
NaCl	-	-	-	-	10	-	-	-	-	-	-	-	-	-
NH_4NO_3	-	-	-	-	-	-	9	-	-	-	-	-	-	-
$(NH_4)_2SO_4$	-	-	10	-	-	-	-	-	-	-	-	-	-	-
NH_4HSO_4	-	-	3	-	-	-	-	-	-	-	-	-	-	-
Total	24	15	13	12	10	10	9	9	9	9	3	1	1	1
SE	5	4	4	3	3	3	3	3	3	3	2	1	1	1

Table 2: Classification of the test particulate standard mixture using HCA, SE denotes the sampling error connected to each class

(e) Based on intensity values, normalized to the maximum, not standardized

	Classified to group No:													
Particle type	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
Algae	10	-	-	-	-	-	-	-	-	-	-	-	-	-
Lichen	11	-	-	-	-	-	-	-	-	-	-	-	-	-
$CaCO_3$	-	-	-	-	-	-	-	6	-	4	-	-	-	-
CaO	-	-	-	-	-	-	-	3	-	3	6	3	-	-
Na-acetate	1	-	-	-	-	-	-	-	2	-	-	-	3	-
Na-formate	-	3	-	-	-	-	-	-	-	-	-	-	-	-
Na-oxalate	-	5	-	-	-	-	-	-	6	-	-	-	-	1
Na_2CO_3	-	8	-	-	-	-	2	-	-	-	-	-	-	-
$NaHCO_3$	-	-	-	-	-	-	7	-	-	-	-	-	-	-
Fe_2O_3	-	-	-	-	10	-	-	-	-	-	-	-	-	-
NaCl	-	-	-	10	-	-	-	-	-	-	-	-	-	-
NH_4NO_3	-	-	-	-	-	9	-	-	-	-	-	-	-	-
$(NH_4)_2SO_4$	-	-	10	-	-	-	-	-	-	-	-	-	-	-
NH_4HSO_4	-	-	3	-	-	-	-	-	-	-	-	-	-	-
Total	22	16	13	10	10	9	9	9	8	7	6	3	3	1
SE	5	4	4	3	3	3	3	3	3	3	2	2	2	1

Calcium carbonate and *calcium oxide* were the two calcium containing particle types studied. Calcium oxide is very likely to take up water vapor and to react with carbon dioxide present in the air, to form calcium carbonate. The average carbon concentration of this particle type is around 7 %, as shown in Table 1, indicating that the reaction occurs, but not in every particle or not completely. Therefore, three groups were usually formed from these two particle types using HCA. The normalization does not induce any difference in the formation of the clusters. Using concentration or intensity values, the CaO and $CaCO_3$ particles are mixed in the first two calcium-containing groups (Classes 4 and 11 in Table 2a and 2c; Classes 8 and 10 in Table 2e). The most relevant clusters were formed from these particle types using the square-root scaling, where $CaCO_3$ and totally reacted CaO formed one group. The other calcium-containing groups are formed only from CaO particles (Classes 2, 9 and 14 in Table 2d).

The most difficult task was the classification of the particle types containing Na, C, and O (and H). When concentration values normalized to the maximum were used for HCA, three major classes were formed from the originally five particle types, mainly consisting of Na-oxalate, Na-carbonate and Na-bicarbonate (Classes 1, 7 and 10 in Table 2a). The Na-acetate and Na-formate particles were classified to one of these groups. This normalization yielded more relevant results than the normalization to the total sum of concentrations, since in that case only two major groups were formed. For these types of particles the intensity values seem to be better input values for HCA than the square root scaled intensities, since almost all of the particles were classified together in the latter case (see Table 2d and 2e).

It was previously shown that for conventional EPMA, the classification of the particles was indifferent either using the normalized X-ray intensities or the calculated concentrations as input data.[15] However, due to the more pronounced matrix and geometric effects for low-*Z* elements, the quantification of each individual particle is necessary prior to their classification. The use of the square root scaled intensities yields most of the cases more relevant classification results than the intensities themselves.

- **Non-Hierarchical clustering analysis**

The standard data set was also processed by NHCA using the MASLOC algorithm; NHCA is an adapted form of HCA that reduces the effect of imposing a strict and hierarchical pattern upon the analyzed dataset.[9] The results obtained for concentration values normalized to the maximum are summarized in Table 3, which shows that there is no significant difference between the classes obtained by HCA and NHCA; they are identical for standardized data and within the sampling error for unstandardized data. This confirms previous research with conventional EPMA.[5]

- **Principal component analysis**

PCA was also applied for the standard data set, as a display method to visualize the data set in the most important dimensions. When the input values are not subjected to standardization using the *z*-transform, the first three components describe more than 90 % of the total variance of the data set. However, if the data are subjected to *z*-transform prior to PCA, at least nine components are requested to fulfill the above condition. Therefore if PCA is used as a dimension reduction method, standardization is not recommended. As an example, the scores of the first three components are plotted in Figure 3, where the input data were concentration values normalized to the maximum. The visualization of the data in the reduced dimensional space can be a basis for the application of supervised classification methods.[14]

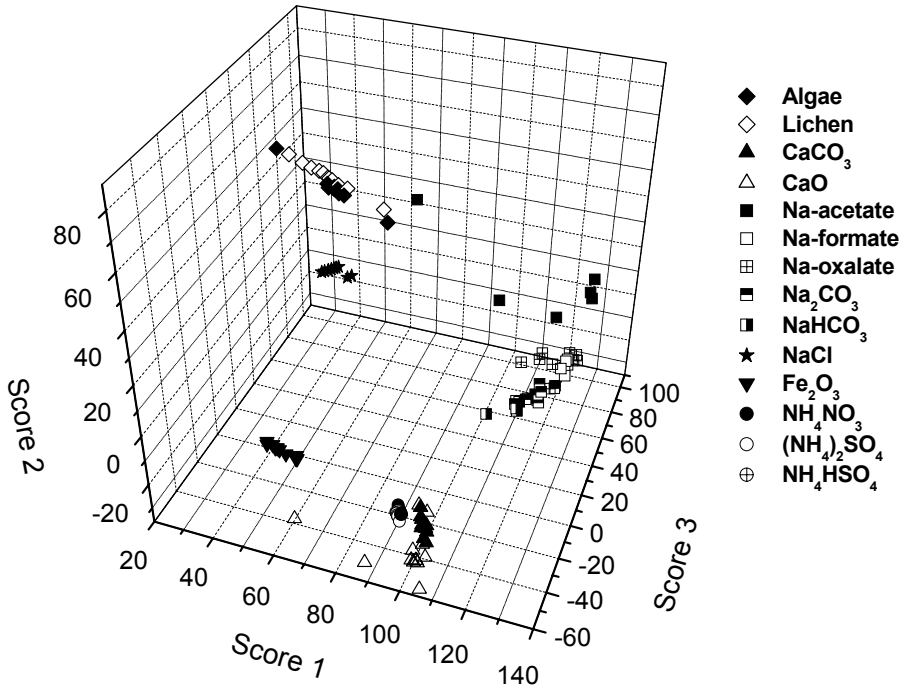

Figure 3: Classification of the test particulate standard mixture using PCA

Table 3: Classification of the test particulate standard mixture using NHCA, SE denotes the sampling error connected to each class

(a) Based on concentration values, normalized to the maximum, not standardized

	Classified to group No:													
Particle type	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
Algae	-	-	-	1	-	-	-	9	-	-	-	-	-	-
Lichen	-	-	-	11	-	-	-	-	-	-	-	-	-	-
$CaCO_3$	-	9	-	-	-	-	-	-	-	1	-	-	-	-
CaO	-	4	-	-	-	-	-	-	-	8	-	-	3	-
Na-acetate	-	-	-	-	-	-	-	1	-	-	1	4	-	-
Na-formate	-	-	-	-	5	-	-	-	-	-	-	-	-	-
Na-oxalate	-	-	-	-	5	-	-	-	-	-	4	-	-	1
Na_2CO_3	8	-	-	-	2	-	-	-	-	-	-	-	-	-
$NaHCO_3$	7	-	-	-	-	-	-	-	-	-	-	-	-	-
Fe_2O_3	-	-	-	-	-	10	-	-	-	-	-	-	-	-
NaCl	-	-	-	-	-	-	10	-	-	-	-	-	-	-
NH_4NO_3	-	-	-	-	-	-	-	-	9	-	-	-	-	-
$(NH_4)_2SO_4$	-	-	10	-	-	-	-	-	-	-	-	-	-	-
NH_4HSO_4	-	-	3	-	-	-	-	-	-	-	-	-	-	-
Total	15	13	13	13	12	10	10	10	9	9	5	4	3	1
SE	4	4	4	4	3	3	3	3	3	3	2	2	2	1

(b) Based on concentration values, normalized to the maximum, standardized

	Classified to group No:													
Particle type	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
Algae	-	-	-	-	2	-	-	-	5	1	1	-	-	1
Lichen	-	-	-	-	8	-	-	-	-	2	1	-	-	-
$CaCO_3$	-	10	-	-	-	-	-	-	-	-	-	-	-	-
CaO	-	8	-	-	-	-	-	7	-	-	-	-	-	-
Na-acetate	5	-	-	-	-	-	-	-	-	1	-	-	-	-
Na-formate	5	-	-	-	-	-	-	-	-	-	-	-	-	-
Na-oxalate	9	-	-	-	-	-	-	-	-	-	-	1	-	-
Na_2CO_3	10	-	-	-	-	-	-	-	-	-	-	-	-	-
$NaHCO_3$	7	-	-	-	-	-	-	-	-	-	-	-	-	-
Fe_2O_3	-	-	-	10	-		-	-	-	-	-	-	-	-
NaCl	-	-	-	-	-	9	-	-	-	-	-	-	1	-
NH_4NO_3	-	-	-	-	-	-	9	-	-	-	-	-	-	-
$(NH_4)_2SO_4$	-	-	10	-	-	-	-	-	-	-	-	-	-	-
NH_4HSO_4	-	-	3	-	-	-	-	-	-	-	-	-	-	-
Total	36	18	13	10	10	9	9	7	5	4	2	1	1	1
SE	6	4	4	3	3	3	3	3	2	2	1	1	1	1

- **Conclusions**

Methods like HCA, NHCA and PCA are considered to be 'unsupervised', since the basic part of the involved procedures consists of mathematical algorithms without intervention by the analyst. Therefore, we had to check the different mathematical aspects of the techniques in order to see which algorithms and options lead to a better data interpretation. As can be seen from the results, clustering and principle component analysis are efficient tools for the straightforward classification of particles based on low-*Z* EPMA data. As the technique is capable of detecting only major and minor elements, standardization using *z*-transform is not recommended prior to applying cluster or principal component analysis. Cluster analysis appears to give more relevant results when using calculated concentration values that are normalized to the maximum, which illustrates again importance of our iterative Monte Carlo quantification method. Standard compounds that are similar to each other, were not always classified to separate groups, but the results were acceptable, since the original X-ray spectra also had poor counting statistics. After PCA without standardization, the visualization of the data in the reduced dimensional space could be a basis for the application of supervised classification methods.

During the work carried out for this doctoral thesis, (N)HCA and PCA were often applied to datasets resulting from the analysis of real samples containing atmospheric particles. Examples of applications will be given in the next chapters, and it will be shown that these methods are not only successful in case of standard particles, but that their application on TW-EPMA data is able to extend our knowledge about environmental particles. However, we think it is necessary to comment on the designation 'unsupervised'. Although the tested methods indeed appear to consist of mathematical algorithms that only need some limited input from the user and will do the rest of the job on their own, reality is rather different. For example, in case of clustering analysis, the user is obliged to follow-up the complete clustering procedure, since he has to decide to which level the clustering has to continue. After all, a clustering procedure is bound to end up with only one cluster containing all objects, if the user doesn't stop the algorithm at the desired number of clusters. Although some mathematical 'stopping rules' were developed to determine the ideal number of clusters based on entropy calculations, it is our experience that in the end it is still the user that has to decide if a clustering is meaningful or not. For example, we often found that aluminosilicate clusters in atmospheric aerosols were split up into many different subclusters, though we would just combine them into one group as 'aluminosilicates' in the end. Often we also experienced that very interesting clusters, obviously consisting of different particle types, could not be split up unless ending up in a huge number of shattered clusters, which had to be combined again. This is of course due to the mathematical nature of the clustering, which does not take into account the chemical or environmental insights that the analyst has. These insights are also required for the identification of the clusters. In the experiments above, we knew which particles to expect, since we had prior knowledge of their compositions. In real environmental samples, the particle compositions are unknown. In the experiment, we also worked with pure particles, while real particles are often mixtures of different compounds. So, it is clear that the cluster identification also requires some interpretation by the analyst. In practice, this is done by using identification tables that are built up from experiences in the past (especially from conventional SEM of course). Whenever the analyst encounters a new and interesting particle type for his specific kind of samples, he puts the typical characteristic composition in the table. When the average composition of a cluster corresponds to one of the types in the identification table, it is easily recognized and identified. The more experience the analyst has, the more diverse his identification tables will be.

