rapport
IVL Svenska Miljöinstitutet AB

Analysis and Development of the Interpretation process in LCA

Mats Almemark, Charlotte Bjuggren, Jessica Granath, Jenny Olsson,
Jonas Röttorp och Lars-Gunnar Lindfors

B 1375
Stockholm, maj 2000

IVL

Organisation/Organization

IVL Svenska Miljöinstitutet AB
IVL Swedish Environmental Research Institute Ltd.

Adress/address

Box 21060
100 31 Stockholm

Telefonnr/Telephone

08-08-587 563 00

RAPPORTSAMMANFATTNING
Report Summary

Projekttitel/Project title

Anslagsgivare för projektet/
Project sponsor

Rapportförfattare/author

Mats Almemark, Charlotte Bjuggren, Jessica Granath, Jenny Olsson,
Jonas Röttorp och Lars-Gunnar Lindfors

Rapportens titel och undertitel/Title and subtitle of the report

Analysis and Development of the Interpretation process in LCA

Sammanfattning/Summary

The objective of this work is to study interpretation as a procedure to use the quantitative results of a life-cycle inventory to compare process alternatives with the aim to conclude, whether or not significant differences exist with regard to the studied issues (individual emissions or impact categories).

As a result of an introductory survey a procedure for quantitative interpretations is suggested, with data-quality scoring, statistical experimental planning, and multivariate data analysis as basic tools.
The procedure has been tested on a case study of treatment of paper packaging waste, either by material recycling or by energy recovery (incineration). The inventory of an earlier study has been used. With the aid of what is called a conceptual model five variables, which could be presumed to have an influence on the environmental impact of paper packaging waste treatment, were identified. The choice of technology, material recycling or energy recovery, was one of these variables. Subsequently 36 scenario calculations, organised in an experimental matrix, were performed. The result was interpreted with the multivariate techniques principal component analysis (PCA), partial least-square modelling (PLS), and uncertainty analysis. The multivariate analysis made it possible to isolate the influence of the variable "choice of technology" on the environmental impact of the system.

As a result of the study it is concluded that the interpretation procedure suggested in the introductory survey, i.e. construction of a conceptual model, sensitivity and uncertainty analysis with multivariate methods, and conclusions based on the results of principal component analysis and partial least-square models, can give easily surveyable descriptions of complicated decision-making situations in cases, where the environmental effects of technology changes depend on several pre-conditions. It is further concluded, that a systematic structuring of methodological choices and the use of factorial experimental designs to organise scenario calculations can minimise the necessary inventory work, and that Monte Carlo simulations in combination with multivariate evaluation and other statistical tests are helpful methods to determine whether or not observed differences between two cases are significant.

Nyckelord samt ev. anknytning till geografiskt område eller näringsgren /Keywords

Bibliografiska uppgifter/Bibliographic data

IVL Rapport/report B1375

Beställningsadress för rapporten/Ordering address

IVL, Publikationsservice, Box 21060, S-100 31 Stockholm
fax: 08-598 563 90, e-mail: publicationservice@ivl.se

Table of contents

Abstract

The objective of this work is to study interpretation as a procedure to use the quantitative results of a life-cycle inventory to compare process alternatives with the aim to conclude, whether or not significant differences exist with regard to the studied issues (individual emissions or impact categories).

As a result of an introductory survey a procedure for quantitative interpretations is suggested, with data-quality scoring, statistical experimental planning, and multivariate data analysis as basic tools.

The procedure has been tested on a case study of treatment of paper packaging waste, either by material recycling or by energy recovery (incineration). The inventory of an earlier study has been used. With the aid of what is called a conceptual model five variables, which could be presumed to have an influence on the environmental impact of paper packaging waste treatment, were identified. The choice of technology, material recycling or energy recovery, was one of these variables. Subsequently 36 scenario calculations, organised in an experimental matrix, were performed. The result was interpreted with the multivariate techniques principal component analysis (PCA), partial least-square modelling (PLS), and uncertainty analysis. The multivariate analysis made it possible to isolate the influence of the variable “choice of technology” on the environmental impact of the system.

As a result of the study it is concluded that the interpretation procedure suggested in the introductory survey, i.e. construction of a conceptual model, sensitivity and uncertainty analysis with multivariate methods, and conclusions based on the results of principal component analysis and partial least-square models, can give easily surveyable descriptions of complicated decision-making situations in cases, where the environmental effects of technology changes depend on several pre-conditions. It is further concluded, that a systematic structuring of methodological choices and the use of factorial experimental designs to organise scenario calculations can minimise the necessary inventory work, and that Monte Carlo simulations in combination with multivariate evaluation and other statistical tests are helpful methods to determine whether or not observed differences between two cases are significant.

1. Introduction

The concept of life-cycle interpretation is defined by ISO as:

”The phase of life-cycle assessment in which the findings of either the inventory analysis or the impact assessment, or both, are combined consistent with the defined goal and scope in order to reach conclusions and recommendations” (ISO 1997). The procedure of interpretation is further elaborated in the draft of ISO 14043. Here the objective of the procedure of interpretation is described as being ”to analyse and report results, reach conclusions, explain limitations and provide recommendations for a life-cycle inventory study or a life-cycle assessment study”.

A request to analyse, conclude and recommend presumes that there is a question to answer or a problem to solve. H. Baumann (1996) has investigated which purposes LCA was put to in Swedish companies. Some of her results are cited in table 1.

Table 1. LCA applications in Swedish companies in 1995 (data from Baumann 1996).

Number of applications	Type of application	Reference no.
14	Analysis of own product	1
n.d.	To learn about LCA	2
11	Product development	3
11	For external use (marketing, labelling...)	4
9	Process development and optimisation	5
9	Choice of suppliers and raw materials	6
5	In-training programmes	7
8	Analysis of line of business	8
6	To meet authorities´ demands	9

In 70 % of the quantified number of cases (applications no. 1, 3, 5, 6, and 8) LCA was used in an application that implies decision-making. In the applications no. 1 and 8 the decision is presumably to find out, if and where in a process chain it might be warranted to search for alternative methods, i.e. to decide where the environmental key issues are. In applications no. 3, 5, and 6 the decision is probably to make a choice between some existing alternatives. In the remaining 30 % of the cases the LCA must not necessarily result in a decision. The procedure of interpretation could then be limited to acquiring a knowledge of the structure and functions of the system and an explanation of the limitations of the results.

In this report we study interpretation as a procedure to use the quantitative results of a life-cycle inventory to compare process alternatives with the aim to conclude, whether or not significant differences exist with regard to the studied issues (individual emissions or impact categories). Valuation methods are not studied in this report.

In a draft for an ISO standard (Committee draft ISO/CD 14043-2) some guidelines for the interpretation process are suggested. These guidelines may be tabulated as in table 2.

Table 2. Suggested interpretation procedure (tabulation based on ISO/CD 140432-2).

1. Identification of significant environmental issues

1.1. Identification and structuring of four types of information:

1. l .l . Results from LCI and LCIA with information of its data quality

1.1.2. Methodological choices (e.g. allocation rules, system boundaries in the LCl)

1.1.3. Possible value system

1.1.4. Role and responsibilities of the stakeholders

1.2. Determination of the significant environmental issues:

1.2.1. Determination if data from LCI and LCIA are sufficient to meet the needs defined in goal and scope

1.2.2. Determination of the relative importance of the inputs and outputs.

1.3. Report on the results.

2. Evaluation

2.1. Completeness check

2.1.1. Determination if missing information are necessary according to the goal and scope of the study.

2.1.1.1. If unnecessary, record why.

2.1.1.2. If necessary, revise either LCI and LCIA or revise the goal and scope. Record how and why.

2.2. Sensitivity check (can be more or less detailed)

2.2.1. Choice of factors to include in the sensitivity analysis

2.2.2. Determination of the possible need of sensitivity analysis and the scope of it.

2.2.3. Performance of sensitivity analysis

2.2.3.1. Quantitative sensitivity analysis (two types to choose from)

2.2.4. Reporting on results from sensitivity test.

2.3. Consistency check, that is deciding

2.3.1. ...if regional and/or temporal differentiations have been consistently applied.

