We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

BOOK
CITATION
INDEX
CLARIVATE ANALYTICS
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Application of Multivariate Data Analyses in Waste Management

K. Böhm, E. Smidt and J. Tintner

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/53975

1. Introduction

First of all, what is multivariate data analysis and why is it useful in waste management?

Methods dealing with only one variable are called univariate methods. Methods dealing with more than one variable at once are called multivariate methods. Using univariate methods natural systems cannot be described satisfactorily. Nature is multivariate. That means that any particular phenomenon studied in detail usually depends on several factors. For example, the weather depends on the variables: wind, air pressure, temperature, dew point and seasonal variations. If these factors are collected every day a multivariate data matrix is generated. For interpretation of such data sets multivariate data analysis is useful. Multivariate data analysis can be used to process information in a meaningful fashion. These methods can afford hidden data structures. On the one hand the elements of measurements often do not contribute to the relevant property and on the other hand hidden phenomena are unwittingly recorded. Multivariate data analysis allows us to handle huge data sets in order to discover such hidden data structures which contributes to a better understanding and easier interpretation. There are many multivariate data analysis techniques available. It depends on the question to be answered which method to choose.

Due to the requirement of representative sampling number of samples and analyses in waste management lead to huge data sets to obtain reliable results. In many cases extensive data sets are generated by the analytical method itself. Spectroscopic or chromatographic methods for instance provide more than 1000 data points for one sample. Evaluation tools can be developed to support interpretation of such analytical methods for practical applications. For specific questions and problems different evaluation tools are necessary. Calculation and interpretation are carried out by the provided evaluation tool.

In this study an overview of multivariate data analysis methods and their application in waste management research and practice is given.

2. Multivariate data analysis in waste management

The main objectives of multivariate data analysis are exploratory data analysis, classification and parameter prediction. Many different multivariate data analysis methods exist in literature. Thus the following list is not exhaustive however subdivided into the mentioned superior categories. It only concentrates on the methods applied in waste management.

Table 1 gives an overview of the existing literature in waste management on multivariate data analysis applied by several authors. It can be summarised that PCA and PLS1 are the most popular multivariate data analysis methods applied in waste management. Details are given in the following sections 2.1 and 2.2. Due to easy traceability of the parameters investigated in the different papers parameter descriptions have been taken as they were mentioned in the original.

In practice there are many software packages available which include different multivariate data analysis methods. Some software tools are: SPSS (www.spss.com\de\statistics), Canoco (www.canoco.com), The Unscrambler (www.camo.com) and the Free Software R-project (www.cran.r-project.org).

	Pattern recognition						Calibration			
Method	PCA	FA	CCA	CA	DA	SIMCA	MLR	PLS1	PLS2	PSR
Chapter		2.1.1		2.1.2	2.1.3	2.1.3	2.2.1		2.2.2	
Compost science	[1-23]	[24]	[25]	[1, 4, 22, 24-31]	[3, 9]	[8, 12]	[29, 32, 33]	[2, 6, 8, 19, 21, 23, 34-47]	[8, 21, 48]	[49]
Municipal solid waste	[50-55]			[56]				[17, 53, 57, 58]		
Landfill research	[59-72]	[65]	[73, 74]	[72, 75]	[66, 71, 76, 77]		[78]	[79, 80] [17, 61, 62, 66, 71, 78]		
Logistics	[81]	[82]		[82]				[83, 84]		

Table 1. Literature review of different multivariate data analysis methods applied in waste management; PCA – Principal Component Analysis, FA – Factor Analysis, CA – Cluster Analysis, CCA – Canonical Correspondence Analysis, DA – Discriminant Analysis, SIMCA – Soft Independent Modelling of Class Analogy, MLR – Multiple Linear Regression, PLS-R – Partial Least Squares Regression, PSR – Penalised Signal Regression

2.1. Pattern recognition

2.1.1. Exploratory data analysis

- Principal Componant Analysis (PCA)

PCA is mathematically defined as an orthogonal linear transformation that arranges the data to a new coordinate system in that the greatest variance by any projection of the data takes place along the first coordinate (called the first principal component), the second greatest variance along the second coordinate, and so on. Theoretically the PCA is the optimum transformation for a given data set in least square terms. That means PCA is used for dimensionality reduction of variables in a data set by retaining those characteristics of the data set that contribute most to its variance. The transformation to the new coordinate system is described by scores (T), loadings (P) and errors (E). In matrix terms, this can be written as $X = T * P + E$. Fig. 1 illustrates the mathematical transformation using PCA. The matrices can be displayed graphically. The scores matrix illustrates the data structure and the loading matrix displays the influence of the different variables on the data structure.
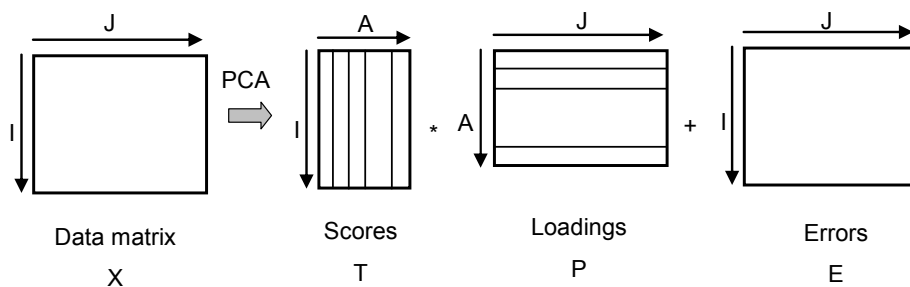

Figure 1. Principle of the PCA (according to Esbensen [85])

PCA displays hidden structures of huge data sets. PCA is applied in different fields of waste management to find out the relevant parameters of a large parameter set. So we can see which properties of a sample are significant and important to answer a particular question. Due to the results obtained time and money can be saved in further research activities.

Many applications can be found in compost science. Zbytniewski and Buszewski [1] applied PCA to reveal the significant parameters and possible groupings of chemical parameters, absorption band ratios and NMR data. Campitelli and Ceppi [3] investigated the quality of different composts and vermicomposts. The collected data were evaluated by means of PCA to extract the significant differences between the two compost types. Gil et al. [4] used PCA to show effects of cattle manure compost applied on different soils. Termorshuizen et al. [13] carried out a PCA based on disease suppression data determined by bioassays in different compost/peat mixtures and pure composts. PCA was applied by Planquart et al. [10] to examine the interactions between nutrients and trace metals in colza (Brassica napus) when sewage sludge compost was applied to soils. LaMontagne et al. [7] applied PCA on terminal restriction fragment length polymorphisms (TRFLP) patterns of different composts to reveal their characteristics with respect to microbial communities. Malley et al. [8] recorded near infrared spectra from cattle manure during composting. The collected spectral data were

evaluated by PCA to show the relationships among samples and changes due to stockpiling and composting. Hansson et al. [6] observed the anaerobic treatment of municipal solid waste by using on-line near infrared spectroscopy. For spectral data interpretation PCA was carried out. Albrecht et al. [2] also performed a PCA for near infrared (NIR) spectra evaluation from an ongoing composting process. Smidt et al. [12] used PCA to show differences in spectral characteristics of different waste materials. Lillhonga et al. [23] used PCA to observe spectral characteristics of different composting processes. Vergnoux et al. [21] applied a PCA on NIR spectra as well as on physico-chemical and biochemical parameters to derive regularities from the data. Nicolas et al. [9] used PCA to evaluate data from an electronic nose. The correlations between the sensor of an electronic nose and chemical substances were determined by Romain et al. [11] using PCA. PCA was applied to observations of a composting process by means of analytical electrofocusing. The electrofocusing profiles were evaluated by Grigatti et al. [5]. PCA was also used by Biasioli et al. [19] to evaluate odour emissions and biofilter efficiency in composting plants using proton transfer reaction-mass spectrometry. Bianchi et al. [18] also used PCA to reduce the complex data set and to analyse the pattern of organic compounds emitted from a composting plant, a municipal solid waste landfill and ambient air. The effect of 14 different soil amendments on compost quality were evaluated using a PCA by Tognetti et al. [20]. Smidt et al. [16] applied PCA to illustrate the influence of input materials and composting operation on humification of organic matter. Böhm et al. [14] and Smidt et al. [15, 17] used PCA to illustrate spectral differences caused by different materials such as biowaste, manure, leftovers, straw and sewage sludge.