Table 4: Example of an extensive identification table (for conventional SEM of atmospheric particles)

CODE	*TYPE*	*CRITERIA (based on relative X-ray intensities)*	
	SOIL DUST/Si-PARTICLES		
B00	Si-rich	60 < Si	(Al, Fe, Ca, S) < 10
B01	Si-Al-rich	60 < Si + Al	10 < (Al, Si) < 60
B02	Si-Fe-rich	60 < Si + Fe	10 < (Fe, Si) < 60
B03	Si-Ca-rich	60 < Si + Ca	10 < (Ca, Si) < 60
B04	Si-Cl-rich	60 < Si + Cl	10 < (Cl, Si) < 60
B05	Si-S-rich	60 < Si + S	10 < (S, Si) < 60
B+	B-combinations	60 < Si + Al + Fe + Ca + S + Cl	10 < (S, Cl, Ca, Fe, Al, Si) < 60
	Ca/Mg-PARTICLES		
C00	(Ca/Mg)-rich	60 < Ca + Mg	(Si, S, P, Cl) < 10
C01	(Ca/Mg)-S-rich	80 < Ca + Mg + S	Si < 10; 10 < S ≤ (Ca/Mg) < 80
C02	(Ca/Mg)-P-rich	80 < Ca + Mg + P	Si < 10; 10 < P ≤ (Ca/Mg) < 80
C03	(Ca/Mg)-Cl-rich	80 < Ca + Mg + Cl	Si < 10; 10 < Cl ≤ (Ca/Mg) < 80
C+	C-combinations	Combinations of types C1-3	
	METAL PARTICLES		
M01	Al-enriched	20 < Al	Si < 15; (other metals) < 20
M02	Fe-enriched	20 < Fe	Si < 15; (othermetals) < 20
M03	Ti-enriched	20 < Ti	(other metals) < 20
M04	Cr-enriched	20 < Cr	(other metals) < 20
M05	Zn-enriched	20 < Zn	(other metals) < 20
M06	Cu-enriched	20 < Cu	(other metals) < 20
M07	Ni-enriched	20 < Ni	(other metals) < 20
M08	Cd-enriched	20 < Cd	(other metals) < 20
M09	Sn-enriched	20 < Sn	(other metals) < 20
M10	Mn-enriched	20 < Mn	(other metals) < 20
M11	Pb-enriched	20 < Pb	(other metals) < 20
M12	Ba-enriched	20 < Ba	(other metals) < 20
M+	M-combinations	Combinations of types M1-12	
	LOW-Z PARTICLES		
L0	C/N/O-particles	Relatively low total sum of X-ray intensities; (...) < 20	
L1	P-rich	60 < P	Ca < 20
L2	K-rich	60 < K	S < 20
L3	Na-rich	60 < Na	(S, Cl) < 10
L4	Cl-rich	60 < Cl	(Na, S, Ca/Mg) < 10; metals < 20
L5	S-rich	60 < S	(K, Na, Cl, Si, Ca/Mg) < 10; metals < 20
L6	Cl-Na-rich	60 < Cl + Na	10 < (Na, Cl) < 60; Si < 15
L7	Cl-Na-S-rich	80 < Cl + Na + S	10 < (S, Na, Cl) < 80
L8	S-Na-rich	60 < S + Na	10 < (Na, S) < 60
L9	S-Cl-rich	60 < S + Cl	10 < (Cl, S) < 60
L10	S-K-rich	60 < S + K	10 < (K, S) < 60
L+	L-combinations	60 < P + K + Na + Cl + S	(P, K, Na, Cl, S) < 50
	INDEFINABLE		
I	Indefinable	Particle cluster is very difficult to split up in identifiable clusters	

In the past, experiments were also carried out to test the use of fuzzy clustering analysis (FCA). Fuzzy clustering goes one step further than non-hierarchical clustering to avoid imposing certain structures on the analyzed dataset. This is done by assigning each object (particle) to more than one cluster. For all particles, membership coefficients are determined to express the degree in which they belong to the different clusters that are obtained. The purpose is to avoid unidentifiable or overlapping clusters, but optimal cluster classification is never guaranteed. In fact, it is often difficult to graphically represent the final result in an orderly way that allows clear interpretation.[5,9]

2.2 Expert identification followed by classification

- **Introduction**

As we concluded in the previous paragraph, the main disadvantage of clustering is to be found in its biggest advantage. As it is a powerful tool for mathematically compressing data to surveyable and understandable clusters, much detailed information is lost. Possible insights that might be gained on the level of individual particles, get lost in crowded clusters for which only average concentrations are obtained. For example, sea salt particles contain, of course, large amounts of sodium and chloride, so that they are often easy to cluster. However, the sea salt in each particle might be agglomerated with different other types of compounds. For understanding the chemistry behind the formation of particles, the obvious identification of a sea salt cluster is of little value. On the other hand, the fact that e.g. 1% of those particles would contain relatively high concentrations of lead, might reveal a lot about the origins of the particles within the sampled aerosol. Clustering is not always able to reveal such details, unless the analyst specifically looks for them. Therefore, although clustering is indeed a nice tool for the fast data processing of numerous particles into neatly arranged cluster tables, it might be worthwhile and more revealing to unravel the composition of each particle individually as well. It was our next objective to develop methods for processing datasets in such a way.

- **Supervised bookkeeping method for 'exact' speciation of each particle separately**

A first method was based on figuring out the composition of each individual particle 'by hand', i.e. the analyst would take a closer look at the elemental composition of each particle to find clues about the presence of certain compounds.[16] Consider some examples of aged sea salt particles, given in the table below.

Table 5: Example of bookkeeping method

	Composition (%at)		
Element	**Particle A**	**Particle B**	**Particle C**
N	0.0	7.9	11.5
O	12.3	21.8	59.5
Na	45.0	37.7	19.3
S	3.1	0.0	9.7
Cl	38.6	32.7	0.0

From our knowledge of aerosol chemistry, it is logical to assume that particle A only contains sodium chloride and sodium sulfate. The small number of elements (O, Na, S, Cl) does not suggest the presence of any other compound. Then the next question would be: what is the relative concentration of these compounds in the particle? The answer to this question can be found by simple deduction. If chloride is only coming from NaCl, a corresponding amount of sodium should part of the sodium concentration. If 3.2% of sulfur is present, two times as much sodium should be present as well. Based on stoichiometry, the sodium concentrations could be split up as follows: 2x3.1% + 38.6% = 44.8%. This closely matches the sodium concentration of 45%. If all of the oxygen belongs to sodium sulfate, as we assume, its concentration should correspond to four times the sulfur concentration, namely 4x 3.1% = 12.4%, which again closely matches the real concentration. So, the presence of different compounds in the particle could indirectly be confirmed by simple bookkeeping. In fact, from these data, we could even calculate the relative concentrations (86.2% NaCl vs. 13.8% Na_2SO_4).

Particle B can be treated similarly, but it contains nitrogen instead of sulfur. The problem now is that we could have three compounds present: sodium chloride, sodium nitrate and ammonium nitrate. Ammonium chloride is not considered, since we know from our experience that is quite volatile and that, even if it were present originally, we could not have measured it anyway due to evaporation losses in the microscope. So, again we start by searching for elements that only occur in one compound only. Chloride is linked to sodium chloride, as should 32.7% of the sodium concentration. The remaining 5.0% of sodium should be linked to the other elements in sodium nitrate, namely 5.0% of nitrogen and 15.0% of oxygen. The remaining 2.9% of nitrogen is assumed to come from ammonium nitrate, which corresponds to 4.4% of oxygen. We still have 2.4% of oxygen left, which can not be assigned to any other compound, except for water. Hydrogen does not produce X-rays under electron bombardment, so it is not detected in our spectra. Water can therefore only be detected through its contribution to the oxygen concentration. Water could either be present as crystal water in certain compounds, or as water adsorbed on the particle.
Although chloride is absent, the composition of particle C is more complex, since both sulfur and nitrogen are present. Therefore, the number of possible compounds increases; on top of all the previously mentioned compounds, ammonium sulfate could be present as well. Even worse is the fact that all elements are present in more than one of the possible compounds, and so there is no easy starting point to resolve the problem. Relative abundances can only be calculated by solving a series of equations.

Table 6: Result of the bookkeeping example

	Relative abundance (%)		
Compound	**Particle A**	**Particle B**	**Particle C**
NaCl	86.2	65.3	0.0
Na_2SO_4	13.8	0.0	50.1
$NaNO_3$	0.0	44.9	24.9
$(NH_4)_2SO_4$	0.0	0.0	17.6
NH_4NO_3	0.0	7.3	3.8
H_2O	0.0	2.9	3.6

This exact, supervised method really showed to be very time-consuming and impossible to apply on large numbers of particles. Taking into account that most atmospheric particles are in fact internal mixtures of many different compounds, it was quite frustrating to see that many particle compositions were difficult to solve in a simple way. The three particles above have only very simple compositions (only five elements are present) and they didn't even contain carbon or heavy metals. It should also be pointed out that the speciation of organic compounds cannot be performed, since hydrogen is not detected. It would be impossible to determine which types of organics are present, let alone to calculate their abundance. The same can be said for soil dust particles, like aluminosilicates that could be present in many different compositions.
The 'exactness' of the calculated speciation should be put in the right spotlight as well, since the calculations are mostly based on assumptions and on trial-and-error. For example, we did not consider bi-compounds like $(NH_4)HSO_4$. Based on our experience in atmospheric chemistry, we could of course omit this compound from the calculations, since the abundance of these compounds is very low. Secondly, the fact that our quantification method is only semi-quantitative, makes speciation also look less exact. Since this supervised method appeared to be rather an exercise in patience based on time-consuming iterations, efforts were undertaken to turn it into an automated, computerized and unsupervised technique.

- **Unsupervised expert identification**

After realizing that 'exact' speciation was out of the question for all particles, we soon tried out another approach that could be seen as the reverse of the one used in clustering analysis. In clustering procedures, the first part consists of grouping similar particles (classification), followed by identification of the clusters based on the expertise of the analyst. This is done by setting up identification rules in an easy-to-use table, as was shown before. The reverse approach would be to thoroughly apply rules like this to the individual particles first, and then to combine similarly identified particles in groups (classes or clusters). The identification rules depend of course on the experience or expertise of the analyst, but once they are created, it is possible to automate the algorithm.[17]
The difficulty is to find rules for obtaining higher degrees of identification. Again we could give the example of sea salt particles that reacted or agglomerated with particles of different compounds. This expert system would not be valuable if it were only able to identify these particles as 'sea salt'. It should be able to distinguish the possible compounds inside each particle, but it should also be able to group particles of similar composition in a comprehensive way. Based on the experience and the objectives of the analyst in his particular field (e.g. sediment analysis, urban particulate matter, biogenic particles, etc.), it is possible to extend the identification to the desired level. For example, below we show the typical codes that were used to identify particle groups obtained after the analysis of samples in the PM2.5 project (see chapter 5).