2.3.2. ...if allocation rules and system boundaries have been consistently applied to all production systems.

2.3.3. ...if a uniform differentiation between foreground and background processes has been used.

2.3.4. ...if the differences and variabilities among the quality and environmental relevance of LCA. indicators have been consistently considered.

2.3.5. ...if weighting has been carried out consistently and in accordance with the stated value/judgement system.

Table 2. Suggested interpretation procedure (tabulation based on ISO/CD 140432-2) (continued)

3. Conclusions, recommendations and reporting.

3.1. Reaching conclusions

3.1.1. Identification of the significant environmental issues.

3.1.2. Evaluation of the methodology and results for completeness,, sensitivity and consistency.

3.1.3. Check that conclusions are consistent with the requirements of the goal and scope of the study.

3.1.4. If above OK, report, otherwise loop back.

3.2. Recommendations

3.2.1. Determination if it is possible to make recommendations being logical and reasonable consequences of the conclusions.

3.2.2. Formulate recommendations.

3.2.3. Report.

3.3. Reporting

3.3.1. Report:
- values adopted
- decisions
- reasonings
- expert judgement

3.3.2. Results from the steps above.

4. Critical review

The goal of the introductory survey is to put the guidelines of table 2 (or some of them) in a concrete form adapted to quantitative interpretation.

1.1. Identification and structuring of information of data quality

The first item of table 2, 1.1.1. ”Results from LCI and LCIA with information of its data quality”, requires us to collect and structure information of data quality. Since interpretation in this report means quantitative comparison, we need data quality information in a form, which may be transformed into a statistical measure of uncertainty, e.g. a standard deviation or a minimum – maximum interval.

Data quality may be defined by a set of data quality indicators, DQIs, which may be both quantitative and qualitative by nature. Table 3 exemplifies DQIs suggested by various sources.

Table 3. Examples of data quality indicators, DQIs.

SETAC [a] quantitative	SETAC [a] qualitative	U.S. EPA [b]	Weidema and Wesnoes [c]	ISO 14040 [d]
Accuracy	Accessibility	Precision	Reliability	Precision
Bias	Applicability/ Suitability/ Compatibility	Data Collection Method and Limitations	Temporal Correlation	Time-related Coverage
Completeness	Comparability	Comparability	Geographic Correlation	Geographic Coverage
Data Distribution	Consistency	Completeness	Completeness	Completeness
Precision	Derived Models	Bias	Technological Correlation	Technology Coverage
Uncertainty	Identification of Anomalies	Acceptability		Uncertainty
	Peer Review	Referenced		Data source
	Representativeness	Representative		Representativeness
	Reproducibility			Reproducibility
	Stability			Consistency
	Transparency			

[a] SETAC 1994. [b] U.S. EPA 1995. [c] Weidema and Wesnoes 1995. [d] ISO 1997.

Weidema and Wesnoes turn their basically qualitative DQIs into a kind of semi-quantitative indicators by scoring them on a scale from 1 to 5 according to some established criteria. (In their system 1 denotes the highest and 5 the lowest quality). The result is a so-called pedigree matrix. For each DQI there are five possible quality levels defined. Kennedy et al. (1997) have pointed out that this is a way of converting the data quality description into a 1 x n vector. For each data element there are n numbers q_i describing the quality of that particular element. If instead of integers a continuous scale from 1 to 5 is introduced, where 1 corresponds to the worst possible quality and 5 to the best possible quality, a quality function may be introduced (Kennedy et al. 1997 and 1996):

$$Q = \sum_{i=1}^{n} q_i \qquad (1)$$

Equation (1) reduces the quality pedigree matrix and the quality 1 x n vector to a single aggregated number, Q. Since $1 \leq q_i \leq 5$, it follows, that $n \leq Q \leq 5n$. Using the SETAC quality indicators as an example $Q_{min} = 17$ and $Q_{max} = 5 \times 17 = 85$. Any intermediate quality value Q attained by a data element with a given quality vector

$(q_1, q_2 ..... q_{17})$ may be expressed as a percentage of the maximum attainable quality according to equation (2):

$$x = \frac{Q - Q_{min}}{Q_{max} - Q_{min}} \text{ x } 100 \qquad (2)$$

x = % attainable quality.

There are of course an infinite number of vectors corresponding to any given intermediate value of Q.

The percentage attainable quality may be translated into an aggregated data quality index with the help of table 4, given by Kennedy et al (1997):

Table 4. Transformation of percentage attainable quality into an aggregated data quality indices according to Kennedy et al. (1997).

Attainable Quality (x), %	Aggregated Data Quality Index (ADQI)
$0 \leq x < 12.5$	1
$12.5 \leq x < 25$	1.5
$25 \leq x < 37.5$	2
$37.5 \leq x < 50$	2.5
$50 \leq x < 62.5$	3
$62.5 \leq x < 75$	3.5
$75 \leq x < 87.5$	4
$87.5 \leq x < 100$	4.5
$x = 100$	5

In conclusion, provided that the LCA practitioner can define a set of applicable quality indicators, which are relevant to the goal and scope of the study, and provided that the practitioner can score them in a realistic way, each data element may be assigned an aggregated data quality index (ADQI, our notation) in the form of a single number.

The discussion above has implicitly assumed, that all data quality indicators have the same weight. It is however possible to apply different weights to the different quality scores. Each score q_i is multiplied by a weight factor w_i. Equation (1) then becomes:

$$Q = \sum_{i=1}^{n} w_i q_i \qquad (3)$$

The calculation of percent attainable quality, equation (2), is adjusted accordingly. The weight factors w_i like the quality scores q_i are real, positive numbers, not necessarily integers.

The next step is to transform the ADQI of each data element into an uncertainty range. If nothing is known about the distribution and the spread of the data element around the value found in the inventory, Kennedy et al.(1996) suggest the use of a probability density function known as the beta distribution. A beta distribution is described by four parameters, the upper and lower end points a and b of the possible range of values of the data element, and the shape parameters α and β, which define the shape of the distribution curve. The higher the values of α and β are, the sharper the distribution curve is, the lower the variance is, and the lower the probability is, that the data element assumes a value close to the end points. $\alpha = \beta$ means, that the distribution is symmetrical. The median is equal to the arithmetical mean. $\alpha \neq \beta$ means that the distribution is skewed. The median is not in the middle of the range of values from a to b.

If no other information about the range of possible values for a given data element is available, Kennedy et al. (1996) suggest the symmetrical beta distribution shown in table 5. The table transforms ADQIs into beta distribution parameters.

Table 5. Transformation of aggregated data quality indices to beta distributions as given by Kennedy et al. (1996) (baseline case).

ADQI	Beta distribution Shape parameters, α, β	Range endpoints, ± %
5	**5, 5**	**10**
4.5	**4, 4**	**15**
4	**3, 3**	**20**
3.5	**2, 2**	**25**
3	**1, 1**	**30**
2.5	**1, 1**	**35**
2	**1, 1**	**40**
1.5	**1, 1**	**45**
1	**1, 1**	**50**

Adopting the statistical methodology described above enables us to transform a quality description of a data element into a statistical measure of the uncertainty of the value of that data element. The shape parameters of the beta distribution define the variance, the range endpoints define the spread. This concludes the first interpretation step, namely to structure the results from the LCI with information of its data quality.