PCA was also applied to illustrate the alteration of municipal solid waste during the biological degradation process reaching stability limits for landfilling as well as to demonstrate similarities and differences of reactor and old landfills based on thermal data [53, 66]. Scaglia and Adani [52] focused on municipal solid waste treatment. They used PCA to create a stability index for quantifying the aerobic reactivity of municipal solid waste. Abouelwafa et al. [54, 55] investigated the degradation of sludge from the effluent of a vegetable oil processing plant mixed with household waste from landfill. Abouelwafa et al. [54] applied PCA on various parameters measured during composting (e.g. pH, electrical conductivity, moisture, C/N, NH_4/NO_3, ash, decomposition in percent, level of polyphenols, lignin, cellulose, hemicellulose, humic acid) to find the main parameters in the decomposition and restructuring phase [54]. Abouelwafa et al. [55] extracted fulvic acids from the samples mentioned above and extended the data set used for PCA by a series of absorption band ratios resulting from of FTIR spectra.

PCA has also been used in landfill research. Mikhailov et al. [62] applied PCA for monitoring data from different landfills. They included parameters such as depth, ash content, volumetric weight, humidity, amounts of refuse in summer and winter as well as the topsoil depth of landfill sections, sewage sludge lenses and the existence of a protection system. Kylefors [61] investigated data of leachate composition using PCA. The idea was to reduce the analytical monitoring program for further investigations. Durmusoglu and Yilmaz [60] used PCA to extract the significant independent variables of the collected data of raw and pre-treated leachate. A comparable work was done by De Rosa et al. [59]. They also investigated the leachate composition of an old waste dump connected to the

groundwater. Olivero-Verbel et al. [63] investigated the relationships between physico-chemical parameters and the toxicity of leachates from a municipal solid waste landfill. PCA was used to find out which parameters were responsible for their toxicity. Jean and Fruget [72] used PCA to compare landfill leachates according to their toxicity and physico-chemical parameters. Ecke et al. [71] showed an example for PCA application in landfill monitoring of data from landfill test cells, leachate and gas data. Smidt et al. [64] investigated landfill materials by means of mid infrared spectroscopy, thermal analysis and PCA. They used PCA to support data interpretation. Van Praagh et al. [70] investigated the potential impacts on leachate emissions using pretreated and untreated refuse-derived material as a cover layer on the top of a municipal solid waste landfill. To interpret leachate characteristics they used PCA. Tintner and Klug [69] used PCA to illustrate how vegetation can indicate landfill cover features. Diener et al. [67] investigated the long-term stability of steel slags used as cover construction of a municipal solid waste landfill by means of a PCA. Smidt et al. [17] used PCA to display spectral characteristics of different landfill types.

Pablos et al. [68] used a PCA to evaluate toxicity bioassays for biological characterisation of hazardous wastes.

Other publications focus on the process monitoring of municipal solid waste incineration residues. Ecke [50] performed PCA on leaching parameters from municipal solid waste incineration fly ash to get an overview of the mobility of metals under certain conditions. Mostbauer et al. [51] carried out PCA to observe the long-term behaviour of municipal solid waste incineration (MSWI) residues.

In the field of waste management logistics PCA is rarely applied. Dahlén et al [81] used PCA to display the impact of waste costs on a weight basis in a specific municipality.

- Factor Analysis (FA)

FA is related to PCA but differs in its mathematical conception [86]. FA is also used to describe the variability of observed variables in terms of fewer variables called factors. That means factor analysis is a tool which reveals unobservable underlying features of a specific phenomenon by previous visible observations. The observed variables are modelled as linear combinations of the factors plus "error" terms. The information about interdependencies can be used to reduce the number of variables in a data set.

In waste management practice PCA is preferentially used. Differences between factor analysis and PCA are found to be small [86]. Srivastava and Ramanathan [65] investigated the groundwater quality of a landfill site in India by means of FA. They explained the observed relationship in simple terms expressed as factors. Bustamante et al. [24] used FA to identify the principal variables associated to the composting of agro-industrial wastes. Lin et al. [82] used FA for selecting the best food waste recycling method.

- Canonical Correspondence Analysis (CCA)

CCA is a multivariate method to explain the relationships between biological communities and their environment [87]. The method is designed to extract environmental gradients from

ecological data sets. By means of the gradients an ordination diagram describing and visualising the diverse habitat preferences of taxa is calculated.

CCA is sometimes used in waste management if, for example, microbial communities or vegetation surveys are analysed. CCA was applied by Franke-Whittle et al. [25] and El-Sheikh et al. [73]. Franke-Whittle et al. [25] applied CCA to illustrate the similarities in microbial communities of three different composting processes. El-Sheikh et al. [73] investigated the ten-year primary succession on a newly created landfill at a lagoon of the Mediterranean Sea. Vegetation surveys where the basis for CCA. Kim et al. [74] applied CCA to investigate the vegetation and the soil of a not properly maintained landfill to suggest restoration alternatives by comparing the vegetation of the landfill to the nearby forests.

2.1.2. Unsupervised pattern recognition

- Cluster analysis (CA)

Clustering is the classification of objects into groups called clusters. Objects from the same cluster are more similar to one another than objects from different clusters. The difference of clusters is based on measured distances without any unit. Cluster analysis can be illustrated graphically in a dendrogram as shown in Fig. 2. The samples 2, 3 and 5 are clustered due to the high degree of similarity as well as the samples 1 and 4. The two clusters show little similarity.
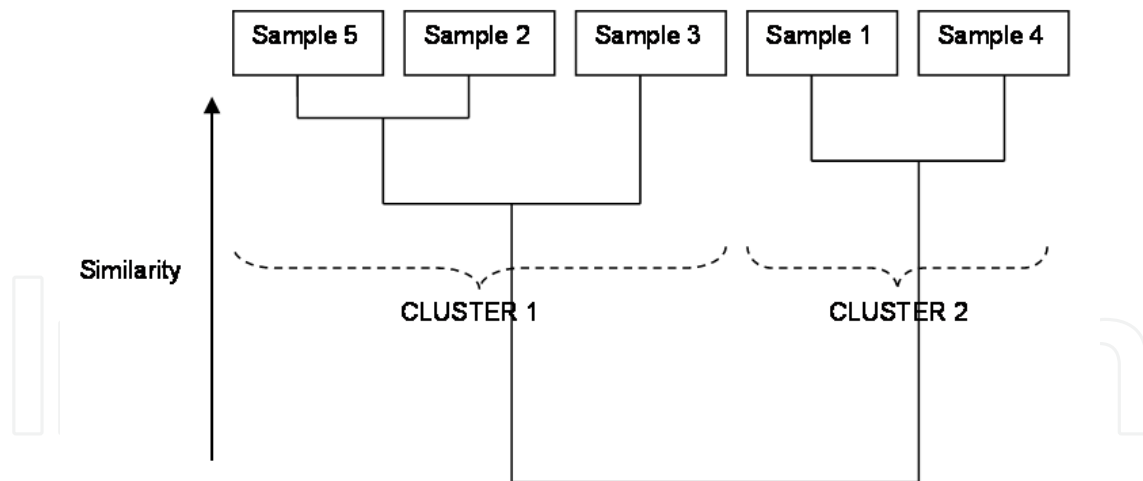

Figure 2. Example of a cluster analysis visualised by a dendrogram

CA was applied in compost science by Zybtniewskie and Buszewski [1]. They applied CA to conventional compost parameters and NMR data to find out the grouping depending on the composting time. He et al. [56] used a hierarchical cluster analysis to show the similarities and differences of UV-Vis and fluorescence spectra of water extractable organic matter, originating from municipal solid waste that had been subjected to different composting times. A hierarchical cluster analysis was also used by He et al. [22] to investigate water-

extractable organic matter during cattle manure composting. Gil et al. [4] displayed dendrograms to illustrate the similarities or differences by application of cattle manure compost to different soils. Bustamante et al. [24] studied physico-chemical, chemical and microbiological parameters of different composts. The evaluation of the composts was conducted by a hierarchical cluster analysis [24].

Lin et al. [82] applied a CA for the selection of optimal recycling methods for food waste.

A stepwise cluster analysis (SCA) was used to describe the nonlinear relationships among state variables and microbial activities of composts by Sun et al. [29]. Sun et al. [30] developed a genetic algorithm aided stepwise cluster analysis (GASCA) to describe the relationships between selected state variables and the C/N ratio in food waste composting.

Furthermore CA has often been used to evaluate microbiological data, especially in compost science [25-28, 31]. Innerebner et al. [26] and Ros et al. [27, 28] used CA to identify related samples and similar groups of microorganisms. Franke-Whittle et al. [25] used CA to show the similarities of Denaturing Gradient Gel Electrophoresis (DGGE) data of three different compost types with proceeding compost maturity. Xiao et al. [31] used a hierarchical cluster analysis of DGGE data to estimate the succession of bacterial communities during the active composting process.

Tesar et al. [75] applied CA to spectral data to illustrate the effect of in-situ aeration of a landfill. Jean and Fruget [72] used CA to compare landfill leachates on the basis of their toxicity and physico-chemical parameters.