Table 7: Identification table for TW-EPMA of PM2.5 samples

CODE	TYPE
Org	Organic particles
Org+N+O	Organic + possibly ammonium nitrate particles
Org+N+S+O	Organic + possibly ammonium sulfate particles
Org+N+S+Cl	Organic or biogenic particles
Org+Na(+Cl)+(N+O/S+O)	Complex mixtures of sodium and ammonium salts (+ sodium chloride)
Org+K	Biogenic + combusted biomass particles
Org+Al+Si	Aluminosilicates, containing some organics
Org+Fe+S	Particles containing iron/sulfur/organic
N+S+O	Ammonium salts
Na+(N+O/S+O)	Sodium nitrate/sulfate (aged sea salt) particles
Na+Cl+(N+O/S+O)	Sodium chloride + nitrate, sulfate (aging sea salt) particles
K+N+O	Potassium nitrate particles
Al+Si	Aluminosilicates
Fe+O	Iron oxide particles
Ca+S/C/P+O	Calcium sulfate/phosphate/carbonate particles

We would like to point out that we did introduce some specific compounds into the identification, as we tried in the rigorous bookkeeping method, but without pretending to know exactly to which extent they really occur. Some assumptions are of course introduced, but always based on the experience of the analyst. On the other hand, it should however be clear that this identification table is still not very extensive. Although it was used for identifying clusters obtained by NHCA in chapter 5, it was not useful for the method that we tried to develop, since it was our goal to offer a thorough analysis at the individual particle level. A typical highly detailed identification table is shown below, as developed and published by Ro *et al.* (2004) for the expert analysis of Asian dust samples.[18]

Table 8: Identification table for TW-EPMA of Asian dust samples [18]

CODE	PARTICLE TYPE
C	carbon-rich
C-O	organic or carbon-rich
C-O-Ca	inorganic salts or aluminosilicate and Ca salt mixture
C-O-N	organic or/and inorganic salts
C-O-Na	organic or/and Na_2CO_3
C-O-Si	organic or/and aluminosilicate
O-Al	aluminosilicate or/and Al_2O_3
O-C	organic
O-C-Ca	inorganic salts or aluminosilicate and Ca salt mixture
O-C-Cl	inorganic salts
O-C-Fe	aluminosilicate or/and Fe_2O_3
O-C-K	inorganic salts or aluminosilicate and K salt mixture
O-C-Mg	inorganic salts
O-C-N	organic
O-C-N-S	ammonium salt
O-C-N-Ca	inorganic salts or/and aluminosilicate
O-C-Na	inorganic salts or/and aluminosilicate
O-C-S	inorganic salts or/and aluminosilicate
O-C-Si	aluminosilicate
O-C-Si-Mg	aluminosilicate and inorganic salts mixture
O-C-Si-Ca	aluminosilicate and inorganic salts mixture
O-C-Si-K	aluminosilicate and inorganic salts mixture
O-C-Si-N	aluminosilicate and inorganic salts mixture
O-C-Si-Na	aluminosilicate and inorganic salts mixture
O-C-Si-S	aluminosilicate and inorganic salts mixture
O-Ca-Si	inorganic salts and aluminosilicate mixture
O-Ca-Al	inorganic salts and aluminosilicate mixture
O-Ca	inorganic salts
O-K-Si	inorganic salts and aluminosilicate mixture
O-K-Al	inorganic salts and aluminosilicate mixture
O-K	inorganic salts
O-Mg-Si	Si and Mg oxide
O-Mg	inorganic salts
O-N-S	ammonium salts
O-N-Ca	inorganic salts
O-N-K	inorganic salts
O-N-Na	inorganic salts
O-Na-C	inorganic salts
O-Na-Cl	inorganic salts
O-Na-N	inorganic salts
O-Na-S	inorganic salts
O-Si	aluminosilicate
O-Si-Na	aluminosilicate or/and inorganic salts
O-Si-Mg	aluminosilicate or/and inorganic salts
O-Si-Na-S	aluminosilicate and inorganic salts
O-Si-Na-N	aluminosilicate or/and inorganic salts
O-Si-K	aluminosilicate or/and inorganic salts
O-Si-K-S	aluminosilicate and inorganic salts
O-Si-K-N	aluminosilicate and inorganic salts
O-Fe	iron oxide
O-Mn	manganese oxide
O-Ti	titanium oxide
O-V	vanadium oxide

The level or degree of identification in Table 8 is quite high and offers great possibilities for many types of atmospheric samples. However, this table is still only part of an initial step in the expert system. The overall algorithm behind the expert system could be described as follows.

First the elemental data for all particles are pre-processed, i.e. elements with atomic concentrations >1% are sorted in decreasing order of their concentrations. Afterwards, each particle is assigned to a certain decision tree according to its most abundant element. The other elements are used for distributing the particles over lower level trees, in order of decreasing concentration. The obtained distribution for a particle results in the assignment to a typical elemental pattern as shown in Table 8. So, the decision-making trees are comparable to the identification rules used in clustering procedures, like we showed in

Table 4 for cluster identification in case of particles analyzed by conventional SEM. Most particles require the use of three or four of their elements to assign them to a certain pattern. In the next step of the procedure, a knowledge database for performing chemical speciation is used for linking the pattern to possible inorganic compounds that could occur in the particle. Similar to the bookkeeping method, relative abundances are be calculated for different possible compounds in the particle. In fact the whole algorithm mimics the thinking process of an experienced analyst trying to work out the same problem.

The expert system focuses of course mainly on the inorganic species, since organic compounds are difficult to distinguish due to the lack of information on hydrogen. However, this doesn't mean that the system is not able to identify organic particles; it is even possible to separate carbon-rich and organic particles from each other based on their carbon content. As pure particles, both types are recognized if the carbon and oxygen concentration account for 90% of the concentrations, but carbon-rich particles can be distinguished by the fact that their carbon content is typically three times higher than their oxygen content. The inorganic particles can be divided into two groups, namely the aluminosilicates and inorganic salts. The aluminosilicate species consist of mainly aluminium and silicon oxides often with one or more minor elements (like Na, K, Mg, Fe, Ca, S, Cl) and some transition metals. The inorganic salts include carbonate, sulfate, nitrate and chloride, in combination with the different cations like sodium, ammonium, magnesium, calcium and potassium.

The method was first tested on nearly the same set of pure, standard particles as the one that was used for testing the statistical methods. From the results in Table 9, it is clear that the method is indeed able to automatically identify the different species: 93% of all particles was correctly identified. The other 7% correspond to the same particles that were also difficult to identify using clustering analysis. Those 23 particles were checked again by their spectra, and it was clear that neither of them really corresponded to the ones for similar particles of their kind. Therefore, it is expected that these particles had a different composition, or that something went wrong during analysis.

The expert system proved to work efficiently for pure particles, so it could be tested on real samples as well. Moreover, it was time to use it for speciation. The speciation of inorganic salts is based on iterative calculations of the possible relative abundances for each compound. In some way, it should be seen as trial-and-error of calculating the concentrations of all possible compounds. The criterion for checking if a certain composition is valid or acceptable is based on a simple rule of equivalence. The amount of anions and cations should be proportional, since their charges should be in balance. The iterative calculations only stop when the calculated composition of inorganic salts meets the equivalence rule (less than 15% difference between the sums of cations and anions). Particles are identified by those compounds that occur in concentrations > 10%; a particle class is defined by the names of the compounds in order of decreasing concentration.

Table 9: Testing the expert system on pure standard particles

TYPE	*Correctly identified*	*Wrongly identified*
Algae	20	
Lichen	21	
$CaCO_3$	13	
Fe_2O_3	21	
Na_2CO_3	14	
NaCl	19	6
NH_4NO_3	18	
$(NH_4)_2SO_4$	19	1
Al_2O_3	18	4
$CaCl_2$	18	2
$CaSO_4$	20	
K_2CO_3	19	1
K_2SO_4	19	1
KCl	15	5
KNO_3	18	2
$MgSO_4$	19	1
Na_2SO_4	19	
SUM	310	23

A detailed discussion about the capabilities of speciation analysis will be given in the North Sea application described in chapter 4, but a shorter example will be given here as an illustration. Consider the particle for which the spectrum and elemental composition are given below. It was taken from an Asian dust sample in ChunCheon, South-Korea, on April 7, 2000.[19] The term 'Asian dust' or 'yellow dust' refers to the huge sand storms that occur periodically in Asia; yellow sand from deserts in Mongolia are transported by the wind over central China to southern Asia and even the USA.[20,21,22,23] The phenomenon could be compared to the transport of Sahara sand to Europe and South America, but the effect is more explicit, since it has an enormous effect on visibility.[24,25] The composition of the particles could tell us more about possible transformations of particles during long-range transport. The composition of 'yellow sand' is well-known, and so it could be compared with that of 'yellow dust' to find out which reactions occurred on the way to Korea.

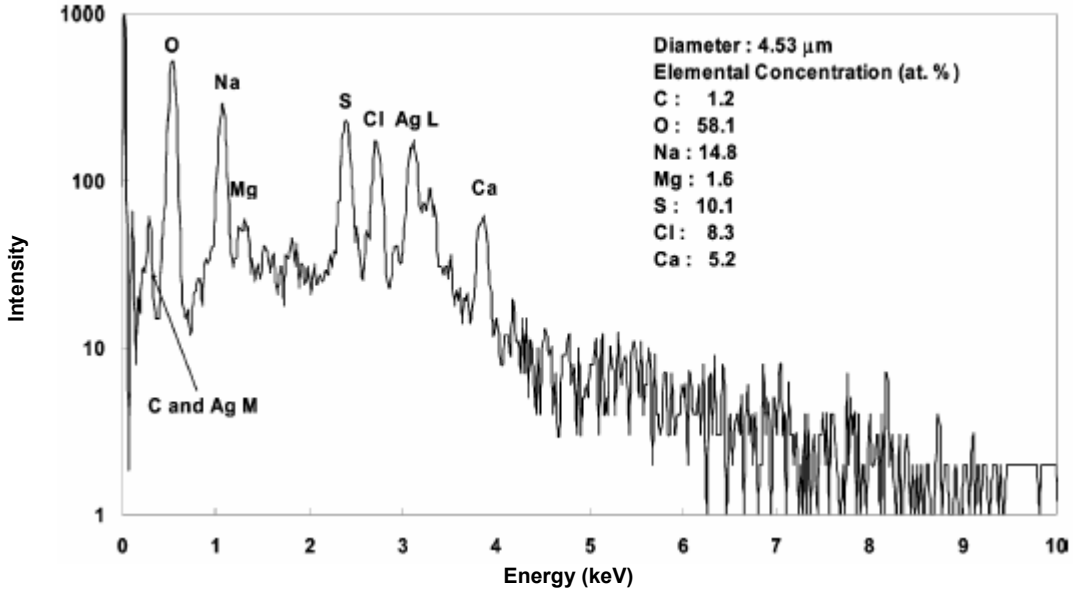

Figure 4: Example of speciation analysis

Based on the data obtained by TW-EPMA, the expert system determined that the composition of the particle is expected to be: 37% Na_2SO_4, 26% $CaSO_4$, 15% NaCl, 11% $CaCl_2$, 8% $MgSO_4$ and 3% $MgCl_2$. This composition also matches the equivalence rule, as the sum of anion compositions closely matches the one from the cations (28.5% versus 28.4% respectively; 2x10.1% + 8.3% *versus* 14.8% + 2x1.6% + 2x5.2%). Therefore, the particle is identified as Na_2SO_4-$CaSO_4$-NaCl-$CaCl_2$. Once all particles for the same sample are analyzed and identified, similar particles are joined in expert clusters. For example, this particle could be assigned to a class identified as (Na,Ca)(SO_4,Cl).

In the case of impactor samples, it could be interesting for the analyst to compare different stages. Therefore, the expert system is able to apply uniform class identifications for all stages, allowing the unambiguous interpretation of results. Additional tools make it also possible to investigate specific particle groups for other detailed properties. For example, it might be interesting to search for the presence of specific elements in aluminosilicate particles. Using the expert system, it should be possible to link the results of single particle analysis and bulk analysis in a better way, as we will shown in chapter 4.

3. Extended particle analysis

3.1 Introduction

One of the targets of criticism in the evaluation of the iterative Monte Carlo simulations is the precondition that the analyzed particles should be homogeneous in composition. It is indeed fair to assume that this is not always the case. Especially naturally grown particles in the environment can have quite complex morphology and composition compared to artificially generated or manmade particles. For example, a particle that broke off a silicon wafer and is found during contamination control in a cleanroom, will be far less complex than an aged sea salt particle that covered 50 km before ending up under the microscope.
Much research was done on the possible reaction mechanisms of microparticles with gaseous species in the air and different chemical and physical models were suggested.[26] For example, the mechanisms for the reaction with SO_2 and NO_x are well-studied and documented in literature, since they are linked to acid rain (a very popular item since the eighties).[27,28,29,30] As an example, the figure below shows a model for the transfer of SO_2 from the gas phase to the aqueous phase of an atmospheric water droplet (1 = transport to the surface of the droplet, 2 = transfer across the air-liquid interface, 3 = formation of aqueous phase equilibra, 4 = transport from the surface to the bulk region, 5 = reaction in the droplet).