1.2 Identification and structuring of methodological choices and system choices

In the preceding section we have dealt with variables, which are data elements and are thus described by continuous real numbers. It is assumed, that the uncertainty of these variables can be described by probability functions. The result of an LCI calculation may, however, be influenced by variables which are not data elements. E.g. if two or more techniques are available to manufacture a studied product, the choice of technique is a variable, which influences the result of the LCI calculation and consequently the result and the interpretation of the LCA. This variable may mathematically be expressed by discrete numbers, like –1, 0 or +1, where each value signifies a specified technique. The value of the variable ”choice of technique” is restricted to a few exact numbers. There is no probability function associated with this variable. The uncertainty is rather in the relevance of each choice, the influence of this particular variable on the various environmental impact parameters, and the uncertainty introduced by neglecting or overlooking a possible choice of technique.

Likewise a methodological choice, like choice of allocation procedure, or the choice whether to use allocation or system expansion, is a variable, which may assume a few discrete numbers, and the uncertainty of which may not be described by ranges of values or standard deviations or probability functions.

To identify and structure the independent variables, which determine the environmental performance of a system, we may use the methods of statistical experimental planning. The first step is to set up a conceptual model of the system, i.e. basically a very simple scheme, which shows the function of the system and its inflows and outflows. The scheme is used as an aid to systematise the practitioners knowledge of the system and its technology into a list of independent parameters, which may be assumed to determine the performance of the system.

As a next step the goal and scope, the value system if any, and other prerequisites of the study are checked in order to exclude parameters which are irrelevant to the study, or the values of which may not be changed.

A conceptual model could in general terms look like figure 1.

Independent variables	**System**	**Outflows**
Techniques, X_1 ----->		=> Emissions $e_j \pm \delta e_j$
System boundaries, X_2 ---->		
Allocation principle, X_3 ----->		
Transport distance, X_4 ----->		=> Products $P_k \pm \Delta P_k$
	(Outflows, Consumptions) =	
	($f_i \pm \delta f_i$)(Inflows, Syst. var.)	
Inflows		=> Wastes $W_l \pm \Delta W_l$
Raw materials, $F_n \pm \Delta F_n$ =>		
Energies, $E_m \pm \Delta E_m$ =>		

Figure 1. General description of a conceptual model with four independent system variables.

A system, e.g. a manufacturing process, receives inflows in the form of raw material flows and m energy flows, and it produces k product flows, j emission flows and l waste flows. The numbers n, m, i, j, k and l are integers. The quantities F_n, E_m, e_j, P_k and W_l are measures of physical quantities and may thus have uncertainties and probability functions. (One exception would be the product flow of the functional unit, which by definition is an exact quantity). In an LCI calculation model, the mathematical function of the system is to calculate the outflows from the inflows and the system parameters. The last-mentioned parameters are symbolised by the functions f_i in figure 1. They may be emission factors, specific energy consumptions, process yields etc. They have uncertainties, and their range of values may be described by probability functions. The system parameters are also dependent on the independent variables X. In principle there is one set of system parameters ($f_i \pm \delta f_i$) for each set of X-values.

In figure 1 we have as an example assumed, that three variables, which are not data elements, have been identified as important parameters, namely the real system parameter "techniques" and the methodological parameters "choice of system boundaries" and "choice of allocation principle". In addition it is assumed that the goal and scope of the study requests the practitioner to study the influence of transport distance. The system and methodological variables X_1 to X_4 may for the purpose of a study be structured according to the methods of statistical experimental planning, i.e. they are varied between a few discrete levels, such as a few defined techniques, a few selected transport distances etc. (In this way the variable X_4 "transport distance" is transformed from a continuous to a discrete variable). If each variable is varied between two levels, denoted as –1 and +1, we would in an experimental study need a minimum of $2^4 = 16$ experiments to investigate the influence of the four variables (see for instance Box et al. 1978). In an LCI the experiments are replaced by scenario calculations. We would thus need a minimum of 16 scenario calculations in order to study the influence of the four variables X_1 to X_4 in a structured way, i.e. using a factorial experimental design.

Performing the scenario calculations according to a factorial experimental design at two levels for each of the four independent variables would yield a result, which may be tabulated as in table 6.

Table 6. Result matrix of a factorially designed scenario calculation.

	Independent variables (X variables)				**Results (dependent variables) (Y variables)**				
Scen. no.	$\mathbf{X_1}$	$\mathbf{X_2}$	$\mathbf{X_3}$	$\mathbf{X_4}$	$\mathbf{F_n}$	$\mathbf{E_m}$	$\mathbf{e_j}$	$\mathbf{P_k}$	$\mathbf{W_l}$
1	+1	+1	+1	+1	F_{n1}	E_{m1}	e_{j1}	P_{k1}	W_{l1}
2	-1	+1	+1	+1	F_{n2}	E_{m2}	e_{j2}	P_{k2}	W_{l2}
3	+1	-1	+1	+1	F_{n3}	E_{m3}	e_{j3}	P_{k3}	W_{l3}
…	…	…	…	…	…	…	…	…	…
16	-1	-1	-1	-1	F_{n16}	E_{m16}	e_{j16}	P_{k16}	W_{l16}

Table 6 contains 16 response (result) vectors of size 1 x p, where the number of elements p = n + m + j + k + l is the number of inflow and outflow parameters to and from the system. Altogether there are 16p response parameters. The interpretation of the result of the factorial scenario calculations requires us to analyse, how each of the 16p response parameters is influenced by the independent variables X_1 - X_4, and to detect and systematise any significant differences between the 16 scenarios. Even for a moderately large system influenced by a small number of independent variables, such an analysis comprises handling of a vast amount of apparently disparate data. The mathematical tool to do such an analysis is multivariate analysis, which comprises principal component analysis (PCA) and partial least-square models (PLS). (For a description of principal component analysis see for instance Chatfield and Collins 1980).

In a PCA model all parameters (both X and Y) are considered to be X-parameters. The result from a PCA provides information about co-relations between parameters, e.g. cluster formations, and also a coarse information about the relations between the independent and the dependent variables.

A PCA-model receives a percentage value showing the amount of variance in the parameters described by the model. It is desirable to achieve a value as high as possible. It is possible to get a 100 % explanation of the variance in the parameters with enough principal components. However it is not interesting to include too many components in the model because then the noise is included as well. The aim with a PCA is to attain a high percentage value with as few principal components as possible.

If we for the sake of clarity of explanation assume that the system of figure 1 and table 6 can be described by only three variables, Y_1, Y_2 and Y_3, then each observation (experiment) is represented by a position in the three-dimensional space. Several experiments result in a swarm of positions. The swarm is approximated by a vector using the least-squares method. This vector is called the first principal component and describes the greatest variance among the observations in the system. Another vector, perpendicular to the first, is also calculated. It describes the direction for the second greatest variance. Mathematically these vectors are linear combinations of the variables. This means, that each variable contributes to the two principal components with a pair of correlation coefficients.

The two calculated vectors will together form the slope of a two-dimensional plane in the three-dimensional space, see figure 2. The plane is fitted to minimise the sum of the quadrants on the distance from the objects to the plane. Thus the plane is the best fit to the results of the experiments.

The two principal components have reduced the dimensions from three to two. The swarm of observations is then projected on to the two-dimensional plane, se figure 2. This plane can be studied and the relations between different observations can be evaluated. The result of the calculations is that the system can be described in a two-dimensional space instead of a three-dimensional.

A PCA can be performed with more than three principal components. A multi-dimensional space can also be summarised to a two-dimensional planes. Several principal components can be calculated but often three or four are sufficient to describe the major part of the variance of a system.