2.1.3. Supervised pattern recognition

All supervised methods are classifications. Classification can be considered as a predictive method where the response is a category variable. Different classification methods exist. There are types of "hard" and "soft" modelling. Hard modelling means that a non-relocatable line between the defined groups exists. One object can only belong to one group. Soft modelling allows an overlapping of the defined classes. An object can belong to both groups [88]. With regard to waste management practice two different classification methods are described in detail.

- Discriminant analysis (DA)

DA is a classification method of hard modelling. Campitelli and Ceppi [3] carried out a DA to distinguish between compost and vermicompost on the basis of parameters such as total organic carbon (TOC), germination index (GI), pH, total nitrogen (TN), and water soluble carbon (WSC). Nicolas et al. [9] performed a DA to classify data of an electric nose according to defined exceeded levels of odour. Ecke et al. [71] investigated samples from three different landfill sites by the biochemical methane potential and used DA for data evaluation. Huber-Humer et al. [77] applied DA to determine methane oxidation efficiency of different materials based on chemical and physical variables. Smidt et al. [66, 76] used DA to differentiate the infrared spectral [76] and thermal patterns [66] of municipal solid waste

incinerator (MSWI) bottom ash before and after CO_2 uptake. A DA on the CO_2 ion current recorded during combustion was applied to illustrate the effect of CO_2 treatment of MSWI bottom ash [66]. DA was also used to illustrate the spectral characteristics of leachate from landfill simulation reactors under aerobic and anaerobic conditions [17].

- Soft independent modelling of class analogy (SIMCA)

SIMCA is a special method of soft modelling recommended by Wold in the 1970s [88]. Objects can belong to one of the defined class, to both classes or to none. Whether SIMCA can be applied on the data set depends on the question to be answered. According to Brereton [88] it is often legitimate in chemistry that an object belongs to more than one class For example a compound may have an ester and an alkene group which are both reflected by an infrared spectrum. Thus they fit in both classes. In natural science it is allowed in most cases for an object to be in line with more than one class simultaneously.

Contrarily in other cases an object can belong only to one class and the application of SIMCA is inappropriate. Brereton [88] gives a good example where the concept of SIMCA is not applicable: A banknote is either forged or not. In many cases there is only one true answer. For such problems SIMCA is not the adequate method.

In compost science Malley et al. [8] and Smidt et al. [12] carried out a SIMCA. Malley et al. [8] classified different decomposition stages of manures by means of near infrared spectroscopy and SIMCA. Smidt et al. [12] carried out a SIMCA to classify different waste materials such as biowaste compost, mechanically-biologically pretreated waste and landfill materials based on their spectroscopic pattern. Smidt et al. [78] used the SIMCA model developed by Smidt et al. [12] to identify different landfill types such as reactor landfill and industrial landfill samples.

2.2. Calibration

2.2.1. Multiple Linear Regression (MLR)

MLR is directed at modelling the relationship between two or more explanatory variables and a response variable by fitting a linear equation to observed data. Every value of the independent variable X is associated with a value of the dependent variable Y, with explanatory or predictive purposes. A direct correlation between Y and X-matrix is performed.

In waste management MLR was applied by Chikae et al. [32] to predict the germination index which was adopted as a marker for compost maturity. Thirty-two parameters of 159 samples were measured. MLR was carried out to reduce this huge parameter set to some significant parameters. Lawrence and Boutwell [79] used MLR for predicting the stratigraphy of landfill sites using an electromagnetic method. Moreno-Santini et al. [80] applied MLR to determine arsenic and lead levels in the hair of residents in a municipality constructed on a former landfill.

Noori et al. [84] compared two different statistical methods (artificial neural networks and MLR based on a PCA) to predict the solid waste generation in Tehran. Cheng et al. [83] used

MLR to predict the factors associated with medical waste generation at hospitals. Sun et al. [29] used MLR to predict mesophilic and thermopilic bacteria in food waste composts. Suehara and Yano [33] applied MLR to predict conventional compost parameters by NIR spectral data.

2.2.2. Partial Least Squares Regression (PLS-R)

- PLS1

PLS-R is used to find out the fundamental relations between two matrices. PLS-R is a bilinear modelling method. The main idea behind it is to calculate the principal components of the X and the Y matrix separately (external correlation) and to develop a regression model between the scores of the principal components (inner correlation). The concept of PLS-R is demonstrated in Fig. 3.

PLS1 is often used to predict time consuming or expensive parameters using an alternative analytical method. Modern analytical tools such as spectroscopic, chromatographic and thermo analytical methods generate data with inherent information on different parameters. With the development of an evaluated prediction model conventional analytical methods can be replaced by easier and/ or faster handling and robust methods.
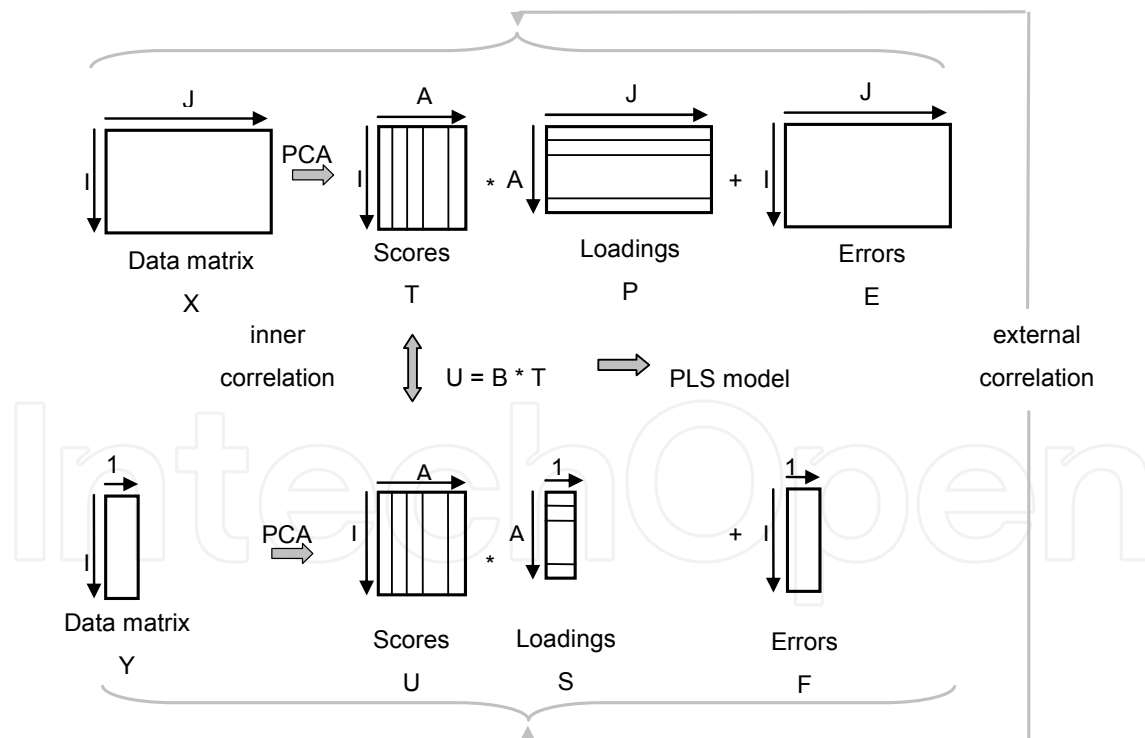

Figure 3. Principles of PLS-R (according to Esbensen [85])

Many authors have developed such prediction models in compost science. Zvomuya et al. [44] predicted phosphorus availability in soils, amended with composted and non-composted cattle manure by means of cumulative phosphorus analysis. Fujiwara and