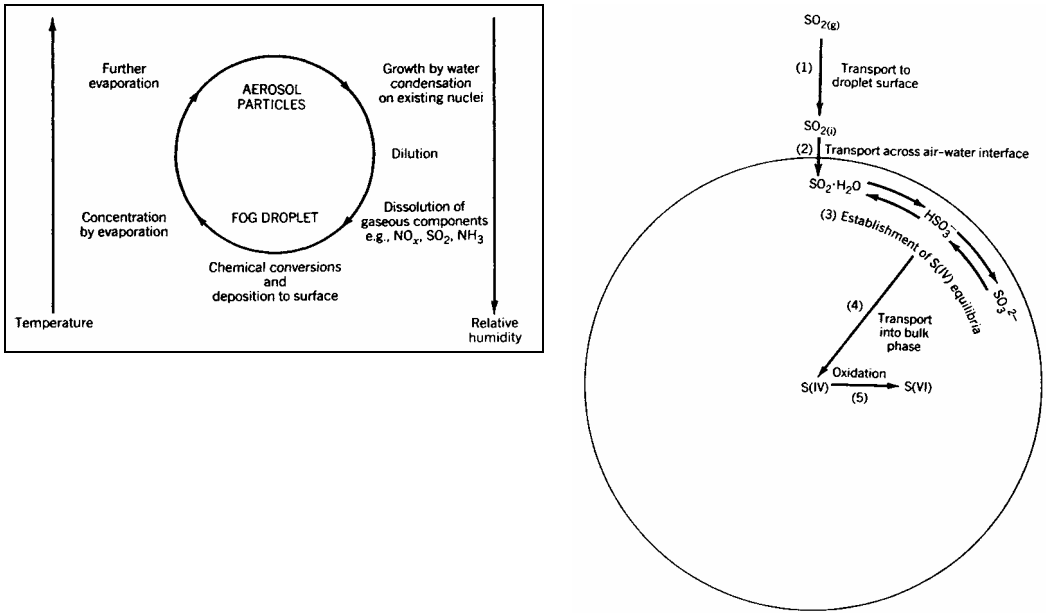

Figure 5: Schematic of steps involved in SO_2 transfer in aerosol particles and fog droplets [31,32]

The same kind of reaction mechanisms were found for reactions with solid, atmospheric particles either after adsorption water at their surface or after deliquescence at a certain relative humidity (i.e. particles turn into "droplets" of diluted salt solution of which the concentration is in equilibrium with the water vapor pressure). [33,34] Studies in controlled laboratory conditions have lead to results regarding the reaction mechanisms of many organic and inorganic particles in different gaseous environments.[35,36,37,38,39] This consequently lead to conclusions about the possible structure and morphology of reacted or mixed aerosols.[40,41,42,43,44,45] The complex morphology of environmental particles has consequences for our research, and some tools were developed to investigate the heterogeneity of a particle volume.

3.2 Grazing-Exit TW-EPMA

In 1999, Tsuji et al. developed a new method for surface or particle analysis using grazing exit detection at MiTAC.[46,47,48,49] The term 'grazing' is used in many analytical techniques to denote the use of small angle geometries for beam incidence or signal detection. When a primary beam impinges a sample, it often occurs that the signal originating from the microsurface or microparticle is overshadowed by the background signal. Therefore, it is very difficult to achieve microanalysis with sufficient analytical sensitivity.

In the field of *X-ray fluorescence* (XRF) analysis, the X-ray total reflection phenomenon is a good example of how geometry is used to achieve surface-sensitive analysis of flat substrates. In total-reflection XRF, an X-ray beam is set to impinge the flat sample substrate at grazing angles less than a certain critical angle, below which total reflection occurs. The primary X-rays are totally reflected, yielding low background intensity, since the penetration depth is restricted to several nanometers. Another method is grazing-emission XRF, in which the same effect is utilized in the process of X-ray detection. Fluorescent X-rays are measured at grazing exit angles, so that only X-rays coming from the top surface of the sample are measured. In *reflection high-energy electron diffraction* (RHEED) and in *scanning electron microscopy* (SEM), the setup of total-reflection-angle X-ray spectroscopy (TRAXS) consists of an electron beam at grazing incident angles and X-ray detection at grazing exit angles. However, the area irradiated by the electrons under grazing-incidence conditions is considerably large, so that it is impossible to perform microanalysis.

However, Tsuji et al. showed that grazing-exit EPMA could be applied for surface or particle analysis. The main principle consists of detecting only the fraction of signals that originate specifically from the region of interest within the analyzed sample. Whenever the electron beam impinges a sample, background signals will emerge due to elastic and inelastic scattering events. However, grazing-exit detection is able to 'select' only those X-rays that are of interest to us. The ideal setup would exist of a detector that is movable over a certain range, so that the exit angle between the sample surface and a straight line to the detector is close to 0°. Since most detector systems, like the one in our SEM and EPMA, are fixed to a certain angle (typically 30° or 40°), it is impossible to move the detector to 0°. Therefore, Tsuji et al. suggested the use of a brass triangular attachment having an inclination of about the same angle as the normal exit or take-off angle of the X-ray detector. In this way, the sample surface could be positioned in line with the detector by using the tilting feature of the sample stage instead of tilting the detector itself. To minimize the solid angle for X-ray detection to approximately 3.4° and to improve angular resolution, the distance between the sample and the detector was set from 30 mm to 100 mm. In addition, a slit of 0.5 mm in width was placed in front of the detector, further reducing the exposed area of the detector surface from 30 mm² to 3 mm².
Figure 6 shows the difference between the basic concept of GE-EPMA and the alternative setup that was developed at MiTAC. The main aspect of grazing-exit measurements is the position of the detector relative to the sample surface, so the detector should be movable to 0° while the primary electrons enter the sample at an incidence angle of 90°. Since moving the detector was not an option for the microscopes at MiTAC, the alternative is based on tilting the sample holder on which the sample is attached onto a triangular element with inclination of 45°. It is obvious that in this way, not only the exit angle, but also the incidence angle is changed every time the tilt angle is altered. At grazing exit, the electron beam impinges under 50°.

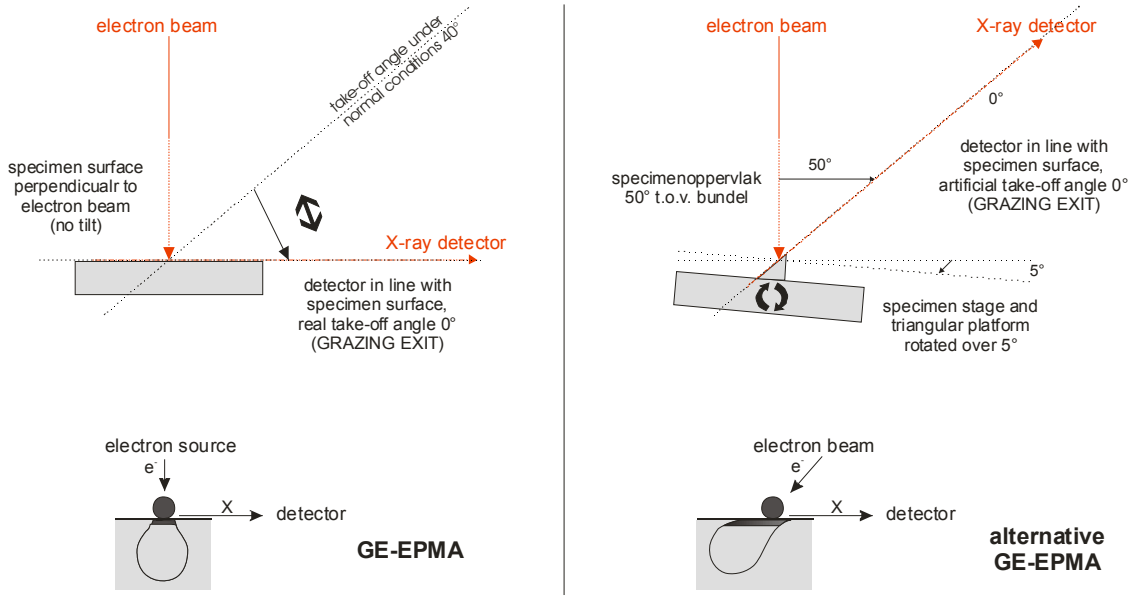

Figure 6: Principle of GE-EPMA and alternative set-up

Tsuji et al. suggested selecting very flat collection substrates like silicon wafer, because 1) the slightest deviation might deflect the emerging X-rays on their way to the detector, and 2) it would guarantee the clear and correct localization of the particles.[46] A gold coating should be deposited as a tracer on the substrate surface, prior to sampling for finding the exact grazing exit angle during analysis afterwards. In the spectra taken at different exit angles (typically from -2.5 to 5.5°), grazing conditions are found when the gold signal has disappeared.

GE-EPMA, as described above, could be a useful technique in the following cases:

1) If the particles of interest are expected to contain some of the elements that are present in the substrate they were collected on, it would be impossible to determine the concentrations of these elements by normal EPMA. In GE-EPMA, it is possible to find conditions were the substrate does not contribute to the spectrum, so that the detected signals can be clearly identified as coming from the particle.
2) When the analyzed particles are very small and their signals are too small to be distinguished from the substrate or background signal, grazing-exit conditions make sure that only the signals of the particle are considered.

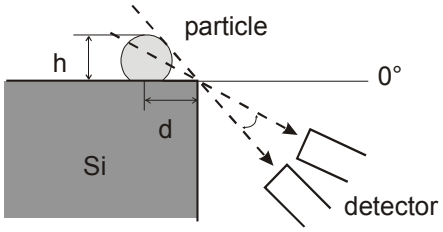

Figure 7: GE-EPMA at negative angles

GE-EPMA is also interesting in another setup. When the detector is positioned at small negative angles, it should be possible to analyze particle surfaces. At grazing exit (0°), X-rays are expected to come from the particle as a whole. When the detector is positioned at negative angles relative to the substrate surface, signals will originate from the uppermost parts of the particle. At a certain angle, we will only merely see the top layer, which makes this technique interesting to study the surface of particles.

3.3 Multiple-Voltage TW-EPMA

The idea behind Mutiple-Voltage EPMA (MV-EPMA) is to investigate heterogeneous individual particles with different primary electron beam energies.[50,51,52,53,54] Since the excitation volume decreases with decreasing beam energy, the information depth will also decrease, so that the obtained spectra will give more information on the chemical compositions of the upper particle regions. It is clear that this depth profile technique can be applied to layered particles, i.e. particles with a coating or a shell around a core. The analysis of such a particle at different accelerating voltages would result in different spectra, that each would give more information on the shell or coating layer depending on the attained information depth.

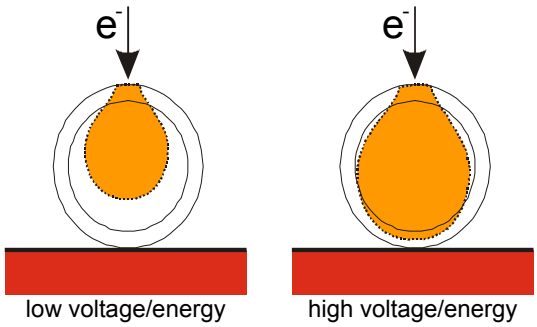

Figure 8: Principle of MV-EPMA: analysis at different voltages

MV-EPMA could be used for the following two problems:

1) If the composition of the particle is unknown, the interpretation of the obtained spectra could be used to find out which elements are present in the top layer.
2) If the composition of the particle core and shell are known, it is possible with iterative calculations to calculate the shell or layer thickness.

For most environmental particles, the first problem can only be solved in a qualitative way, because there are too many unknowns: either the layer thickness is unknown or it is very difficult to find out if the elements are present in either the shell or the core, or in both. Therefore, a quantitative determination of the elemental composition of the particle regions is impossible. Based on some assumptions or foreknowledge about the particle composition, the second question can be solved by iterative Monte Carlo simulations. As explained in the previous chapter, our simulations were adapted for layered particles (hexahedral, spherical or hemispherical).

Of course, there are also other limitations to the technique depending on the size of the particle, the thickness of the layer and the low-voltage capabilities of the instrument.