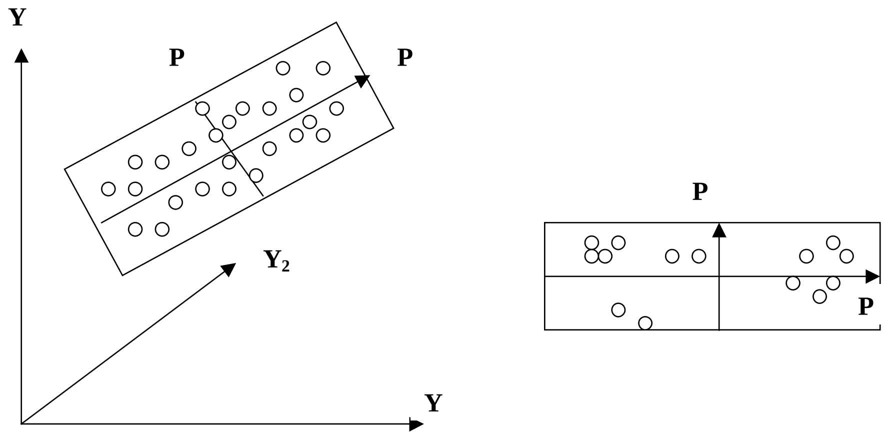


Figure 2. Principal component analysis for a set of experiments, where each result may be described by three parameters.

The first result of the PCA analysis is a graphic view of the interdependence of the X- and Y-parameters. If the above-mentioned correlation coefficients for each parameter is projected onto the plane defined by the pair of principal components, each parameter will be represented by a point in the plane. The co-ordinates of that point will be determined by the two correlation coefficients of the parameter. The distance and the direction to the point from the origin will describe how the corresponding parameter is influenced by the variation of the independent variables. E.g. a point close to the origin of the plane means that the parameter represented by that point is not or only to a small extent influenced by the independent variables, which drive the variations of the system.

In a PLS-model the variance in X- and Y-parameters is quantified. When there are a lot of Y-parameters usually a few are chosen to be included in the model. If the PCA analysis shows, that the variables are grouped in discernible clusters, one should select one or two Y-parameters to represent each cluster. The PLS model is a linear equation of the form shown in equation 4.

$$Y = \sum_{i} \gamma_i X_i + \varepsilon$$

γ = coefficient, ε = residual

A PLS connects the Y- and X-spaces. X-parameters of great importance to the Y-parameters are detected. The theory behind PLS-modelling may be further studied in an article by Geladi and Kowalski (1985).

When using multivariate methods the calculated models are an approximation to the original data set. The residuals are the distance from the calculated vectors to the observations. The principal components are calculated to minimise this distance, i.e. the residuals. The residuals represent variation in the data material not explained by the model. If the residuals are large that implies that the fit of the model is less good.

The residuals should be randomly distributed with means that the unexplained information in data should consist of pure noise. If a systematic pattern is shown it indicates that there still is some undescribed systematic variation in data. To check if the residuals are randomly distributed a number of different methods are used, like plots of observed against predicted values and normal probability residuals plots.

Multivariate methods are looking at variations in data. Like all evaluation methods it is important that the data quality is fairly good to be able draw valid conclusions from the evaluation. In an multivariate evaluation it is important to include all parameters that possibly could have an effect on the system. Any parameters that are found not to have an impact may be excluded from the data material throughout the modelling. Constructing a conceptual model is a good way to make sure that no important parameters in the system are forgotten.

Up to this point we have treated the results of table 6 as if they were exact numbers without uncertainties. In reality, of course, the precision and the accuracy of the values of theY-parameters are determined by the data quality of the inventory, i.e. the results suffer from uncertainties. As a consequence each experiment and each parameter will be represented by a cloud in the principal-component plane rather than by a sharp point.

Le Téno (Le Téno 1997) has studied principal component analysis as a tool to visualise data and as a support for decision-making. In a case study he compared data-base data for different electricity production methods, and he used projections of result vectors and parameter coefficients onto the same principal component plane. The result was a picture, which visualised how the individual production methods covaried with the individual emission parameters. With the help of the picture one could, at least qualitatively, find for instance the best compromise between low greenhouse gas emissions and low soil pollution.

In a pre-study to the project reported here Bjuggren et al. (1998) used a somewhat different approach to compare the results of scenario calculations for the production and consumption of milk and milk packages. They used an existing LCI-model to study the influence of four discrete X-variables, namely Technology Level, Cut-off, Allocation, and Data Type. Each variable was varied at two levels, and the calculation of the sixteen resulting scenarios was organised basically as described above in table 6. The result vectors were analysed with principal component analysis. A total of four principal components and three principal component planes were needed to describe the variance of the system. The independent and the dependent variables (i.e. the correlation coefficients of these variables) but not the result vectors were projected onto the three principal component planes. In this way the dependent variables (Y-variables) could be grouped into clusters, and the covariance, or lack of covariance, of the clusters with the X-variables could be visualised. The study was not carried on to a partial least-square model, nor was uncertainty of the data considered (deterministic study).

1.3 Conclusions of the introductory survey

Based on the preceding sections, the recommendations in table 7 may be formulated.

Table 7. Procedure for LCA studies with a quantitative interpretation phase.

1. Identification of significant issues.

1.1. Methodological choices. Based on the goal and and scope of the study and a technical knowledge of the system, set up a conceptual model of the system and identify the technical and methodological variables, the independent variables, which determine the performance of the system.

1.2. Results from LCI with information of its data quality. Select suitable data quality indicators and, during the inventory, try to obtain expert help to evaluate and score the quality of each data element. Calculate uncertainty ranges.

2. Evaluation

2.1. Completeness check. Determine if missing information, such as data gaps, data quality gaps, information gaps on technical and methodological choices, are crucial to the goal and scope of the study.

2.2. Sensitivity analysis. Determine if a sensitivity analysis, that is a study of the influence of identified technical and methodological variables, is necessary. If yes, design a factorial scenario calculation plan. Carry out the calculations in a deterministic way, i.e. without considering data uncertainty. Analyse the result with PCA and PLS. Determine the influence of the independent variables.

2.3. Uncertainty analysis. Determine whether or not an uncertainty analysis, i.e. replicate calculations of scenarios with varying values of selected data elements, is necessary. If yes, make replicate calculations of at least one experiment with selected Y-parameters, representative of identified clusters. Determine if the spread of the replicates is larger than the variance between the different scenarios.

3. Conclusions

3.1. Data quality check From the uncertainty analysis, determine whether the data quality is sufficient or not.

3.2. Reaching conclusions. If 3.1. is yes, conclude, that is determine whether or not there are significant differences between the scenarios, and the cause of such differences.

The procedure suggested in table 7 pertains in the first place to a LCA study intended to describe the effects of changed conditions on a system, or intended to compare two different systems producing the same function, not so much to an inventory study intended for environmental labelling (like Type III). Linear correlations, such as PLS, is a permissible approximation in a LCA model, when changes are small.

In order to test some of the procedures suggested in table 7, especially points 1.1, 2.2 and 2.3 (conceptual model and multivariate data analysis) and 3.1 – 3.2 (reaching conclusions) a test study using data from a published case study, has been carried out.

2. The selected case study

2.1 System description

Since the goal and scope of the study reported here is to try out new interpretation methods, we have used a published LCA study as a test case, and accepted the inventory of that study. Whether the data are complete and accurate or not is less relevant to our purpose. We use the data as a model. The goal and scope of the selected LCA case was to compare the environmental consequences of different methods to dispose of paper packaging waste (Finnveden et al. 1994). Particularly the question whether material recycling is better for the environment than incineration was addressed. The systems for material recycling and for incineration are outlined in figure 3.