Murakami [35] applied near infrared spectroscopy to estimate available nitrogen in poultry manure compost. Huang et al. [36] also used near infrared spectroscopy to estimate pH, electric conductivity, volatile solids, TOC, total N, the C:N ratio and the total phosphorus content. Furthermore they determined nutrient contents such as K, Ca, Mg, Fe and Zn of animal manure compost using near infrared spectroscopy and PLS1 [37]. Malley et al. [8] developed prediction models for total C, organic C, total N, C:N ratio, K, S and P by means of near infrared spectroscopy and PLS1. Morimoto et al. [43] carried out carbon quantification of green grass tissue using near infrared spectroscopy. Hansson et al. [6] predicted the concentration of propionate in an anaerobic process by near infrared spectra. Albrecht et al. [2] developed calibration models between spectral data and C, N, C:N ratio and composting time. Michel et al. [42] predicted chemical and biological properties of composts such as organic C (C_{org}), total N, C:N ratio, age, microbial biomass (C_{mic}), $C_{mic}:C_{org}$, basal respiration, enzymatic activity and plant suppression using near infrared spectroscopy. Ludwig et al. [39] also used near infrared spectroscopy to predict pH, electric conductivity, P, K, NO_3^- and NH_4^+ and phytotoxicity. Ko et al. [38] predicted heavy metal contents of Cr, As, Cd, Cu, Zn and Pb by means of near infrared spectroscopy and PLS1. They hypothesised that heavy metals are detectable by NIR when they are complexed with organic matter. Capriel et al. [34] found out that mid infrared spectroscopy is a rapid method to estimate the effect of nitrogen and relevant parameters such as total C, total N, the C:N ratio and the pH of biowaste compost. Meissl et al. [40] used PLS1 and the mid infrared region to predict humic acid contents in biowaste composts. Furthermore they determined humic acid contents by near infrared spectroscopy [41]. Sharma et al. [47] developed prediction models for conventional compost parameters, especially ammonia, pH, conductivity, dry matter, nitrogen and ash using NIR and Vis-NIR spectroscopy. Lillhonga et al. [23] used PLS-R for compost parameter prediction based on NIR spectra. They developed models for the parameters: time, pH, temperature, NH_3/NH_4^+, energy (calorific value) and moisture content. Galvez-Sola et al. [45] used PLS1 to predict different compost quality parameters such as pH, electric conductivity, total organic matter, total organic carbon, total N, C/N ratio as well as nutrients contents (N, P, K) and potentially pollutant element concentrations (Fe, Cu, Mn and Zn) from near infrared spectra. Vergnoux et al. [21] applied a PLS1 to predict physico-chemical and biochemical parameters from NIR spectra. Physico-chemical parameters comprised age, organic carbon, organic nitrogen, C/N, total N, fulvic acids (FA), humic acids (HA) and HA/FA. The soluble fraction, lignin and biological maturity index were summarised as biochemical parameters. Mikhailov et al. [62] used PLS1 to predict maturity and stability based on conventionally measured data. Kylefors [61] developed prediction models for leachate concentrations of specific organic substances in leachate by means of conventional leachate analysis and PLS1. Biasioli et al. [19] used PLS1 to predict odour concentrations in composting plants by proton transfer reaction-mass spectrometry (PTR-MS). Mohajer et al. [46] used a PLS1 to generate a model to predict the microbial oxygen uptake in sludge based on different physical compost parameters.

Böhm et al. [57] used PLS1 to predict the respiration activity (RA4) based on FT-IR spectra of mechanically-biologically pretreated (MBT) waste. The potential of thermal data of MBT waste was shown by Smidt et al. [53]. They applied PLS1 to predict the calorific value, total organic carbon (TOC) and respiration activity (RA4). Smidt et al. [17] also developed a prediction model for the calorific value based on spectral data. Biasioli et al. [58] used PLS1 to predict odour concentration from MSW composting plants based on PTR-MS.

Ecke et al. [71] performed detoxification of hexavalent chromium to less toxic trivalent chromium in industrial waste and applied a PLS model to identify the relevant factors. Smidt et al. [78] predicted the biological oxygen demand and the dissolved organic carbon (DOC) of old landfill materials from spectral data. They also used PLS-R to predict the total organic carbon and total nitrogen based on thermal data [78]. Furthermore PLS-R was used to predict respiration activity (RA4) from MS data of old landfill materials [66]. Smidt et al. [17] developed a prediction model for the DOC and the TOC from spectral data of landfill materials.

- PLS2

PLS2 is a variant of the PLS-R method where several Y-variables are modelled simultaneously. An advantage of this method is to find possible correlations or co-linearity between the Y-variables.

Malley et al. [8] developed prediction models for pH, total N, nitrate and nitrite, total C, organic C, C:N ratio, P, available P, S, K and Na by means of near infrared spectroscopy and PLS2. Suehara et al. [48] used PLS2 for simultaneous measurement of carbon and nitrogen content of composts using near infrared spectroscopy. Vergnoux et al. [21] applied PLS2 to predict physico-chemical (moisture, temperature, pH, NH4-N) and biochemical parameters (hemicellulose and cellulose) from NIR spectra.

- Penalised signal regression (PSR)

This special regression method is described in Galvez-Sola et al. [49]. Galves Sola et al. [49] used this method to predict the phosphorus content in composts.

3. Selected examples from literature using multivariate data analysis in waste management

In the following chapter four selected examples using multivariate data analysis in waste management are described in detail. To illustrate the application of principal component analysis (PCA) the study by Mikhailov et al. [62] is presented. He carried out multivariate data analysis for the ecological assessment of landfills. The second example illustrates the application of partial least squares regression (PLS-R). Michel et al. [42] applied PLS-R to predict conventional parameters by spectroscopic data. Ros et al. [27] applied a cluster analysis to data of polymerase chain reaction coupled with denaturing gradient gel electrophoresis (PCR-DGGE) to observe the long-term effects of compost amendment on soil microbial activity. A soft independent model of class analogy (SIMCA) was applied by Malley et al. [8]. They used SIMCA to classify different composts according to their spectroscopic characteristic.

3.1. Principal component analysis (PCA)

3.1.1. Objective of the study

The objective of the study by Mikhailov et al. [62] was to evaluate the stability of landfills based on many conventional parameters such as ash content, temperature, volume weight, pH, humidity and depth. They supposed that a multivariate approach could provide a more efficient data interpretation. Therefore they compared conventional and multivariate data analysis methods.

3.1.2. Method of evaluation and results

In a first step Mikhailov et al. [62] collected conventional data to describe landfill stability. They investigated 3 different landfills in Russia, one illegal dump, an old poorly-run dump and a modern well-run landfill. They focused on geodesic surveys to obtain the overall object properties such as size, volume and different layers. Furthermore they investigated the physical and chemical properties of the samples collected in different depths of the landfill. The physical and chemical properties include ash content, humidity, and acidity. Using the conventional collected data they carried out a PCA for each landfill site. They included the ash content, temperature, volume weight, pH, humidity and depth. The PCA for the two landfills in Bezenchuk and Kinel are presented in the study [62]. Based on the data pool Mikhailov et al. [62] could identify two important sources of waste around Bezenchuk, a poultry farm and a granary. In addition to regular domestic refuse, the agricultural and industrial wastes were disposed illegally in this dump. Kinel on the other hand is a modern, well operated landfill, in which both domestic and industrial wastes are disposed. These assumptions were confirmed by chemometric investigations based on PCA. The mentioned PCAs show clustering of the different classes. The results of the PCA of the third investigated landfill are not shown in their study. Otradny was shown to be a poorly maintained landfill. Clear separation of layers by means of the scores plot was not possible. They found out that the information by the landfill manager and the results obtained did not correspond.

3.1.3. Conclusion

Mikhailov et al. [62] concluded that multivariate data analysis is an appropriate tool for ecological monitoring. They pointed out that chemometric methods provide the possibility to explore the structure of waste disposal by identification of specific areas.

3.2. Partial Least Square Regression (PLS1)

3.2.1. Objective of the study

The verification of compost quality has to be monitored consistently. However this is time-consuming and laborious. Due to the fact that NIR is a simple, accurate and fast technique used for routine analysis Michel et al. [42] hypothesised that NIR could be used for parameter prediction. The objective of the study was to use NIR spectroscopy to determine chemical and biological properties.

3.2.2. Method of evaluation and results

The first step was to define compost quality. Michel et al. [42] defined compost quality by C and N contents, suppression of pathogens, stability/ maturity and biological parameters, especially organic carbon (C_{org}), total N (N_t), C:N ratio, age, microbial biomass (C_{mic}), C_{mic}:C_{org}, basal respiration, enzymatic activity and suppression of plant disease. Spectroscopic data from 98 composts samples as well as the mentioned conventional parameters were collected. Fundamental relations between two matrices can be found by means of PLS1. Michel et al. [42] applied a PLS1 to express conventional parameters by spectral data. They designed for each conventional parameter a PLS1. Table 2 summarises the collected data and results obtained by Michel et al. [42]. The standard error of cross-validation (SECV) and the coefficient of determination (r^2) indicate the quality of prediction. The SECV provides information on the prediction error, r^2 demonstrates the quality of correlation. Composting age and basal respiration show the highest r^2. The specific enzymatic activity and the suppressive effect show the lowest r^2. It should be emphasised that biological tests that are carried out with the original wet compost are more susceptible to interferences due to the heterogeneity of the material. Michel et al. [42] concluded that especially compost age and basal respiration are clearly reflected by the NIR spectrum and feature the best results. By contrast, the specific enzyme activity and suppressive effects show the worst prediction results. The assigned correlations are illustrated in the paper [42].