1) If the particle is very small, the particle volume will always be contained inside the excitation volume (even at the lowest voltages). The spectra will then only give more information about the substrate and not about the particle layer.
2) If the layer is very thin, its contribution to the spectra at low voltages will not be high enough to be noticed. On the other hand, if the layer is too thick so that it always contains the excitation volume, we will have no information about the core of the particle.
3) The instrument should produce high-quality spectra at all voltages. FEG-SEM will of course provide a higher quality and a broader voltage range than our probe, but we experienced that spectra down to 5 kV were satisfying enough for our analysis.

3.4 Experiments

We did some experiments on standard particles, as will be explained below. The experiments will be discussed in chronological order, and the reader will learn that we started with GE-EPMA before ending up with MV-EPMA. It will also be explained why we prefer MV-EPMA over GE-EPMA for the kind of experiments that we carried out.

- **$CaCO_3$-$CaSO_4$ and $CaCO_3$-$Ca(NO_3)_2$**

Artificial heterogeneous $CaCO_3$-$CaSO_4$ and $CaCO_3$-$Ca(NO_3)_2$ particles were synthesized, i.e. particles with $CaSO_4$ or $Ca(NO_3)_2$ in the surface region and $CaCO_3$ in the core.[55] Standard $CaCO_3$ were prepared on silicon substrate, after which they were exposed to sulfuric or nitric acid vapors. This was done by attaching the prepared substrate upside-down to the lid of a glass Petri dish filled with acid. The particles were exposed to the vapors during 1-3 minutes with nitric acid or 24-48 h with sulfuric acid. In our first experiments, we analyzed the particles by normal-exit TW-EPMA, followed by GE-EPMA using the MiTAC alternative set-up. The starting point of our experiment was to check if we could indeed generate layered particles by the proposed method (exposure to acid vapors in a Petri dish). Our first goal was therefore to check if we could find any layered particles and to see what GE-EPMA could offer us in view of depth analysis. Van Ham et al. analyzed our samples with TOF-SIMS as well to test the capabilities of this technique for the same type of analysis but then by making use of ion sputtering.
Analysis by TW-EPMA at conventional exit angles showed that many particles had completely reacted to nitrate particles in a couple of minutes due to the fast reactions of nitric acid with the carbonate. However, some particles were still found to have some remaining carbonate underneath a layer of nitrate. The exposure to sulfuric acid was less efficient, and so it took a very long time to find any layered particles (days instead of minutes). After finding the right exposure times, we could indeed observe the presence of carbonate and nitrate/sulfate in the X-ray spectra. Van Ham et al. reported that TOF-SIMS could clearly observe the presence of $CaSO_4$, but that the analysis of $Ca(NO_3)_2$ was more difficult, which appears to be in contradiction to our results. However, a reason might be found in the higher volatility of nitrate compounds. Since the applied vacuum in the TOF-SIMS instrument was higher than in our EPMA and since no cold stage was used, the nitrates might have evaporated before the actual analysis.

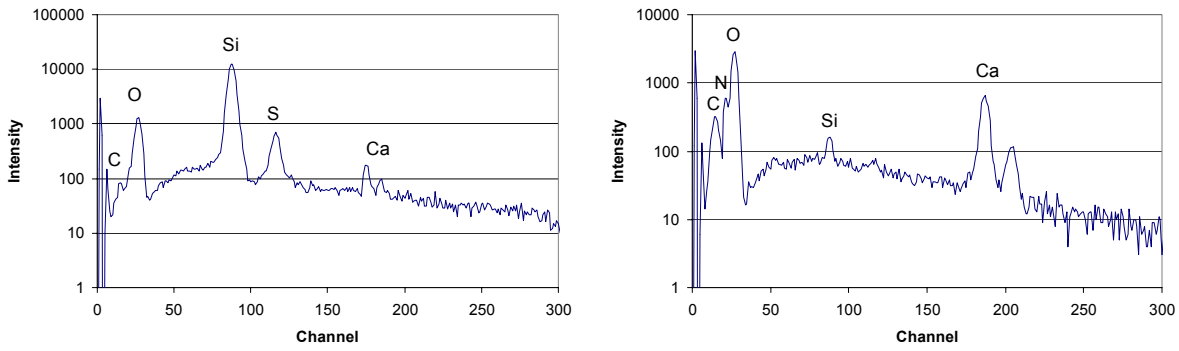

Figure 9: Spectra of typical layered particles: $CaCO_3$-$CaSO_4$ left, $CaCO_3$-$Ca(NO_3)_2$ right

Then we applied GE-EPMA: the exit angle was increased step-by-step until grazing-exit conditions were reached. In the graph below, we show the result for a $Ca(NO_3)_2$ particle; the ratio of the nitrogen and carbon intensity is plotted as a function of the exit angle.

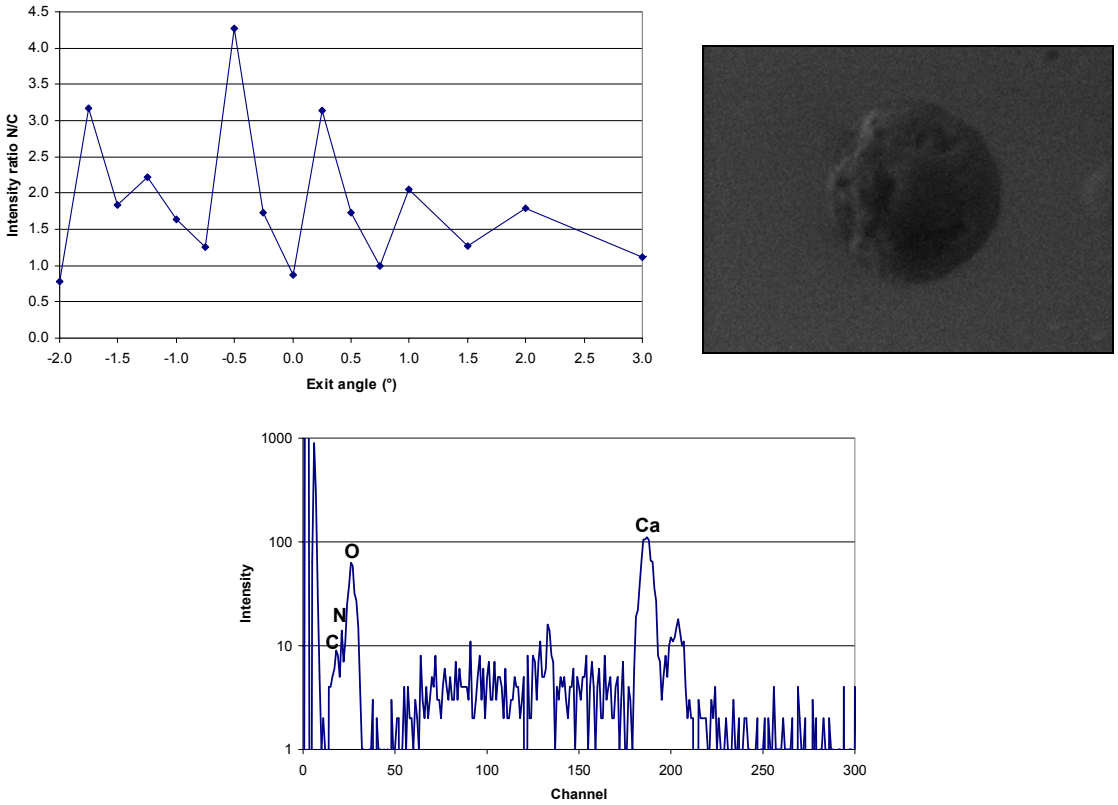

Figure 10: GE-EPMA analysis of $CaCO_3$-$Ca(NO_3)_2$ particle (5 µm, 3 µm from the substrate edge) (ratio N/C: upper left; picture: upper right; spectrum at exit angle of -0.5°: bottom)

In the case of the particle above, we see that the intensity ratio N/C fluctuates a around every 0.5°, but that a maximum is found around the exit angle of -0.5°. The fact that the intensity ratio varies a lot, can be explained by at least two reasons:

1) Nitrates (and to a lesser extent also sulfates) are known to be beam sensitive. Subsequent irradiation causes heat-up of the sample and evaporation of its volatile compounds. The cold stage is able to reduce the beam damage, since it creates a sublimation or cold-finger effect which re-attracts evaporating particle fragments to the sample on the copper triangular stage element.
2) In our alternative setup, the sample and the analyzed particle are tilted for obtaining the right exit conditions. This means that with every turn, another part of the particle is irradiated. If the particle has a rough morphology, the spectra may vary according to the obstacles which have to be passed by the escaping X-rays.

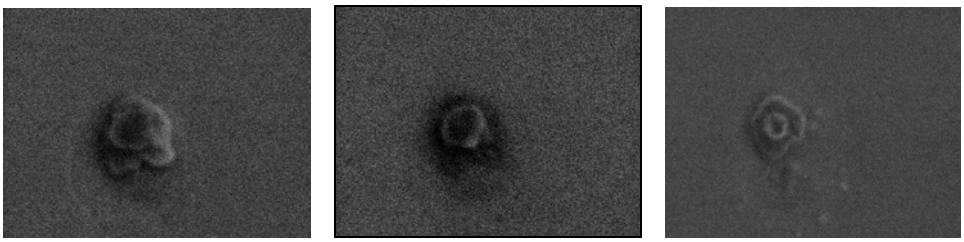

Figure 11: Beam damage to irradiated $CaCO_3$-$Ca(NO_3)_2$ particle

The peak maximum at the negative exit angle around -0.5° confirms the fact that we have produced a layered particle. At negative angles, we start to "see" the uppermost part of the particle. If the ratio increases, the carbon intensity decreases or the nitrogen intensity increases (or both), and therefore, the top region is expected to contain more nitrogen.

The fact that the ratio drops back to lower values at smaller exit angles, can be explained by the screening by the substrate. The lower the exit angle, the less we see the particle, because it is increasingly covered by the substrate edge. It is important to point out that we often found the substrate signal to slightly increase again, although it should have dropped to zero at 0°. This intensity increase is probably caused by X-rays traveling through the substrate edge.

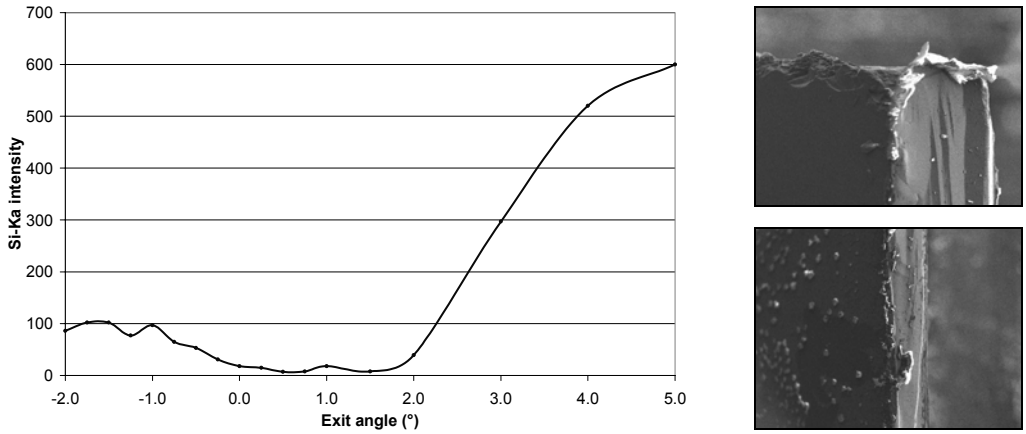

Figure 12: Intensity of silicon at different exit angles (left) and image of silicon edges (right)

During our first experiments, we quickly experienced that the alternative setup used at MiTAC has some disadvantages compared to the setup used by K. Tsuji in Japan. In the alternative setup in Antwerp, the analyst has to change the exit angle manually using the tilting feature of the sample stage. In the setup that K. Tsuji developed in Japan, the detector is moved in a motorized and more accurate way, so that this procedure can be run automatically. In Antwerp, the smallest angle step is about 0.25°, which corresponds to setting the tilt angle in-between two marks of the indicating instrument. This is quite large compared to the smallest step on the detector stage in Japan, namely 0.5 μm, which corresponds to an angle that is about 1000 times smaller ($2.7x10^{-40}$).

Another disadvantage of the alternative in Antwerp is found in that fact that the samples are moved instead of the detector. In this way, the particles are irradiated at a different spot every time the exit angle is changed. This means that the spectra of particles with rough morphology might be very difficult to interpret due to unexpected changes in X-ray intensities that cannot be linked to the grazing-exit effect.