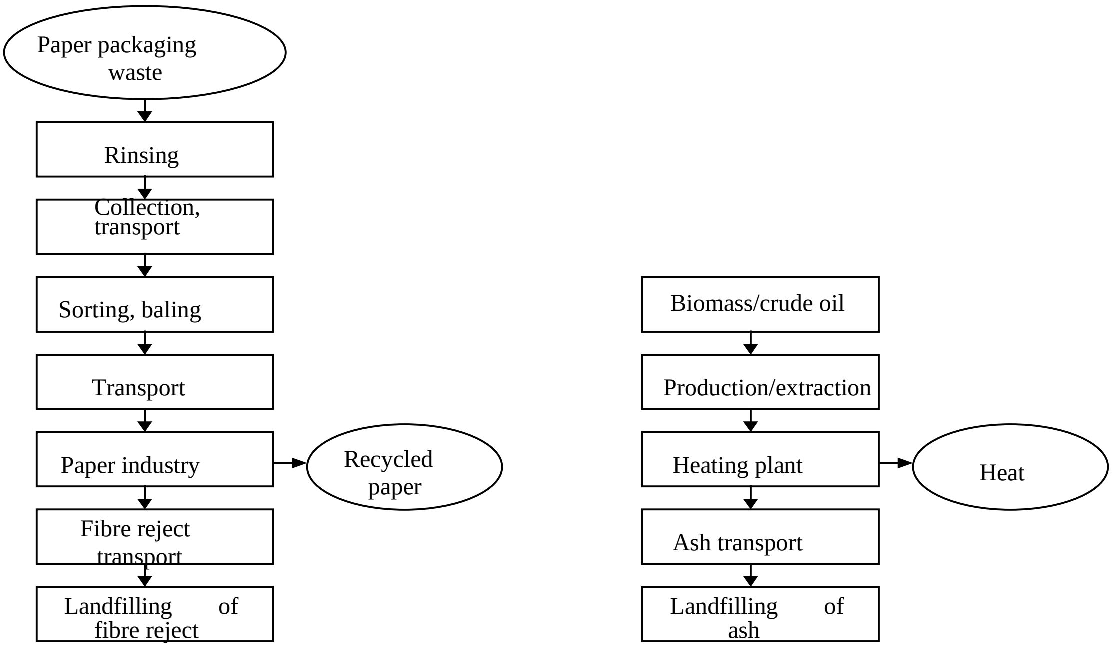
The actual problem is to compare different and not directly comparable systems for treatment of paper packaging waste. With not directly comparable means that the systems do not fulfil the same function. For example, one system produces energy from the paper waste, another produces new paper. This problem can be approached in a number of different ways. The approach in Sundqvist et al is to use system expansion.

In the different scenarios, three functional units are studied. The treatment of 1 kg paper packaging waste, is the main function of the system. The system must also produce new paper, corresponding to 1 kg paper packaging waste and energy (district heat), corresponding to 1 kg paper packaging waste. We neglect, that 1 kg of paper packaging waste may require somewhat different amounts of raw material and produce somewhat different amounts of heat, depending on the composition of the packaging waste.

Sundqvist et al studies treatment of paper packaging waste in five different regions. Here, Skara has been chosen as test case for the multivariate analysis. In addition, a fictitious generic test case was created by averaging the data for incineration and paper production (virgin and recycled) for four regions (Skara, Uppsala, Linköping – Mjölby and Örebro).

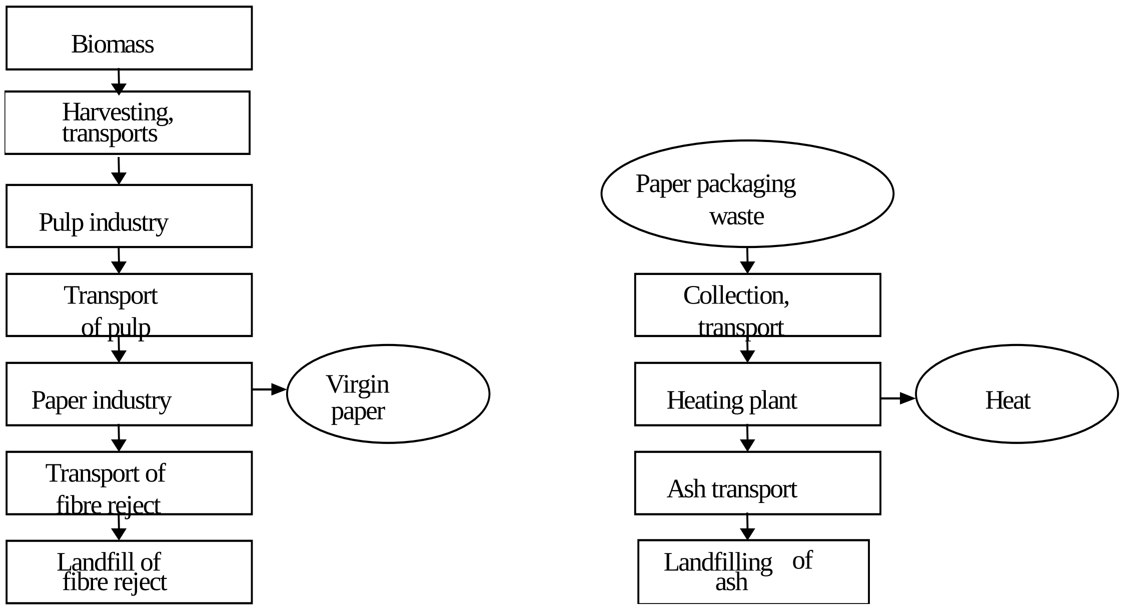
2.1.1 Model structure

The model consists of several modules, for example transportation, collection, paper industry, landfill, heating plant etc. Each module consumes resources (biomass, diesel, oil, uranium etc.), generates emissions to air (SO_2, NO_x, HC etc.) and water (BOD, COD, suspended solids etc.) and different types of wastes (ash, fibre reject etc.). The model flowsheets are presented below.



Material recycling of paper with parallel heat production from another fuel than wastepaper.

Figure 3. System for disposal of paper waste in two different ways, material recycling or energy recovery.



Incineration of (energy recovery from) wastepaper with parallel production of virgin paper from biomass (wood).

Figure 3 (cont.). System for disposal of paper waste in two different ways, material recycling or energy recovery.

2.2 Design of the study

2.2.1 Methodological choices

As stated in section 2.1.1. the goal of the selected LCA case was to compare the environmental consequences of different methods to dispose of paper packaging waste. For the purpose of our study the scope is limited to two disposal methods, material recycling and incineration with thermal energy production. Following the procedure of table 7, the conceptual model of figure 4 may be set up.

The disposal technology, variable X_4 in table 4, is one obvious independent variable, but we may easily identify several others, which will influence the environmental performance of the system and the result of an LCA analysis more or less. For the purpose of this study we select the variables X_1 - X_3 and X_5 in figure 4, in addition to X_4.

Independent variables	**System**	**Outflows**
Type of data, specific/generic X_1 -->	Disposal of paper packaging waste	=> Emissions $e_j \pm \delta e_j$
Heat production, oil/biofuel X_2 -->		
Transport distance, X_3 -->		
Disposal technology, recycling/incineration X_4 -->		⇨ Products $P_k \pm \Delta P_k$ new paper + thermal heat
Paper waste composition, X_5 ->	(Outflows, Consumptions) = $(f_i \pm \delta f_i)$(Inflows, Syst. var.)	
Inflows		=> Wastes $W_l \pm \Delta W_l$
Functional unit, 1 kg of paper packaging waste F_0 =>		
Raw materials, $F_n \pm \Delta F_n$ =>		
Energies, $E_m \pm \Delta E_m$ =>		

Figure 4. Conceptual model of the system ”Disposal of paper packaging waste”.

The main functional unit of the system, F_0, is treatment of 1 kg of paper waste. All calculations in the system are based on this unit.

The different response parameters are divided into parameter categories as shown below. The specific response parameters included in the model are listed in table 9.

2.2.2 Results of the LCI, data quality, completeness check

For the purpose of this study of the interpretation process we will assume that the inventory is complete and adequate. Within the scope of this study it has not been possible to select data quality indicators and calculate uncertainty ranges. Uncertainty ranges for some parameters will be assumed in order to demonstrate uncertainty analysis.

2.2.3 Sensitivity analysis

In this case the issue material recycling versus incineration is obviously only one of several variables. There is at the outset nothing to tell, that this variable is the most important one. It may well be insignificant compared to the influence of other variables. A sensitivity analysis, i. e. a study of the influence of other variables, is clearly necessary. We will use a factorial experimental design to carry out this sensitivity analysis.