	n	Mean	Range	Outliers removed	SECV	r^2
Age [d]	98	183.6	82.0 - 268.0	6	16.7	0.82
C_{org} content [%]	97	26.0	16.4 - 41.5	5	2.32	0.77
N_t content [%]	97	1.4	1.0 - 2.1	4	0.11	0.67
C:N ratio	97	18.2	12.2 - 29.1	4	1.51	0.71
C_{mic} [$\mu g\ g^{-1}$]	98	4986	774 - 8587	5	954	0.68
C_{mic}:C_{org} [$mgCmicgCorg^{-1}$]	97	18.6	4.0 - 29.4	4	4.00	0.63
Basal respiration [$\mu g\ C\ g^{-1}\ d^{-1}$]	47	574.8	252.0 - 966.0	2	49.2	0.88
qCO_2 [μgCO_2-C mg $C_{mic}^{-1}\ d^{-1}$]	47	9.7	4.2 - 17.1	1	1.98	0.83
Hydrolysis of fluorescein diacetate (FDA-HR) [$\mu g\ g^{-1}h^{-1}$]	98	517.9	256.0 - 879.0	5	74.7	0.75
Specific enzyme activity [$\mu gFDA\ mgC_{mic}^{-1}h^{-1}$]	98	118.7	48.6 - 370.9	6	48.6	0.49
Suppression 5‰ (rating) [%]	98	57.3	8.0 - 101.0	2	19.3	0.71
Suppression 5‰ (fresh weight) [%]	98	59.1	14.0 - 103.0	3	18.7	0.47

Table 2. Excerpt of table 1 and 2 by Michel et al. [42], SECV = standard error of cross-validation, r^2 = the coefficient of determination

3.2.3. Conclusion

Michel et al. [42] concluded that NIR spectroscopy was a capable method to predict various chemical and biological parameters using PLS regression. They believe NIR spectroscopy to be capable of monitoring compost quality.

3.3. Cluster analysis (CA)

3.3.1. Objective of the study

The objective of the study by Ros et al. [27] was to find out the long-term effects of composts on soil microbial communities. Different types of compost were applied over a period of 12 years. DNA was extracted by Ros et al. [27] from differently treated soils. The microbial community was described by polymerase chain reaction coupled with denaturing gradient gel electrophoresis (PCR-DGGE). They used multivariate data analysis to show the differences or similarities of microbial communities using DGGE data.

3.3.2. Method of evaluation and results

A polymerase chain reaction coupled with denaturing gradient gel electrophoresis (PCR-DGGE) was performed to characterize the microbial community. In Fig. 4 a DGGE fingerprint is shown. For the interpretation of such fingerprints statistical tools are necessary. DGGE data were converted into a binary system for cluster analysis (Fig. 4). As mentioned above, cluster analysis visualises the similarity between the samples in a dendrogram.

Ros et al. [27] show the cluster analysis of the DGGE profiles of 16S rDNA from the whole bacterial community. The cluster analysis illustrates the segregation of two soil groups. The clusters are caused by two different amendments. One cluster comprises the soil with compost and nitrogen application, the second cluster represents the soil with amendment of different composts (compost + nitrogen as mineral fertiliser).
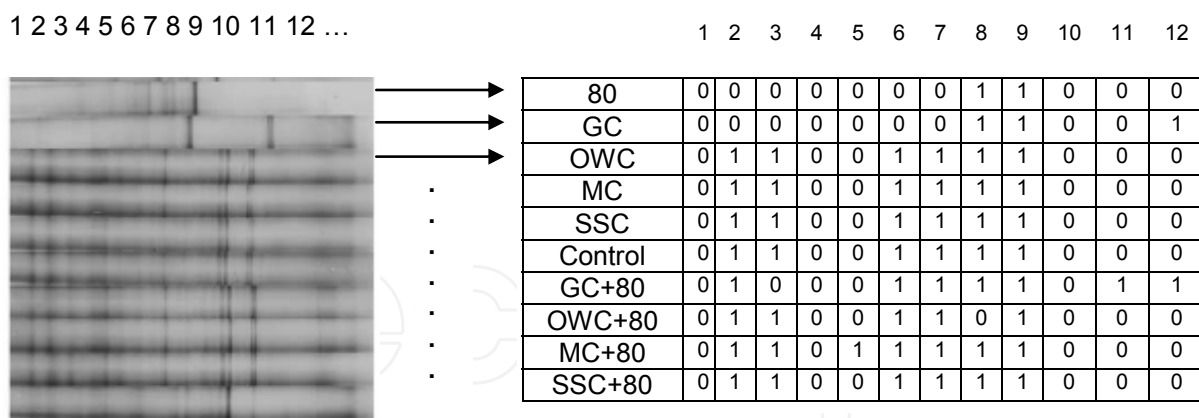
1 2 3 4 5 6 7 8 9 10 11 12 …

	1	2	3	4	5	6	7	8	9	10	11	12
80	0	0	0	0	0	0	0	1	1	0	0	0
GC	0	0	0	0	0	0	0	1	1	0	0	1
OWC	0	1	1	0	0	1	1	1	1	0	0	0
MC	0	1	1	0	0	1	1	1	1	0	0	0
SSC	0	1	1	0	0	1	1	1	1	0	0	0
Control	0	1	1	0	0	1	1	1	1	0	0	0
GC+80	0	1	0	0	0	1	1	1	1	0	1	1
OWC+80	0	1	1	0	0	1	1	0	1	0	0	0
MC+80	0	1	1	0	1	1	1	1	1	0	0	0
SSC+80	0	1	1	0	0	1	1	1	1	0	0	0

Figure 4. DGGE fingerprint and an example of a binary DGGE data matrix

3.3.3. Conclusion

Ros et al. [27] concluded that the differences between soils with compost with additional nitrogen fertiliser, and the second cluster comprising compost, control and mineral fertiliser soils are stronger than the influence of the different compost types. Furthermore they hypothesised that a certain microbial community inherent to the different composts is irrelevant after 12 years of compost application. Based on the cluster analyses of the PCR-

DGGE data, they concluded that the combined application of compost and nitrogen affected soil properties regarding microbial communities much more.

3.4. Soft independent modelling of class analogy (SIMCA)

3.4.1. Objective of the study

Malley et al. [8] used a portable near infrared (NIR) spectrometer to investigate changes of biogenic waste materials during composting. The idea of this study was to observe the composting process continuously in an easy and inexpensive way using NIR spectroscopy.

3.4.2. Method of evaluation and results

First of all many spectra were collected by Malley et al. [8]. The interpretation of spectral data requires experience in spectral interpretation. To provide rapid interpretation of the measured infrared spectra Malley et al. [8] applied the classification method SIMCA. The SIMCA model allows the assignment of a new sample to a defined class. A SIMCA model is always based on the PCAs of the various defined classes. Malley et al. [8] defined 3 different classes: raw manure (M), stockpiled manure (S) and manure compost (C). In the study 2 years of composting were observed (2000 and 2001). Figure 2 by Malley et al. [8] shows the scores plot of the PCA based on the spectral data of the three different classes in the year 2001. The PCA demonstrates a clear grouping of the 3 classes manure, stockpiled manure and manure compost.

Malley et al. [8] illustrated the results of the SIMCA by means of a Coomans plot. In figure 3 by Malley et al. [8] they show the Coomans plot for the investigations of 2001. The vertical and horizontal lines in the Coomans plot mark the 5 % level of significance. That means that 95 % of the samples that truly belong to this group are found within the line. Due to the fact that compost lies on the opposite side of the vertical line from the raw and stockpiled samples Malley et al. [8] concluded that compost is significantly different from the other two classes. The groups of raw manure and stockpiled manure are overlapping. Thus Malley et al. [8] concluded that they did not differ significantly. Nevertheless some raw samples were different. With these results Malley et al. [8] demonstrated that spectroscopic data and multivariate data analysis, especially SIMCA provides a sensitive analysis to differentiate between the products of stockpiles and compost.

3.4.3. Conclusion

Malley et al. [8] concluded that NIR spectroscopy and the multivariate data analysis method SIMCA can be a rapid, inexpensive method for assessing a composting process.

4. Critical discussion of multivariate statistical methods

In fact there are some statistical restrictions, which cannot be solved easily. The simple situation starts with the general linear model. This model usually has a character variable y depending on one or more predictor variables $x_1, x_2, …, x_k$:

In case of cross-classified two-way analysis of variance (equal subclass numbers):

$$y_{ijk} = \mu + a_i + b_j + w_{ij} + e_{ijk}, (i = 1, \quad , a; j = 1, \ldots, b; k = 1, \ldots, n) \qquad (1)$$

μ is the general mean, a_i are the main effects of factor A, b_j are the main effects of factor B, w_{ij} are the interactions between A_i and B_j, e_{ijk} are the random error terms.