One of the problems for GE-EPMA at negative angles is that the particles of interest should be positioned at short distances from the edge of the substrate, which is not always possible. An even bigger problem is the fact that the substrate edge should be clean and sharply cut, which seemed almost impossible to achieve, even with very fine, diamond cutting knives. As can be seen in Figure 12, the edge is often not sharp at all and it might contain particulate impurities even after cleaning the substrates with organic solvents in an ultrasonic bath. The X-rays that emerge from a particle at grazing or negative angles might be absorbed by these objects or by other interfering particles before they reach the detector.

These conclusions indicate that GE-EPMA is not so straightforward (certainly not in the setup at MiTAC), and that the method takes a lot of time. It requires patience from the analyst while carrying out the analysis itself and afterwards during the data interpretation. As can be seen in the figures above, the obtained signals are quite rough and it is difficult to make a clear interpretation of the results. GE-EPMA in our set-up at MiTAC does not really offer prospect of quantification possibilities. Therefore, we started looking for other ways to study particle heterogeneity: MV-EPMA.

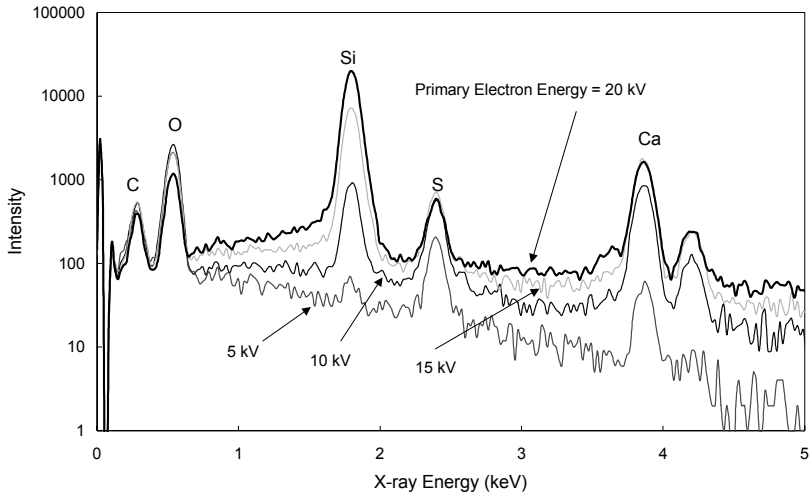

Figure 13: Spectra of a $CaCO_3$-$CaSO_4$ particle taken at different accelerating voltages

X-ray spectra, obtained at 5, 15, and 20 kV electron-accelerating voltages for a heterogeneous $CaCO_3$-$CaSO_4$ spherical particle of 1.5 μm diameter, are shown in Figure 13. The measured characteristic X-ray intensities for the elements in the particle vary differently with the variation of primary electron beam energies. From the observation of the different trends for the elements (characteristic X-ray intensity variation according to the variation of electron beam energies), these X-ray spectra certainly contain information on chemical species and heterogeneity of the particle. In Figure 14, simulated spectra calculated by our modified Monte Carlo program are shown. The similarity between the simulated and experimental spectra is remarkably obvious. The Monte Carlo calculation almost perfectly simulates the X-ray intensity variations for the elements according to the variation of the primary electron beam energies.

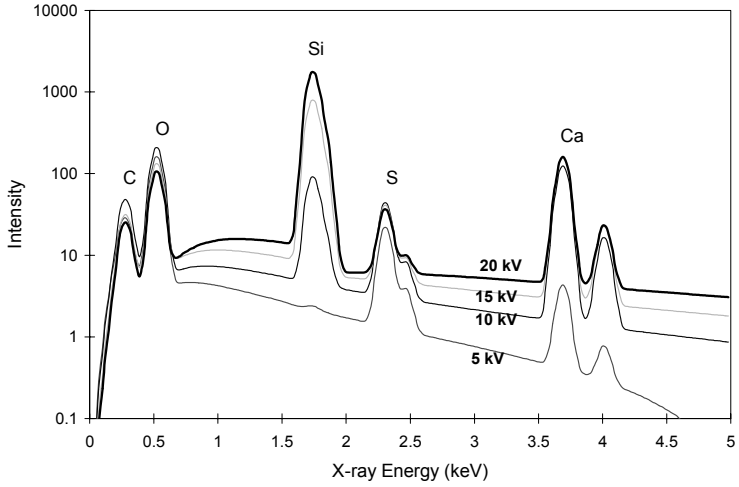

Figure 14: Simulated X-ray spectra for a $CaCO_3$-$CaSO_4$ particle at different voltages

From the spectra above it would still be difficult to make a clear interpretation of which elements are present in the shell or in the core. However, the thickness of the known $CaSO_4$ surface region of the artificially generated $CaCO_3$-$CaSO_4$ particles can be determined by Monte Carlo calculations. In Figure 15, ratios of simulated-to-measured intensities with the variation of the $CaSO_4$ surface thickness are shown for a spherical $CaCO_3$-$CaSO_4$ particle with a diameter of 1.5 μm irradiated with a primary electron beam energy of 15 keV.

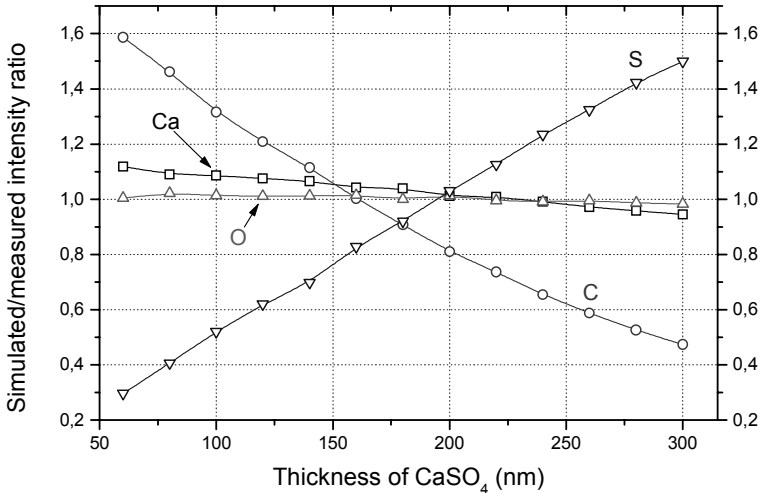

Figure 15: Simulated/measured x-ray intensity ratios versus $CaSO_4$ surface thickness

For oxygen, the ratios of simulated-to-measured intensities are relatively constant with the variation of the $CaSO_4$ thickness, mainly because of their small compositional differences between the two chemical species. However, the ratios for carbon and sulfur between the simulated and measured intensities are strongly dependent on the thickness of the surface $CaSO_4$ region. Furthermore, the ratios for sulfur decrease as the thickness of $CaSO_4$ region decreases, whereas the ratio for carbon increases as the thickness of $CaSO_4$ region decreases. For the heterogeneous $CaCO_3$-$CaSO_4$ particles, the sulfur species is in the surface region and carbon is in the core region. Therefore, if the assumed $CaSO_4$ surface thickness for the Monte Carlo calculation is thicker than the real one, then the calculated intensities are larger than the measured ones for sulfur, whereas they are smaller for carbon. From the result in Figure 15, the best match between the simulated and measured data is in the range of 160-200 nm thickness of the surface region.

- **Carbon-coated glass spheres**

The validity of our technique for estimating layer thickness and for layer applications in general was investigated more systemically using standard soda-lime glass particles (SPI #2716) coated by carbon evaporation. These particles were consecutively coated four times with carbon layers, using a carbon coater for scanning electron microscopy purposes. The estimated carbon thickness for different numbers of carbon layers obtained by the proposed method is tabulated in Table 10. The obtained results show good agreement between the values obtained at different accelerating voltages, supporting the applicability of the proposed method.

Table 10: Estimated thickness (nm) of coated carbon layers

Accelerating voltage (kV)	*Number of carbon layer depositions*			
	1	*2*	*3*	*4*
5	35	65	105	135
10	35	70	95	145
15	30	65	95	150
20	25	55	85	140

- **NaCl-$NaNO_3$-Na_2SO_4 particles**

A similar experiment for MV-EPMA was done with sodium chloride particles that were exposed to nitric and sulfuric acid vapors to form particles with a NaCl core and a $NaNO_3$ or Na_2SO_4 shell.

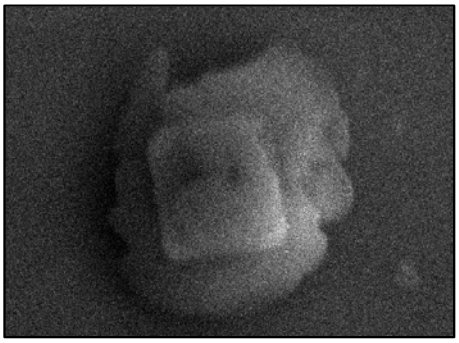

Figure 16: Sodium chloride particle after exposure to nitric acid vapors

In the figure below, spectra are shown for a sodium chloride particle that has been exposed to nitric acid vapors for a short time (5 x 7 x 1.5 μm). For calculating the thickness of the layer, measurements were done at 5, 10 and 20 kV. The spectra are shown for each voltage, and some typical simulation graphs were added for 5 and 20 kV. After iterative calculations, it was concluded that the particle layer is 0.2 μm thick. This result is plausible, but could of course not be confirmed.

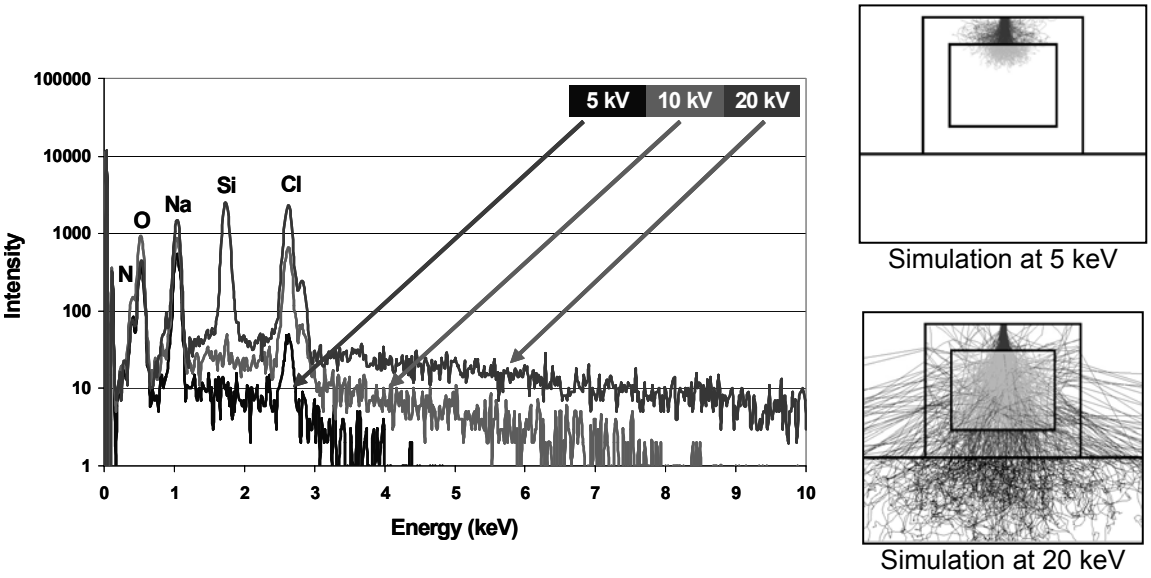

Figure 17: Thickness analysis of NaCl-$NaNO_3$ particle: spectra left, simulations right

In interesting problem in view of quantification is the fact that we don't know during the analysis of real particles if particles have homogeneous or heterogeneous compositions. As an illustration, we performed some simulations for a cubic, layered particle with a rib size of 5 μm, and having a layer thickness between 0.25 and 1.25 μm. The simulated intensities were then treated in the iterative quantification as if they were coming from a homogeneous particle. The resulting graphs clearly show that the variation of the outcome can be quite large.