The identified independent variables, the X-parameters, are listed and characterised as continuous or discrete in table 8. Each will be varied at two levels with the exception of variable X5, which will be varied at three levels.

Table 8 X-parameters varied in the study.

Variable	Type of variable	Explanation of levels
X1 Type of input data	Two levels, -1 or +1, no pdf*	(-1) specific data (+1)generic data
X2 Heat production from oil or biofuel	Two levels, -1 or +1, no pdf	(-1) biomass (+1) oil
X3 Distance to paper industry	Continuous, may have a pdf, varied at two levels	(-1) 106 km (+1) 300 km
X4 Choice of technique	Two levels, -1 or +1, no pdf	(-1) material recycling (+1) incineration (energy recovery)
X5 Composition of the paper packaging waste	Continuous, may have a pdf, varied at three levels	(-1) 100% cardboard (0) 50% of each (+1) 100% liquid cardboard

*pdf = probability density function.

The X-parameters X2 (resource for heat production) and X3 (distance to paper industry) are dependent on X4 (choice of technique), see description of X-parameters below. X2 and X3 only exist when X4 is at a low level (material recycling). If X4 would be included in the matrix the X-parameters would not be independent of each other as required. This means that X2 and X3 can not be used in the same data matrix as X4. Thus there was a need for two separate data matrices to be able to evaluate all of the X-parameters. Further description of the matrices are shown under headline Data matrices below.

The X-parameters, in the two sets, were varied at two different levels with one centre point in a full factorial design. A full factorial design includes all possible combinations of the varied parameters. Each combination is called an experiment. The tables created by the factorial design were used to make new runs within the LCA-model. It resulted in values of the response parameters for each experiment. The experiments of the factorial designs are shown, together with the calculated values of the response parameters, in the data matrices in the Appendix . One of the main reasons for a factorial design, prior to multivariate analysis, is that it makes it possible to detect interaction effects between the X-parameters.

Input data (X1)

Input data to the LCA can either be average data for the country or data for a specific case. Case specific data are available for paper production from virgin or recycled fibres. X1 = -1 corresponds to case specific data for Skara and X1 = +1 represents generic data.

Heat production (X2)

When paper is recycled another energy source for heat production is needed. Thus this variable only exists when material recycling (X4 = -1) is used. Alternative energy sources can be oil or biofuel. Biofuel as a resource for heat production is represented by X2 = -1.X2 = +1 corresponds to heat production by oil.

Distance to paper industry (X3)

This parameter represents different transport distances to the paper industry when material recycling is used (X4 = -1). X3 = -1 represents the normal case with a distance of 106 km. X3 = +1 corresponds to a distance of 300 km.

Technique (X4)

In the LCA-study there are two available techniques to use when handling paper waste, material recycling and energy recovery. Material recycling is represented by X4 = -1 and energy recovery by X4 = +1.

Composition of paper packaging waste (X5)

The composition of the paper packaging waste may be varied continuously with various amounts of cardboard and liquid cardboard. In the original cases the recycled packaging paper is composed of 100 % cardboard or 100 % liquid cardboard. In this study X5 will be varied at three different levels. X5 can be composed of 100 % of paper (-1), 100 % of liquid cardboard (+1) or 50 % of each (0).

2.2.4 Response parameters

The LCA-model generates 29 different response parameters (Y-parameters). All of them are found in the original study. In table 9 they are divided into parameter categories.

Table 9. Y-parameters generated by the LCA-model divided into parameter categories.

Production parameters (MJ or kg)	**Energy consumption (non renewable resources) (MJ)**	**Energy consumption (renewable resources) (MJ)**	**Resource consumpt. (dm^3 or kg)**	**Emiss. to air (kg)**	**Emiss. to water (kg)**	**Waste (kg)**
Heat energy = E_H_energy	Oil = NRE_Fueloil	Biofuel = RE_Biofuel	Rinsing water = R_Water	SO_2 = G_SO2	BOD = AQ_BOD	Ash = W_Ash
Heat energy from oil = E_H_oil	Coal = NRE_Coal	Bark = RE_Bark	Resource biomass = R_Biom	HCl = G _HCl	COD = AQ _COD	Reject = W_Reject
Heat energy from biofuel = E_H_bio	Diesel = NRE_Diesel	Hydropower = RE_Hydrop		CH_4 = G _CH4	Suspended solids = AQ _TSS	
Produced amount of paper = P_paper	Natural gas = NRE_Natgas			CO = G _CO		
	Peat = NRE_Peat			NO_X = G _NOX		
	Uranium = NRE_Uran			Dust = G _Dust		
				CO_2 = G _CO2		
				N_2O = G _N2O		
				HC = G _HC		

2.2.5 Data matrices

As mentioned above two separate datamatrices were created. They consist of values of the X-parameters (+1,0,-1) from the factorial designs, and response parameters from the performed LCA-calculations. The complete data matrices are found in the Appendix.

Data matrix 1

Data matrix 1 contains four of the five X-parameters. The factorial design resulted in 24 experiments. X4 (choice of technique) is held constant at the low level (-1, material recycling) in this data matrix, as explained earlier. Data matrix 1 evaluates variations in the X-parameters shown below.

X1 = input data
X2 = heat production source
X3 = distance to paper industry
X5 = composition of paper packaging waste

Data matrix 2

This data matrix includes three out of five X-parameters. 12 experiments were created by the factorial design. X2 and X3 are held constant at the low levels (-1, heat production from biofuels, transport distance 106 km) in this matrix, as explained earlier. The following X-parameters are included in data matrix 2:

X1 = input data
X4 = choice of technique
X5 = composition of paper packaging waste

In the matrices the experiments were arranged in rows while the X-parameters and the response parameters (R) were represented by columns (figure 5). The response parameters were gathered according to the parameter categories in table 9.

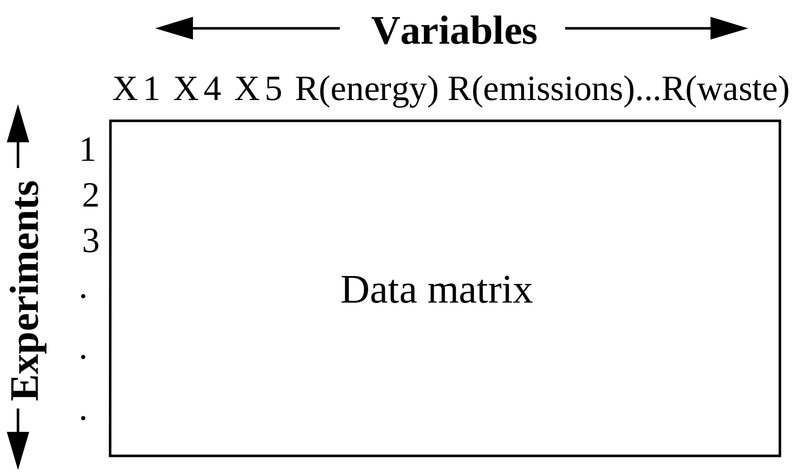


Figure 5. The datamatrix used for multivariate evaluation of X1, X4 and X5.

3. Results and evaluation of the Multivariate Analysis

In the following sections the results of our multivariate analysis will be presented in three parts:

- Principal Component Analysis, PCA.
- Partial Least Square modelling, PLS.
- Uncertainty analysis.

3.1 Principal Component Analysis

3.1.1 Data matrix 1

A Principal Component Analysis (PCA) has been made, composed of three principal components explaining a total of 92,4 % of the variance of the X-parameters. X3 (the distance to paper industry) was only explained to an extent of 0,5 %, which implies that X3 varies independently of the other parameters. Both X2 and X5 were explained by the first two components to an extent of at least 98 %. X2 has its greatest variance explained by component one and X5 by component two. X1 was mainly explained by component three. The total explanation of X1 was 21 %. Due to the poor explainations, X1 and X3 are not discussed in detail in case 1. Although X1 and X3 were not considered to be important in the PCA-model they may be significant for the explanation of the Y-parameters in the PLS, see case 2.