In case of multiple linear regression:

$$y_j = \beta_0 + \beta_1 x_{1j} + \beta_2 x_{2j} + \ldots + \beta_k x_{kj} + e_j, (j = 1, \ldots, n), \qquad (2)$$

y_j is the j-th value of y depending on the j-th values $x_{1j}, \ldots x_{kj}$;

e_j are error terms with $E(e_j) = 0$, $var(e_j) = \sigma^2$ (for all j), $cov(e_{j'}, e_j) = 0$ for $j' \neq j$

The simple case assumes a linear dependency. The statistical parameters (the model coefficients) of the model can be estimated, y can be estimated for given values $x_1, \ldots x_k$. Assuming that the e_j are normally distributed, confidence intervals can be calculated for each model coefficient and finally tests of hypotheses about the model coefficients can be performed. By this procedure each variable can be tested whether its influence on the variable y is significantly different from 0 or not. The type I and type II error can be stated. Furthermore optimal designs for the experiments and surveys can be calculated [89]. Several assumptions are typically made regarding the distribution of the populations and regarding homoscedasticity. Furthermore the problem of extreme values and outliers respectively is critical, especially in environmental measurements. Increasing the number of regressors and factors respectively also increases the error terms.

For some univariate models robust and powerful alternatives regarding the distribution assumptions and regarding homoscedasticity [90-92] already exist. In the case of cross classification there is still no satisfying, powerful alternative. Many multiple regressors methods (multiple regression models, logistic regression models, discriminant analysis, cross classification models) need independent variables.

In chemometrics some of these problems are highly relevant. Usually the number of regressor variables exceeds the number of samples, which excludes most of the common oligovariate models. Many of the regressor variables are highly collinear. Due to these reasons dimension reduction methods are used such as correspondence analysis or factor analysis. The new factors in the latter are strictly independent from one another and can therefore be used in conventional models. There are several possibilities to extract these factors, like Principal Components or Maximum Likelihood. A possibility to model discrete variables is the classification by means of cluster analysis. These clusters can be tested later by contingency tables. Both steps (factor analysis and cluster analysis) lead to descriptive variables of the data set. Just as all descriptive methods in statistics they do not serve as tests against hypothesis of pure chance. There is no risk assessment of the results. Testing of the new descriptive variables implies the understanding of these new variables. By loading the original variables onto the new variables sometimes the interpretation can be done easily. Then models with these variables can be established (PCR or PLS-R) with several quality

parameters (e.g. correlation coefficient). A test of significance for the cross-validated r² was performed by Wakeling and Morris [93]. In this paper critical values of r² occurring just by chance alone are tabulated for one to three dimensional models at a significance level of 5 % based on Monte Carlo simulations. A comparable method was used by Stahle and Wold [94] to develop a polynomial approximation of the test statistic for the two-class problem and the number of objects, the number of variables, the percentage variance explained by the first component in X and the percentage of missing values.

$$cvd/sd = \sqrt{PRESS/RSS} \qquad (3)$$

cvd: cross-validated deviances
sd: standard deviation
PRESS: prediction error sum of squares
RSS: residual sum of squares

Unfortunately the definition of hypothesis regarding the regression coefficients still refers to the new components and provides no results regarding the original variables. There is no statistical possibility to prove whether the extraction method is optimal. Other methods of dimension reduction are already in use (e.g. Boosting, Random forest). Robust alternatives for PLS-R are also available [95].

As long as there are no satisfying testing routines, the results of the presented multivariate methods have to be interpreted very carefully. There is an inherent risk of over-interpretation, especially when using descriptive methods such as PCA or cluster analysis. There is no definition of the error probability of the results. That means whatever interpretation of the picture is done, it could be just pure coincidence and there is no information about the risk. The only possibility to overcome these problems would be to analyse a large number of samples and in case of regression models to validate these models.

5. Summary

In waste management research and practice often huge data sets for statistical evaluation are required to verify the findings. This request concerns both the natural scientific and the logistic field of waste management. Huge data sets can be generated on the one hand by vast numbers of investigated parameters and samples and on the other hand by modern analytical methods such as spectroscopic, chromatographic methods or thermal analysis.

Multivariate data analysis can help to explore data structures of the investigated samples. Another advantage is that the results can be displayed graphically. Furthermore, validated models can serve as adequate evaluation tools for practical application. Different software types are offered to develop such evaluation tools.

In this study the most important multivariate data analysis methods applied in waste management were described in detail and documented by a literature review. It could be demonstrated that Principal Component Analysis (PCA) and Partial Least Square

Regression (PLS-R) are the most applied methods in waste management. PCA was used to find hidden data structures, groupings and interrelationships of data. In most cases PLS-R was applied to predict parameters using new analytical instruments that allow faster and cheaper analyses.

In general it can be stated that multivariate data analysis was successfully applied in all experiments. Several authors compared different multivariate methods to determine which one provided the best results. Depending on the data set and the question to be answered the appropriate method must be identified.

Author details

K. Böhm[*]
Institute of Waste Management, Department of Water, Atmosphere and Environment, University of Natural Resources and Life Sciences, Vienna, Austria

E. Smidt and J. Tintner
Institute of Wood Science and Technology, Department of Material Sciences and Process Engineering, University of Natural Resources and Life Sciences, Vienna, Austria

6. References

[1] Zbytniewski R, Buszewski B (2005) Characterization of natural organic matter (NOM) derived from sewage sludge compost. Part 2: Multivariate techniques in the study of compost maturation. Bioresour. Technol. 96: 479-484.

[2] Albrecht R, Joffre R, Gros R, Le Petit J, Terrom G, Périssol C (2008) Efficiency of near-infrared reflectance spectroscopy to assess and predict the stage of transformation of organic matter in the composting process. Bioresour. Technol. 99: 448-455.

[3] Campitelli P, Ceppi S (2008) Chemical, physical and biological compost and vermicompost characterization: A chemometric study. Chemometrics Intell. Lab. Syst. 90: 64-71.

[4] Gil MV, Calvo LF, Blanco D, Sánchez ME (2008) Assessing the agronomic and environmental effects of the application of cattle manure compost on soil by multivariate methods. Bioresour. Technol. 99: 5763-5772.

[5] Grigatti M, Cavani L, Ciavatta C (2007) A multivariate approach to the study of the composting process by means of analytical electrofocusing. Waste Manage. 27: 1072-1082.

[6] Hansson M, Nordberg A, Mathisen B (2003) On-line NIR monitoring during anaerobic treatment of municipal solid waste. Water Sci. Technol. 48: 9-13.

[7] LaMontagne MG, Michel Jr FC, Holden PA, Reddy CA (2002) Evaluation of extraction and purification methods for obtaining PCR-amplifiable DNA from compost for microbial community analysis. J. Microbiol. Meth. 49: 255-264.

---

[*] Corresponding Author

[8] Malley DF, McClure C, Martin PD, Buckley K, McCaughey WP (2005) Compositional analysis of cattle manure during composting using a field-portable near-infrared spectrometer. Communications in Soil Science and Plant Analysis 36: 455-475.

[9] Nicolas J, Romain AC, Wiertz V, Maternova J, Andre P (2000) Using the classification model of an electronic nose to assign unknown malodours to environmental sources and to monitor them continuously. Sensors and Actuators, B: Chemical 69: 366-371.

[10] Planquart P, Bonin G, Prone A, Massiani C (1999) Distribution, movement and plant availability of trace metals in soils amended with sewage sludge composts: Application to low metal loadings. Sci. Total Environ. 241: 161-179.

[11] Romain AC, Godefroid D, Kuske M, Nicolas J (2005) Monitoring the exhaust air of a compost pile as a process variable with an e-nose. Sensors and Actuators, B: Chemical 106: 29-35.

[12] Smidt E, Meissl K, Schwanninger M, Lechner P (2008) Classification of waste materials using Fourier transform infrared spectroscopy and soft independent modeling of class analogy. Waste Manage. 28: 1699-1710.

[13] Termorshuizen AJ, van Rijn E, van der Gaag DJ, Alabouvette C, Chen Y, Lagerlöf J, et al. (2006) Suppressiveness of 18 composts against 7 pathosystems: Variability in pathogen response. Soil Biol. Biochem. 38: 2461-2477.

[14] Böhm K, Smidt E, Tintner J (2011) Modelled on Nature - Biological Processes in Waste Management. In: Kumar S, editor. Integrated Waste Management. Rijeka, Croatia: InTech. p. 153-178.

[15] Smidt E, Tintner J, Böhm K, Binner E (2011) Transformation of biogenic waste materials through anaerobic digestion and subsequent composting of the residues – A case study. Dyn. Soil Dyn. Plant 5: 63-69.