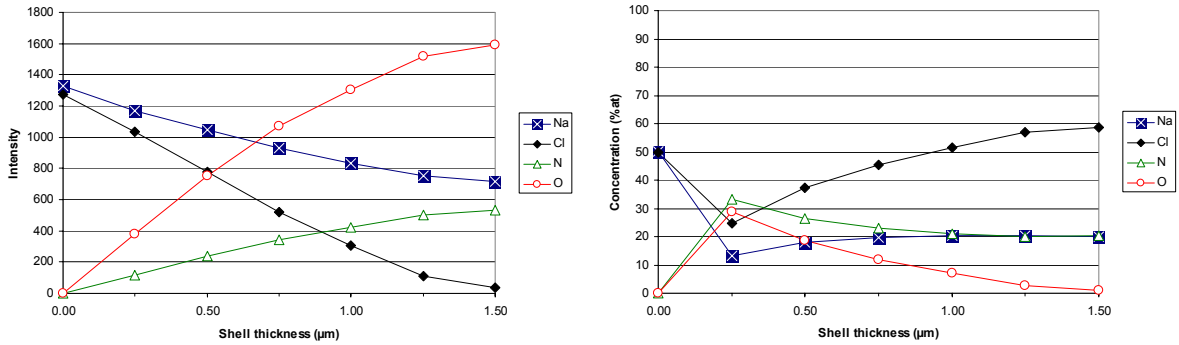

Figure 18: Layered particle falsely treated as homogeneous particle during quantification (left: intensity simulation for layered particle; right: calculation as if homogeneous)

We carried out some experiments to compare homogeneously mixed particles (having a certain composition ratio) with particles that were exposed to acid vapors (having a core surrounded by a shell of which the thickness depends on the exposure time). The mixed homogeneous particles were produced with the air brush method, using a solution of the involved compounds with the required concentration ratio.
Figure 19 and Figure 20 show the differences between the spectra at 5, 10 and 20 kV for both particle types, and for nitrates and sulfates respectively. In the graphs, the chloride peaks are marked, because they can be used as an indicator for deciding if a particle is homogeneous or not. At the lowest voltage the chloride peak of the layered particle has almost disappeared, because we only get information about the top layer.

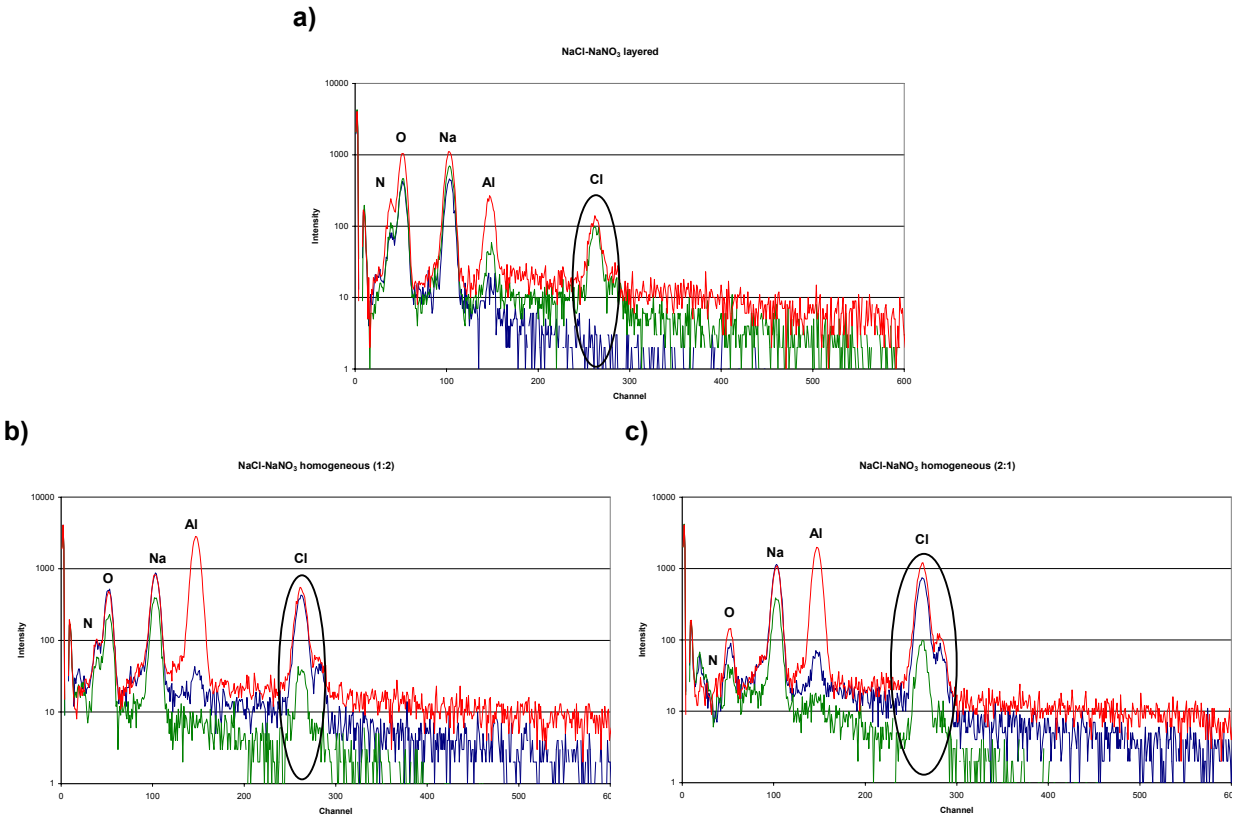

Figure 19: $NaCl$-$NaNO_3$ particles; a) layered, b) homogeneous 1:2, c) homogeneous 2:1

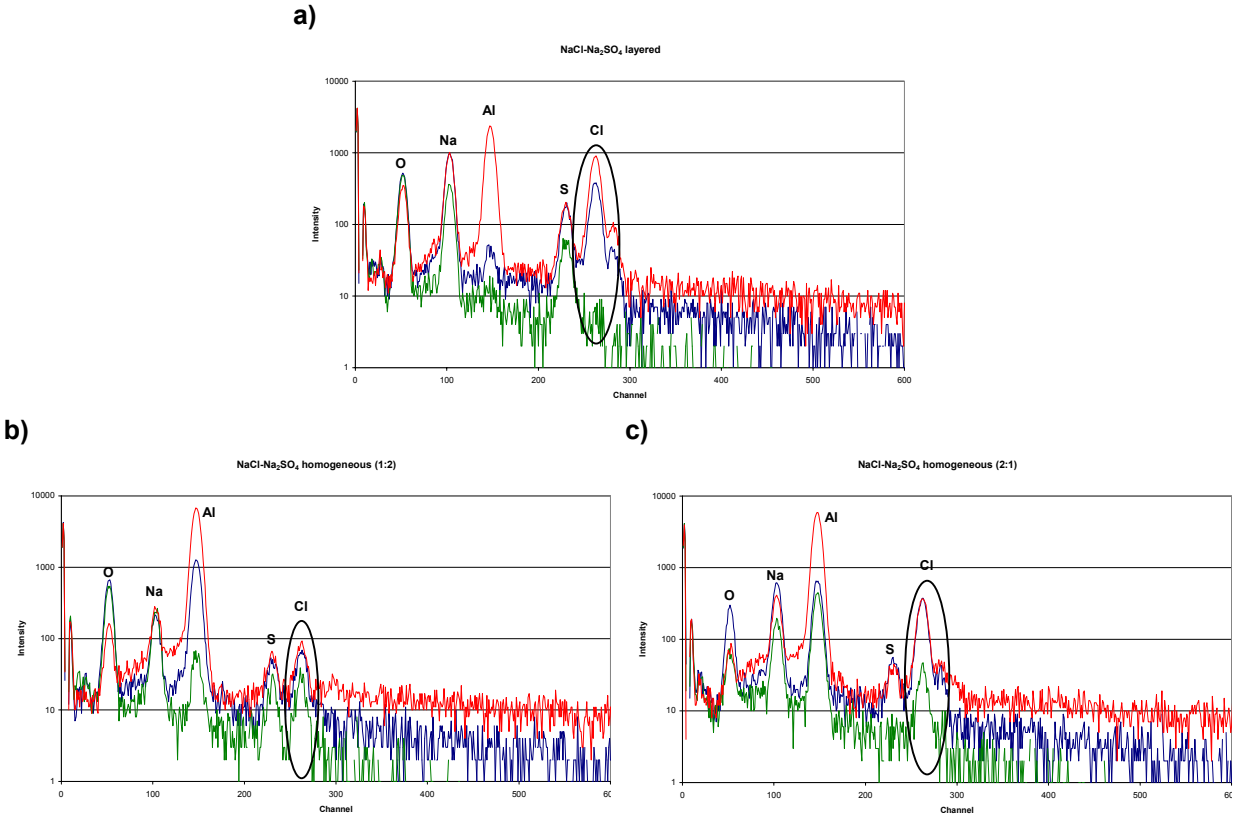

Figure 20: NaCl-Na_2SO_4 particles; a) layered, b) homogeneous 1:2, c) homogeneous 2:1

Although these results clearly show that it is possible to answer questions about the homogeneity and structural composition of particles, we should point out that the applications are still limited. In the cases above, we knew that we had to deal with very simple particle composition (only two compounds and four relevant elements; possible core/shell structure), so it was also obvious to look out for changes in the chloride peak. From this peak variation, it was quite obvious to determine if the analyzed particles had a core/shell composition or not. However, if we would have to analyze and investigate the composition of real environmental particles (like aged sea salt), we would have to deal with more elements and compounds. Moreover, the limitations mentioned under section 3.3 would also play an important role. The spectra above were taken from rather large particles (5-8 μm) and the layers of the heterogeneous particles were obtained after exposure times that allowed for the formation of a relatively thick layer. If the ratio of the particle core diameter and shell thickness would be relatively large (or small), problems would arise because the signals from the shell would be either too small or too large, causing difficulties for the spectral interpretation. The quantification problem is even larger: it is extremely difficult to put all the required interpretations into an algorithm that can handle the particle characterization all by itself. Therefore, MV-EPMA was, so far, only used in manual and qualitative mode. However, the technique clearly has high potential as a tool for studying the chemical and morphological transformations of atmospheric particles.

4. Conclusions

We tested different methods for getting the most out of the obtained light element concentrations values. Our goal in developing the iterative Monte Carlo simulations was obviously to obtain new insights for the characterization of environmental particles.

First we applied different mathematical tools for the statistical treatment of large datasets, like the ones we would obtain from the analysis of environmental samples. Experiments with artificial particles showed that the calculation of light element concentrations by our iterative Monte Carlo quantification procedure indeed produces an added value for differentiating between different particle types. These methods are able to mathematically reduce our large datasets to smaller sets of specific parameters that characterize the analyzed aerosols.
Secondly, the purely mathematical data treatment was extended with an expert system for the elucidation of particle composition. The mental algorithms used to identify particle types were translated into an automatic procedure that could do the same job of the analyst in an unsupervised and computerized way. This method starts on the individual particle level and ends with the classification of similar particles into typical groups.

Two methods were developed and tested for additional characterization of the internal particle composition. Grazing-Exit EPMA (GE-EPMA) was first tested for the analysis of particle layers. Although the technique is quite unique and could offer specific information about core-shell structured particles, some disadvantages due to beam damage and the instrumental setup strongly limited the possibilities and made data interpretation quite difficult. It is, however, important to mention that the results obtained for the analysis of $CaCO_3$-$CaSO_4$ and $CaCO_3$-$Ca(NO_3)_2$ particles were successfully used by Bekshaev et al. to determine particle heights.[56]
Another technique is Multiple-Voltage EPMA (MV-EPMA) in which particles are analyzed at different voltages to vary the information depth, so that layered particles could be qualitatively differentiated from homogeneous particles. The technique was also used to determine the surface layer thickness of artificial, heterogeneous particles, but beam damage again played a disadvantageous role.
It is clear that the application of GE-EPMA or MV-EPMA to real atmospheric particles will probably be limited to qualitative analysis due to the complexity involved in the analysis, since both their complex chemical composition and their structural heterogeneity cannot be known a priori. It was however shown that both techniques, in combination with light-element analysis, are able to extract information that could not be obtained before.
A logical conclusion, however, is that we will still only be able to obtain a qualitative idea about particle heterogeneity and that the results of the quantification procedure applied to real environmental samples should be carefully evaluated. We also only discussed the very simple case of layered particles; other, more complicated structures are also likely to exist (e.g. particles with a matrix disturbed by occasional zones of other composition, or agglomerations of small particles that appear to be a single, solid particle). For other (industrial) applications where particles are produced under controlled conditions, it is expected that the method could offer (semi-)quantitative insights into the particle composition and structure.