The PCA-model explained all response parameters (except three) by at least 98 %. NRE_diesel, AQ_BOD and G_NO_x were explained to an extent of about 85 %. Two response parameters, E_H_energy and R_Biom, had zero variance due to the fact that X4 is constantly at the low level, material recycling. In the material recycling case, the heat production (E_H_energy) is constant and the biomass consumption for paper production(R-Biom) is zero. Thus, these parameters were excluded from the matrix.

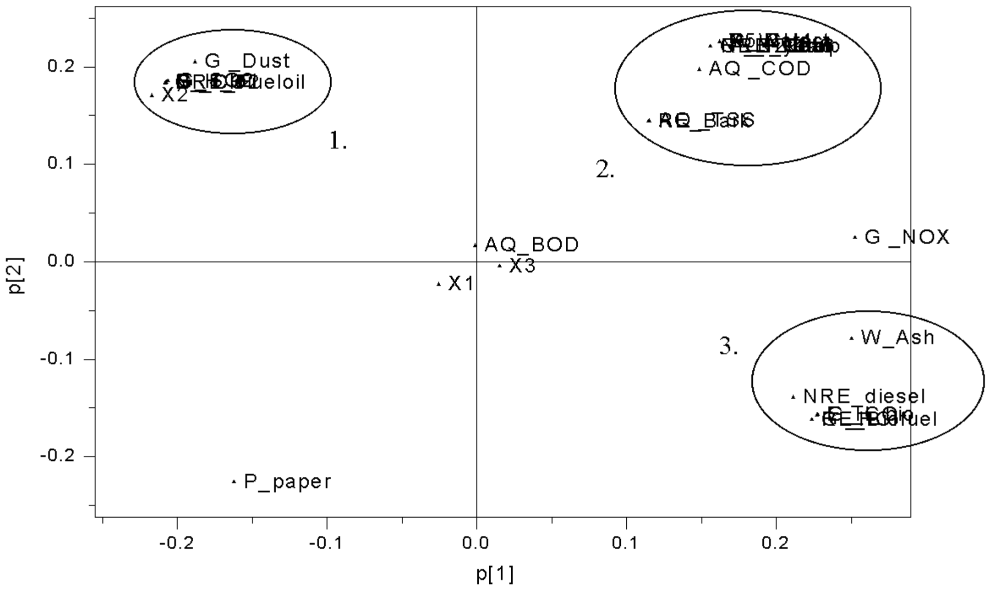


Figure 6. Projection plane showing the first and second principal components of data matrix 1.

In the projection of the first and second principal components there are three distinct clusters, se figure 6. Three clusters are formed. The clusters are circled and numbered from 1-3. P_Paper and G_NO_X are not included in any clusters. The parameters constituting each cluster are shown in table 10 below.

Table 10. Clusters identified in data matrix 1.

Cluster 1	**Cluster 2**	**Cluster 3**
G_Dust	G_N_2O	G_HCl
G_CO_2	G_CH_4	G_CO
G_CH	NRE_Uran	E_H_bio
G_SO_2	NRE_Peat	RE_Biofuel
E_H_oil	NRE_NatG	NRE_diesel
NRE_Fuel oil	NRE_Coal	W_Ash
	RE_Hydro	
	R_Water	
	AQ_TSS	
	AQ_COD	
	W_Reject	

The parameters are not gathered in clusters according to the defined parameter categories. In the projection showing principal component 1 and 2, AQ_BOD together with X1 and X3 are located at origo. This indicates that they have a low variance in these dimensions.

By studying the projections, co-variations between X-parameters and response parameters are discovered. Both cluster 1 and 3 are strongly correlated to X2.

X3 showed not to be of importance to the response parameters. For X3 this means that the increased emissions from the longer transport are relatively small compared to the emissions from the total system and does not have an impact on the overall result.

X2 is positively correlated to the parameters in cluster 1, thus they have high values when oil is used for heat production (X2 = +1). This result is easily explained, all parameters in cluster one is combustion related. The consumption of fuel oil (NRE-Fuel oil) increases naturally when combusting oil, as does the production of heat energy from oil (E_H_oil). The reason for the increase of some combustion related emissions (dust, CO_2, HC and SO_2) when combusting oil is clearly seen when studying the LCA-model. When comparing the modules for production of heat, the combustion related emissions in cluster one all have considerably higher emission factors for heat production through oil combustion than with biofuel combustion.

Cluster 3 co-varies with X2 in an opposite way. X2 at a low level (heat production from biofuel) produce high values on the parameters in cluster 3 and vice versa. This is explained in a similar way. The consumption of biofuel (RE_Biofuel) and the heat production through biofuel (E_H_bio) is of course increasing when X2 has its low level. For the combustion related parameters in cluster three, HCl and CO, again we compare the LCA-modules for heat production from oil and biofuel. It is easily seen why HCl and CO is located in cluster three; the emission factors used for these emissions are far higher for biofuel combustion than for combustion of oil. The W_Ash is explained in the same way. Combustion of biofuel produce much more ash than combustion of oil, therefore the co-variation. The last parameter in cluster three, NRE_diesel, the different diesel consumptions in the heat production systems in the LCA-model must be considered. When studying the system, it shows that a dominating part of the diesel consumption is from the biofuel precombustion. This is the main reason why the response parameter NRE_diesel is located in cluster 3 and co-varies with X2. The precombustion in the oil chain is included in another fossil fuel parameter, NRE_Fuel oil.

Cluster 2 and P_paper are both strongly correlated with X5. The parameters in cluster 2 receives high values when X5 is at a high level (liquid cardboard). The correlation between X5 and P_paper is not important, only natural, when handling 1 kg paper packaging waste, more recycled paper is produced than when handling 1 kg liquid cardboard waste However, to explain the response parameters in cluster two, and their dependence on X5, another study of the LCA-model is needed. The correlation with W_Reject is easily seen. If liquid cardboard is recycled (X5 has a high value), more reject from the recycling industry is produced (W_Reject has a high value). The energy carriers uranium, peat, natural gas, coal and hydropower (NRE_Uran, NRE_Peat, NRE_NatG, NRE_Coal and RE_Hydro) are all sources for electricity production. Thus, when X5 is at the high level (liquid cardboard waste), the electricity consumption of the system is high. This fact depends on that the liquid cardboard waste (milk beverage etc) is rinsed before it is collected for recycling. The rinsing, consequently, explains why the water consumption co-varies with X5. For COD and TSS, which co-vary with X5, the explanation were found in the rinsing and the recycling process. Handling of liquid cardboard results in higher COD and TSS emissions. Surprisingly, BOD varies differently, and independent of X5. This depends on that BOD is more affected of the data source, X1 (generic or specific data).

The air emissions in cluster 2, N_2O and CH_4, can also be explained by studying the LCA-model. The N_2O emissions origin mainly from the electricity production. Since the electricity production co-vary with X5, N_2O also co-vary with X5. CH_4 emissions origin mainly from the landfill of reject from the recycling industry. The reject amount co-vary with X5 as discussed above and thus, so does CH_4.

Principal component 1 always explains the largest variance in data and the second component explains the second largest variance. The result from the PCA-model implies that X2 (resource for heat production) is the most important parameter (of the four parameters studied in this case) in explaining the variation in data. The second most important X-parameter is X5. How the X-parameters effect the response parameters will be further studied in the PLS-model for data matrix 1.

Data matrix 2

Data matrix 2 was used to perform a PCA-model. The PCA-model consisted of three principal components explaining a total of 89,9 % of the variance in data. Most of the response parameters were explained to an extent of 90 %. A few exceptions were NRE_diesel, RE_bark, G_CO_2 and AQ_TSS.