[16] Smidt E, Meissl K, Tintner J, Binner E (2008) Influence of Input Materials and Composting Operation on Humification of Organic Matter. In: Hao X, editor. Dyn. Soil Dyn. Plant. Special Issue 1: p. 50-59.

[17] Smidt E, Böhm K, Schwanninger M (2011) The Application of FT-IR Spectroscopy in Waste Management. In: G.S. N, editor. Fourier Transforms - New Analytical Approaches and FTIR Strategies. Rijeka, Croatia: InTech. p. 405-430.

[18] Bianchi G, Celeste G, Palmiotto M, Davoli E (2010) Source identification of odours and VOCs from a composting plant by multivariate analysis of trace volatile organic compounds. Chemical Engineering Transactions 23: 279-284.

[19] Biasioli F, Aprea E, Gasperi F, Mark TD (2009) Measuring odour emission and biofilter efficiency in composting plants by proton transfer reaction-mass spectrometry. Water Sci. Technol. 59: 1263-1269.

[20] Tognetti C, Mazzarino MJ, Laos F (2011) Comprehensive quality assessment of municipal organic waste composts produced by different preparation methods. Waste Manage. 31: 1146-1152.

[21] Vergnoux A, Guiliano M, Le Dréau Y, Kister J, Dupuy N, Doumenq P (2009) Monitoring of the evolution of an industrial compost and prediction of some compost properties by NIR spectroscopy. Sci. Total Environ. 407: 2390-2403.

[22] He XS, Xi BD, Jiang YH, He LS, Li D, Pan HW, et al. (2012) Structural transformation study of water-extractable organic matter during the industrial composting of cattle manure. Microchem. J. in press: DOI: 10.1016/j.microc.2012.1006.1004.

[23] Geladi P, Sethson B, Nystro?m J, Lillhonga T, Lestander T, Burger J (2004) Chemometrics in spectroscopy: Part 2. Examples. Spectrochimica Acta - Part B Atomic Spectroscopy 59: 1347-1357.

[24] Bustamante MA, Suarez-Estrella F, Torrecillas C, Paredes C, Moral R, Moreno J (2010) Use of chemometrics in the chemical and microbiological characterization of composts from agroindustrial wastes. Bioresour. Technol. 101: 4068-4074.

[25] Franke-Whittle IH, Knapp BA, Fuchs J, Kaufmann R, Insam H (2009) Application of COMPOCHIP Microarray to Investigate the Bacterial Communities of Different Composts. Microbial Ecol. 57: 510-521.

[26] Innerebner G, Knapp B, Vasara T, Romantschuk M, Insam H (2006) Traceability of ammonia-oxidizing bacteria in compost-treated soils. Soil Biol. Biochem. 38: 1092-1100.

[27] Ros M, Klammer S, Knapp B, Aichberger K, Insam H (2006) Long-term effects of compost amendment of soil on functional and structural diversity and microbial activity. Soil Use Manage. 22: 209-218.

[28] Ros M, Pascual JA, Garcia C, Hernandez MT, Insam H (2006) Hydrolase activities, microbial biomass and bacterial community in a soil after long-term amendment with different composts. Soil Biol. Biochem. 38: 3443-3452.

[29] Sun W, Huang GH, Zeng GM, Qin XS, Sun XL (2009) A stepwise-cluster microbial biomass inference model in food waste composting. Waste Manage. 29: 2956-2968.

[30] Sun W, Huang GH, Zeng GM, Qin XS, Yu H (2011) Quantitative effects of composting state variables on C/N ratio through GA-aided multivariate analysis. Sci. Total Environ. 409: 1243-1254.

[31] Xiao Y, Zeng GM, Yang ZH, Ma YH, Huang C, Shi WJ, et al. (2011) Effects of Continuous Thermophilic Composting (CTC) on Bacterial Community in the Active Composting Process. Microbial Ecol. 62: 599-608.

[32] Chikae M, Ikeda R, Kerman K, Morita Y, Tamiya E (2006) Estimation of maturity of compost from food wastes and agro-residues by multiple regression analysis. Bioresour. Technol. 97: 1979-1985.

[33] Suehara K, Yano T (2004) Bioprocess monitoring using near-infrared spectroscopy. Advances in biochemical engineering/biotechnology 90: 173-198.

[34] Capriel P, Ebertseder T, Popp L, Gutser R (1999) IR-spectroscopy: A rapid method to estimate the nitrogen effect and relevant parameters of biocomposts. IR-Spektroskopie: Eine Schnellmethode zur Prognose der N-Wirkung und relevanten Parameter von Biokomposten 162: 149-153.

[35] Fujiwara T, Murakami K (2007) Application of near infrared spectroscopy for estimating available nitrogen in poultry manure compost. Soil Science and Plant Nutrition 53: 102-107.

[36] Huang G, Han L, Liu X (2007) Rapid estimation of the composition of animal manure compost by near infrared reflectance spectroscopy. J. Near Infrared Spectrosc. 15: 387-394.

[37] Huang G, Han L, Yang Z, Wang X (2008) Evaluation of the nutrient metal content in Chinese animal manure compost using near infrared spectroscopy (NIRS). Bioresour. Technol. 99: 8164-8169.

[38] Ko HJ, Choi HL, Park HS, Lee HW (2004) Prediction of heavy metal content in compost using near-infrared reflectance spectroscopy. Asian-Australasian Journal of Animal Sciences 17: 1736-1740.

[39] Ludwig B, Schmilewski G, Terhoeven-Urselmans T (2006) Use of near infrared spectroscopy to predict chemical parameters and phytotoxicity of peats and growing media. Scientia Horticulturae 109: 86-91.

[40] Meissl K, Smidt E, Schwanninger M (2007) Prediction of humic acid content and respiration activity of biogenic waste by means of Fourier transform infrared (FTIR) spectra and partial least squares regression (PLS-R) models. Talanta 72: 791-799.

[41] Meissl K, Smidt E, Schwanninger M, Tintner J (2008) Determination of humic acids content in composts by means of near- And mid-infrared spectroscopy and partial least squares regression models. Appl. Spectrosc. 62: 873-880.

[42] Michel K, Bruns C, Terhoeven-Urselmans T, Kleikamp B, Ludwig B (2006) Determination of chemical and biological properties of composts using near infrared spectroscopy. J. Near Infrared Spectrosc. 14: 251-259.

[43] Morimoto S, McClure WF, Crowell B, Stanfield DL (2003) Near infrared technology for precision environmental measurements: Part 2. Determination of carbon in green grass tissue. J. Near Infrared Spectrosc. 11: 257-267.

[44] Zvomuya F, Helgason BL, Larney FJ, Janzen HH, Akinremi OO, Olson BM (2006) Predicting phosphorus availability from soil-applied composted and non-composted cattle feedlot manure. Journal of Environmental Quality 35: 928-937.

[45] Galvez-Sola L, Moral R, Perez-Murcia MD, Perez-Espinosa A, Bustamante MA, Martinez-Sabater E, et al. (2010) The potential of near infrared reflectance spectroscopy (NIRS) for the estimation of agroindustrial compost quality. Sci. Total Environ. 408: 1414-1421.

[46] Mohajer A, Tremier A, Barrington S, Martinez J, Teglia C, Carone M (2009) Microbial oxygen uptake in sludge as influenced by compost physical parameters. Waste Manage. 29: 2257-2264.

[47] Sharma HSS, Kilpatrick M, Lyons G, Sturgeon S, Archer J, Moore S, et al. (2005) Visible and near-infrared calibrations for quality assessment of fresh phase I and II mushroom (Agaricus bisporus) compost. Appl. Spectrosc. 59: 1399-1405.

[48] Suehara KI, Nakano Y, Yano T (2001) Simultaneous measurement of carbon and nitrogen content of compost using near infrared spectroscopy. J. Near Infrared Spectrosc. 9: 35-41.

[49] Galvez-Sola L, Morales J, Mayoral AM, Marhuenda-Egea FC, Martinez-Sabater E, Perez-Murcia MD, et al. (2010) Estimation of phosphorus content and dynamics during composting: Use of near infrared spectroscopy. Chemosphere 78: 13-21.

[50] Ecke H (2003) Sequestration of metals in carbonated municipal solid waste incineration (MSWI) fly ash. Waste Manage. 23: 631-640.

[51] Mostbauer P, Lechner P, Meissl K (2007) Langzeitverhalten von MVA-Reststoffen - Evaluierung von Testmethoden. Vienna: facultas.wuv. 158 p.

[52] Scaglia B, Adani F (2008) An index for quantifying the aerobic reactivity of municipal solid wastes and derived waste products. Science of the Total Environment 394: 183-191.

[53] Smidt E, Böhm K, Tintner J (2010) Application of various statistical methods to evaluate thermo-analytical data of mechanically-biologically treated municipal solid waste. Thermochim. Acta 501: 91-97.