References

[1] Bondarenko, I., P. Van Espen, B. Treiger, R. Van Grieken and F. Adams, *Classification of coal mine dust particles through fuzzy clustering of their energy-dispersive electron microprobe X-ray spectra*, Microbeam analysis, 3, 33-37 (1994)

[2] Treiger, B., I. Bondarenko, P. Van Espen, R. Van Grieken, F. Adams, *Classification of mineral particles by nonlinear mapping of electron microprobe energy-dispersive X-ray spectra*, Analyst, 119, 971-974 (1994)

[3] Bondarenko, I., H. Van Malderen, B. Treiger, P. Van Espen and R. Van Grieken, *Hierarchical cluster analysis with stopping rules built on Akaike's information criterion for aerosol particle classification based on electron probe X-ray microanalysis*, Chemom. Intell. Lab. Sys., 22, 87-95 (1994)

[4] Bondarenko, I., B. Treiger, R. Van Grieken and P. Van Espen, *IDAS, a Windows based software package for cluster analysis*, Spectrochim. Acta B, 51, 441-456 (1991)

[5] Treiger, B., I. Bondarenko, H. Van Malderen and R. Van Grieken, *Elucidating the composition of atmospheric aerosols through the combined hierarchical, non-hierarchical and fuzzy clustering of large electron probe microanalysis datasets*, Anal. Chim. Acta, 317, 33-51 (1995)

[6] Shattuck, T.W., M.S. Germani and P.R. Buseck, *Multivariate statistics for large datasets: applications to individual aerosol particles*, Anal. Chem., 63, 2646-2656 (1991)

[7] Kim, D., and P.K. Hopke, *Classification of individual particles based on computer-controlled scanning electron microscopy data*, Aerosol Sci. Tech., 9, 133-151 (1988)

[8] Biggins, P.D.E., and R.M. Harrison, *Characterization and classification of atmospheric sulfates*, J. Air Poll. Control Assoc., 29, 838-840 (1979)

[9] Massart, D., and L. Kaufmann, *The Interpretation of Analytical Chemical Data by the Use of Cluster Analysis*, Wiley, New York, 1983

[10] Gordon, A.D., *Classification*, Chapman and Hall, Florida, 1999

[11] Malinowski, E.R., *Factor Analysis in Chemistry*, John Wiley, New York, 1991

[12] Osán, J., J. de Hoog, A. Worobiec, C.-U. Ro, K.-Y. Oh, I. Szalóki and R. Van Grieken, *Application of chemometric methods for classification of atmospheric particles based on thin-window electron probe microanalysis data*, Anal. Chim. Acta, 446, 211–222 (2001)

[13] Van Espen, P., *DPP, a program for the processing of analytical data*, Anal. Chim. Acta, 165, 31-46 (1984)

[14] Van Espen, P., *Quantitative microbeam analysis of particles*, Mikrochim. Acta, 114/115, 129-142 (1994)

[15] Bernard, P.C., R.E. Van Grieken and D. Eisma, *Classification of estuarine particles using automated electron microprobe analysis and multivariate techniques*, Environ. Sci. Technol., 20, 467-473 (1986)

[16] Ro, C.-U., J. Osán, I. Szalóki, K.-Y. Oh, H. Kim and R.E. Van Grieken, *Determination of chemical species in individual aerosol particles using ultra-thin window EPMA*, Environ. Sci. Technol., **34**, 3023-3030 (2000)

[17] Xie, Y., and P.K. Hopke, *Airborne particle classification with a combination of chemical compsotion and shape index utilizing an adaptive resonance artificial neural network*, Environ. Sci. Technol., 28, 1921-1928 (1994)

[18] Ro., C.-U., H. Kim and R. Van Grieken, *An expert system for chemical speciation of individual particles using low-Z particle EPXMA data*, Anal. Chem., 76, 1322-1327 (2004)

[19] Ro, C.-U., K.-U. Oh, H. Kim, Y. Chun, J. Osán, J. de Hoog and R. E. Van Grieken, *Chemical speciation of individual atmospheric particles using low-Z electron probe X-ray microanalysis: characterizing 'Asian Dust' deposited with rainwater in Seoul, Korea*, Atmos. Environ., 35, 4995-5005 (2001)

[20] Kyotani, T., and S. Koshimizu, *Identification of individual Si-rich particles derived from Kosa aerosol by the alkali elemental composition*, Bull., Chem. Soc. Jpn., 74, 723-729 (2001)
[21] Chen, L.-L., G.R. Carmichael, M.-S. Hong, H. Ueda, S. Shim, C.H. Song, Y.P. Kim, R. Arimoto, J. Prospero, D. Savoie, K. Murano, J.K. Park, H.-g. Lee and C. Kang, *Influence of continental outflow events on the aerosol composition at Cheju Island, South Korea*, J. Geophys. Res., 102, 28551-28574 (1997)
[22] Ro, C.-U., K.-Y. Oh, H. Kim, Y. P. Kim, C. B. Lee, K.-H. Kim, J. Osan, J. de Hoog, A. Worobiec and R. Van Grieken, *Single particle analysis of aerosols at Cheju Island, Korea, using low-Z electron probe X-ray microanalysis: a direct proof of nitrate formation from sea-salts*, Environ. Sci. Technol., 35, 4487-4494 (2001)
[23] Gomes, L., and D.E. Gillette, *A comparison of characteristics of aerosol from dust storms in Central Asia with soil-derived dust from other regions*, Atmos. Environ., 27A, 2539-2544 (1993)
[24] Avial, A., M. Alarcón and I. Queralt, *The chemical composition of dust transported in red rains*, Atmos. Environ., 32, 179-191 (1998)
[25] Levin, Z., E. Ganor and V. Gladstein, *The effect of desert particles coated with sulphate on rain formation in the eastern Mediterranean*, J. Appl. Meteorol., 35, 1511-1523 (1996)
[26] Kouimtzis, T., and C. Samara, *Airborne Particulate Matter*, Springer, Heidelberg (1995)
[27] Fenter, F.D., F. Caloz and M.J. Rossi, *Experimental evidence for the efficient dry deposition of nitric acid on calcite*, Atmos. Environ., 29, 3365-3372 (1995)
[28] Fenter, F.F., F. Caloz and M.J. Rossi, *Kinetics of nitric acid uptake by salt*, J. Phys. Chem., 98, 9801-9810 (1994)
[29] Leu, M.-T., R.S. Timonen, L.F. Keyser and Y.L. Yung, *Heterogeneous reactions of $HNO_3(g) + NaCl(s) \rightarrow HCl(g) + NaNO_3(s)$ and $N_2O_5(g) + NaCl(s) \rightarrow ClNO_2(g) + NaNO_3(s)$*, J. Phys. Chem., 99, 13203-13212 (1995)
[30] Mori, I., M. Nishikawa and Y. Iwasaka, *Chemical reaction during the coagulation of ammonium sulphate and mineral particles in the atmosphere*, Sci. Total Environ., 224, 87-91 (1998)
[31] Munger, J.W., J.M. Waldman, D.J. Jacob and M.R. Hoffmann, *Fog water chemistry in an urban atmosphere*, J. Geophys. Res., 88, 5109-5123 (1983)
[32] Freiberg, J.E., and S.E. Schwartz, *Mass-transport limitation to the rate of reaction of gases in liquid droplets: application to oxidation of SO_2 in aqueous solutions*, Atmos. Environ., 15, 1145-1154 (1981)
[33] Davies, J.A. and R.A. Cox, *Kinetics of the heterogeneous reaction of HNO_3 with NaCl: effect of water vapour*, J. Phys. Chem., 102, 7631-7642 (1998)
[34] Winkler, P., *The growth of atmospheric aerosol particles with relative humidity*, Phys. Scripta, 37, 223-230 (1988)
[35] Roth, B., and K. Okada, *On the modification of sea-salt particles in the coastal atmosphere*, Atmos. Environ., 32, 1555-1569 (1998)
[36] Dentener, F.J., G.R. Carmichael, Y. Zhang, J. Lelieveld and P.J. Crutzen, *Role of mineral aerosol as a reactive surface in the global troposphere*, J. Geophys. Res., 101, 869-889 (1996)
[37] Carson, P.G., M.V. Johnston and A.S. Wexler, *Real-time monitoring of the surface and total composition of aerosol particles*, Aeros. Sci. Techn., 26, 291-300 (1997)
[38] Underwood, G.M., T.M. Miller and V.H. Grassian, *Transmission FTIR and Knudsen cell study of the heterogeneous reactivity of gaseous nitrogen dioxide on mineral oxide particles*, J. Phys. Chem., 103, 6184-6190 (1999)

[39] Underwood, G.M., P. Li, C.R. Usher and V.H. Grassian, *Determining accurate kinetic parameters of potentially important heterogeneous atmospheric reactions on solid particle surfaces with Knudsen cell reactor*, J. Phys. Chem., 104, 819-829 (2000)

[40] Cheng, R.J., D.C. Blanchard and R.J. Cipriano, *The formation of hollow sea-salt particles from the evaporation of droplets of seawater*, Atmos. Res., 22, 15-25 (1988)

[41] Ivanov, V.P., D.I. Kochuby, K.P. Kutzenogii and N.S. Bufetov, *Surface composition of atmospheric aerosols*, React. Kinet. Catal. Lett., 64, 97-102 (1998)

[42] Li, P., K.A. Perreau, E. Covington, C.H. Song, G.R. Carmichael and V.H. Grassian, *Heterogeneous reactions of volatile organic compounds on oxide particles of the most abundant elements present in the earth's crust: surface reactions of acetaldehyde, acetone and propionaldehyde on SiO_2 , Al_2O_3 , Fe_2O_3 , TiO_2 and CaO*, J. Geophys. Res., 106 , 5517-5529 (2001)

[43] Song, C.H. and G.R. Carmichael, *The aging process of naturally emitted aerosol (sea-salt and mineral aerosol) during long range transport*, Atmos. Environ., 33, 2203-2218 (1999)

[44] Buseck, P.R., and M. Pósfai, *Airborne minerals and related aerosol particles: effects on climate and the environment*, Proc. Natl. Acad. Sci. USA, 96, 3372-3379 (1999)

[45] Radojevic, M., *SO_2 and NO_x oxidation mechanisms in the atmosphere* in Radojevic, M., and R. Harrisson, *Atmospheric Acidity*, Elsevier, Essex (1992)

[46] Tsuji, K., R. Nullens, K. Wagatsuma and R. Van Grieken, *Elemental X-ray images obtained by grazing-exit electron probe microanalysis (GE-EPMA)*, J. Anal. At. Spectrom., 14, 1711-1713 (1999)

[47] Tsuji, K., K. Wagatsuma, R. Nullens and R. Van Grieken, *Grazing exit electro probe microanalysis for surface and particle analysis*, Anal. Chem., 71, 2497-2501 (1999)

[48] Tsuji, K., Y. Murakami, K. Wagatsuma and R. Van Grieken, *Surface studies by grazing-exit electron probe microanalysis (GE-EPMA)*, X-ray Spectrom., 30, 123-126 (2001)

[49] Tsuji, K., Z. Spolnik, K. Wagatsuma, J. Zhang, R. E. Van Grieken, *Enhancement of electron-induced X-ray intensity for single particles under grazing-exit conditions*, Spectrochim. Acta, B54, 1243-1251 (1999)

[50] Pouchou, J.-L., *X-ray micoanalysis of stratified specimens*, Anal. Chim. Acta, 283, 81-97 (1993)

[51] Kiss, K., *Quantitative electron probe analysis of low-atomic number samples with irregular surfaces*, Appl. Spectr., 37, 19-25 (1983)

[52] Hnizdo, V., W.E. Wallace, *Monte Carlo analysis of the detection of clay occlusion of respirable quartz particles using multiple voltage scanning electron microscopy*, Scanning, 24, 264-269 (2002)

[53] Ammann, N., and P. Karduck, *Quantitative depth profile analysis by EPMA combined with Monte Carlo simulation*, Surf. Interf. Anal., 22, 54-59 (1994)

[54] Berner, A., and G. Proaktor, *Quantitative EPMA of element depth distribution*, Mikrochim. Acta, 114/115, 195-203 (1994)

[55] Goodman, A.L., G.M. Underwood and V.H. Grassian, *A laboratory study of the heterogeneous reaction of nitric acid on calcium carbonate particles*, J. Geophys. Res., 105, 29053-29064 (2000)

[56] Bekshaev, A., J. de Hoog and R. Van Grieken, *Grazing-emission electron probe microanalysis of particles near the substrate edge*, Spectrochim. Acta B, 56, 2385-2395 (2002)