[54] Abouelwafa R, Ait Baddi G, Souabi S, Winterton P, Cegarra J, Hafidi M (2008) Aerobic biodegradation of sludge from the effluent of a vegetable oil processing plant mixed with household waste: Physical-chemical, microbiological, and spectroscopic analysis. Bioresour. Technol. 99: 8571-8577.

[55] Abouelwafa R, Amir S, Souabi S, Winterton P, Ndira V, Revel JC, et al. (2008) The fulvic acid fraction as it changes in the mature phase of vegetable oil-mill sludge and domestic waste composting. Bioresour. Technol. 99: 6112-6118.

[56] He XS, Xi BD, Wei ZM, Guo XJ, Li MX, An D, et al. (2011) Spectroscopic characterization of water extractable organic matter during composting of municipal solid waste. Chemosphere 82: 541-548.

[57] Böhm K, Smidt E, Binner E, Schwanninger M, Tintner J, Lechner P (2010) Determination of MBT-waste reactivity - An infrared spectroscopic and multivariate statistical approach to identify and avoid failures of biological tests. Waste Manage. 30: 583-590.

[58] Biasioli F, Gasperi F, Odorizzi G, Aprea E, Mott D, Marini F, et al. (2004) PTR-MS monitoring of odour emissions from composting plants. Int. J. Mass Spectrom. 239: 103-109.

[59] De Rosa E, Rubel D, Tudino M, Viale A, Lombardo RJ (1996) The leachate composition of an old waste dump connected to groundwater: Influence of the reclamation works. Environmental Monitoring and Assessment 40: 239-252.

[60] Durmusoglu E, Yilmaz C (2006) Evaluation and temporal variation of raw and pre-treated leachate quality from an active solid waste landfill. Water, Air, and Soil Pollution 171: 359-382.

[61] Kylefors K (2003) Evaluation of leachate composition by multivariate data analysis (MVDA). J. Environ. Manage. 68: 367-376.

[62] Mikhailov EV, Tupicina OV, Bykov DE, Chertes KL, Rodionova OY, Pomerantsev AL (2007) Ecological assessment of landfills with multivariate analysis - A feasibility study. Chemometrics Intell. Lab. Syst. 87: 147-154.

[63] Olivero-Verbel J, Padilla-Bottet C, De la Rosa O (2008) Relationships between physicochemical parameters and the toxicity of leachates from a municipal solid waste landfill. Ecotox. Environ. Safe. 70: 294-299.

[64] Smidt E, Tintner J, Meissl K (2007) New approaches in landfill assessment. In: Velinni AA, editor. Landfill Research Trends. New York: Nova Science Publisher, Inc. p. 291.

[65] Srivastava SK, Ramanathan AL (2008) Geochemical assessment of groundwater quality in vicinity of Bhalswa landfill, Delhi, India, using graphical and multivariate statistical methods. Environmental Geology 53: 1509-1528.

[66] Smidt E, Böhm K, Tintner J (2011) Monitoring and assessment of landfills using simultaneous thermal analysis. Sustain. Environ. Res. 21: 247-252.

[67] Diener S, Andreas L, Herrmann I, Ecke H, Lagerkvist A (2010) Accelerated carbonation of steel slags in a landfill cover construction. Waste Manage. 30: 132-139.

[68] Pablos MV, Fernandez C, Babin MD, Navas JM, Carbonell G, Martini F, et al. (2009) Use of a novel battery of bioassays for the biological characterisation of hazardous wastes. Ecotox. Environ. Safe. 72: 1594-1600.

[69] Tintner J, Klug B (2011) Can vegetation indicate landfill cover features? Flora 206: 559-566.

[70] van Praagh M, Persson KM, Karlsson P (2009) Potential environmental impacts of using refuse derived material for landfill capping. Waste Manage. Res. 27: 471-488.

[71] Ecke H, Bergman A, Lagerkvist A (1998) Multivariate data analysis (MVDA) in landfill research. Journal of Solid Waste Technology and Management 25: 33-39.

[72] Jean G, Fruget JF (1994) Comparison of ecotoxicological and physico-chemical data by use of multivariate analyses and graphical displays. Chemosphere 28: 2249-2267.

[73] El-Sheikh MA, Al-Sodany YM, Eid EM, Shaltout KH (2012) Ten years primary succession on a newly created landfill at a lagoon of the Mediterranean Sea (Lake Burullus RAMSAR site). Flora 207: 459-468.

[74] Kim KD, Lee EJ, Cho KH (2004) The plant community of Nanjido, a representative nonsanitary landfill in South Korea: Implications for restoration alternatives. Water, Air, and Soil Pollution 154: 167-185.

[75] Tesar M, Prantl R, Lechner P (2007) Application of FT-IR for assessment of the biological stability of landfilled municipal solid waste (MSW) during in situ aeration. Journal of Environmental Monitoring 9: 110-118.

[76] Smidt E, Meissl K, Tintner J, Ottner F (2010) Interferences of carbonate quantification in municipal solid waste incinerator bottom ash: Evaluation of different methods. Environ. Chem. Lett. 8: 217-222.

[77] Huber-Humer M, Tintner J, Bohm K, Lechner P (2011) Scrutinizing compost properties and their impact on methane oxidation efficiency. Waste Manage. 31: 871-883.

[78] Smidt E, Böhm K, Tintner J (2010) Evaluation of old landfills - A thermoanalytical and spectroscopic approach. J. Environ. Monitor. 13: 362-369.

[79] Lawrence TA, Boutwell GP (1990) Predicting stratigraphy at landfill sites using electromagnetics. In: Landva A, Knowles DG, editors. Geotechnics of Waste Fills-Theory and Practice; 10-13. September 1989; Pittsburgh, PA, USA: ASTM Special Technical Publication; p. 31-40.

[80] Moreno-Santini V, Mansilla-Rivera I, García-Rodríguez O, Rodríguez-Sierra CJ (2012) A Pilot Study Determining Hair Arsenic and Lead Levels in Residents of a Community Established on a Former Landfill in Puerto Rico. B. Environ. Contam. Tox.: 1-5.

[81] Dahlén L, Vukicevic S, Meijer JE, Lagerkvist A (2007) Comparison of different collection systems for sorted household waste in Sweden. Waste Manage. 27: 1298-1305.

[82] Lin C, Wu EMY, Lee CN, Kuo SL (2011) Multivariate Statistical Factor and Cluster Analyses for Selecting Food Waste Optimal Recycling Methods. Environ. Eng. Sci. 28: 349-356.

[83] Cheng YW, Sung FC, Yang Y, Lo YH, Chung YT, Li KC (2009) Medical waste production at hospitals and associated factors. Waste Manage. 29: 440-444.

[84] Noori R, Abdoli MA, Ghazizade MJ, Samieifard R (2009) Comparison of Neural Network and Principal Component-Regression Analysis to Predict the Solid Waste Generation in Tehran. Iran. J. Public Health 38: 74-84.

[85] Esbensen K (2002) Multivariate Data Analysis - in practice. Esbjerg: Alborg University. 598 p.

[86] Velicer WF, Jackson DN (1990) Component Analysis versus Common Factor Analysis: Some issues in Selecting an Appropriate Procedure. Multivar. Behav. Res. 25: 1-28.

[87] ter Braak CJE, Verdonschot PEM (1995) Canonical correspondence analysis and related multivariate methods in aquatic ecology. Aquat. Sci. 57: 255-289.

[88] Brereton RG (2002) Chemometrics: Data analysis for the laboratory and chemical plant. Chichester, England: John Wiley & Sons Ltd. 489 p.

[89] Rasch D, Verdooren LR, Gowers JI (1999) Fundamentals in Design and Analysis of Experiments and Surveys. Munich, Vienna: Oldenbourg Wissenschaftsverlag GmbH. 253 p.

[90] Moder K (2007) How to keep the Type I Error Rate in ANOVA if Variances are Heteroscedastic. Austrian Journal of Statistics 36: 179-188.

[91] Moder K (2010) Alternatives to F-Test in One Way ANOVA in case of heterogeneity of variances (a simulation study). Psychological Test and Assessment Modeling 52: 343-353.

[92] Rasch D, Kubinger K, Moder K (2011) The two-sample t test: Pre-testing its assumptions does not pay off. Stat. Pap. 52: 219-231.

[93] Wakeling IN, Morris J (1993) A test of significance for partial least squares regression. J. Chemometr. 7: 291-304.

[94] Ståhle L, Wold S (1987) Partial least squares analysis with cross-validation for the two-class problem: a Monte Carlo study. J. Chemometr. 1: 185-196.

[95] Wakeling IN, Macfie HJH (1992) A robust PLS procedure. J. Chemometr. 6: 189-198.