X1 had a total explanation of 69 %, mostly explained by component two. X4 had its greatest variance in component one (96 %) and a total explanation of 98,5 %. X1 and X4 have their greatest variations in projection of the first and second principal components. The remaining X-parameter, X5, was explained to an extent of 84,9 % of which 82,2 % was explained by component three. In a data set with these X-parameters the response parameter E_H_oil had no variance (X2 is held constant at a low level – biofuel consumption) and was excluded from the data set.

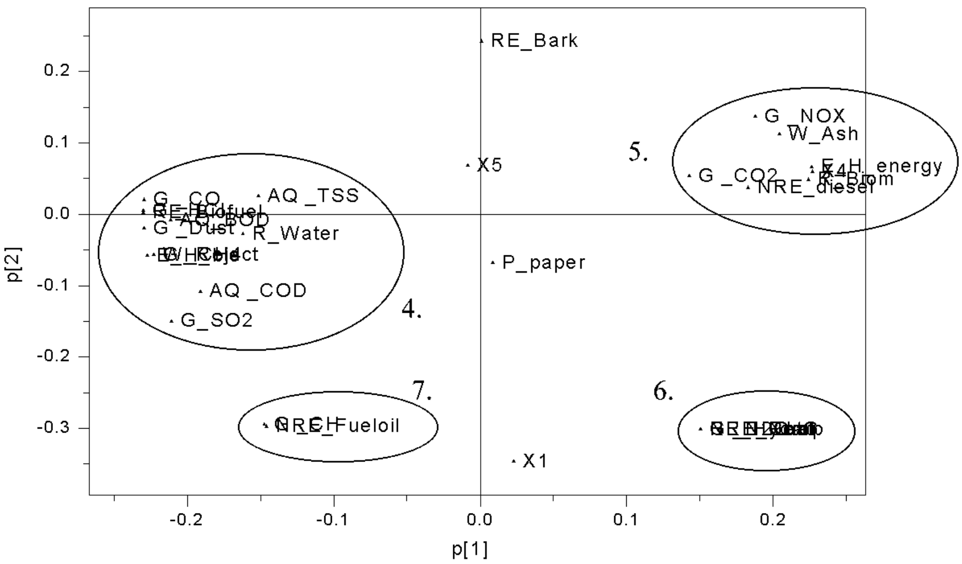


Figure 7. Projection plane showing the first and second principal components of data matrix 2.

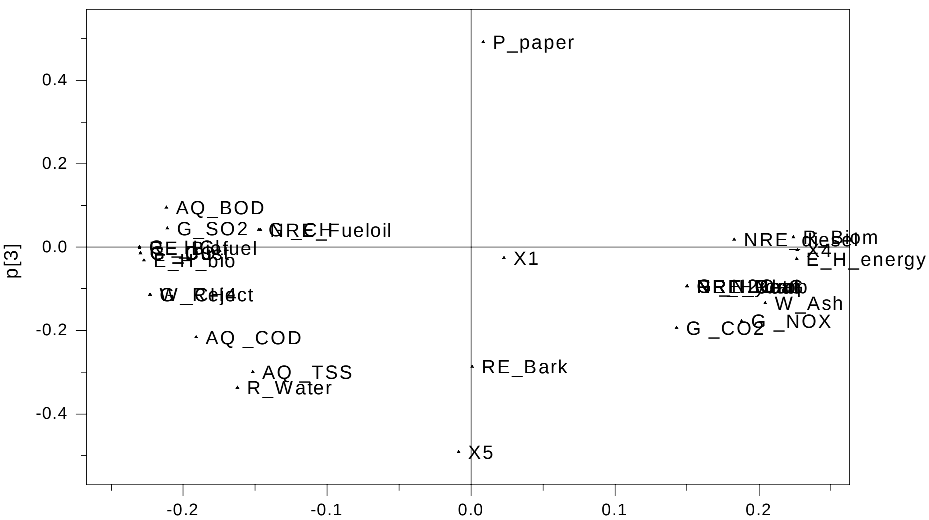


Figure 8. Projection plane showing the first and third principal components of data matrix 2.

Four clusters were identified in the first PCA-projections. They are numbered and circled in the figures. The clusters are numbered 4, 5, 6 and 7 to make it easier to keep apart the results from the two data matrices.

The clusters are not found in the projection of first and third principal component. In this projection X5 is preferably studied. Since X5 does not co-vary with any clusters it is not interesting to try to identify any new clusters in this projection.

Table 11. Clusters identified in data matrix 2.

Cluster 4	**Cluster 5**	**Cluster 6**	**Cluster 7**
E_H_Bio	E_H_energy	NRE_Uran	NRE_Fuel oil
RE_Biofuel	NRE_diesel	NRE_Coal	G_CH
R_Water	R_Biom	NRE_Peat	
G_CO	G_CO_2	NRE_NatG	
G_CH_4	G_NO_x	RE_Hydro	
G_HCl	W_Ash	G_N_2O	
G_Dust			
G_SO_2			
AQ_TSS			
AQ_BOD			
AQ_COD			
W_Reject			

The response parameters are to some extent gathered according to the parameter categories in table 11. Cluster 4 contains five gas emissions and all emissions to water. Cluster 6 is mainly constituted of non renewable resource parameters. Parameters not included in any cluster are P paper and RE_Bark.

Co-variations between X-parameters and clusters are shown in the PCA-projection, figure 7. X4 co-varies with cluster 5 and 4. X4 at a high level (energy recovery) is positively correlated to the parameters in cluster 5. E_H_energy and R_Biom are easy to explain. E_H_energy is heat energy produced by combusting paper packaging waste, so consequently, X4 at a high level means a high value on E_H_energy. R_Biom is biomass used for producing virgin cardboard, which of course has a high value when X4 is at a high level. When combusting paper packaging waste, we have to use biomass to produce new paper packagings. For the other response parameters in cluster 5, a study of the LCA modules must be done.

The reason why W_Ash co-varies with X4 is that the majority of the ash produced in the system originates from the combustion of the packaging waste, thus, a high value on X4 gives a high value on W_Ash. The diesel consumption also co-vary with X4. Although there are many diesel consumption sources, the main reason for the correlation between the consumption and X4 is the high diesel consumption in the harvesting and transport of biomass to the virgin paper production. For the gaseous emissions, CO_2 and NO_x, which obtain high values when X4 is at its high value (energy recovery), the interpretation is more complex. The reason for CO_2 co-varying with X4 is that when liquid cardboard waste (X5 high) is treated, there will be a correlation between X4 and G_CO_2. This influence is strong enough to eliminate the fact that when cardboard waste is combusted (X5 low), the result is completely the opposite, i.e. CO_2 decreases with increasing X4, if X5 is *low*. For NO_x, the main reason for the co-variation with X4 is the relatively high NO_x emission factor used for packaging waste combustion.

The response parameters in cluster 4 has a negative correlation with X4, thus the response parameters receive low values when X4 is at high level. The opposite is valid for X4 at a low level (material recycling). Some of the parameters in cluster 4 are easily explained. RE_Biofuel is simply the consumption of biofuel for heat production, which of course is lower when X4 is high. When combusting packaging waste, less combustion of biofuel is needed. E_H_Bio is the heat production coming from biofuel, which is explained in the same way. The water consumption is lower when paper packagings are energy recovered, in that case no rinsing is needed. The gaseous emissions, G_CO, G_CH_4, G_HCl, G_Dust and G_SO_2 are all decreasing when combusting paper packaging (X4 high). For CH_4 the main reason is the relatively high emissions from landfilling of recycling industry reject. For the other gaseous emissions, the main reason is that the combustion of biofuel has higher emission factors for these emissions than combustion of the packaging waste. The aqueous emissions TSS, BOD and COD also decrease when combusting the packaging waste (X4 high). This fact depends mainly in the higher emissions from the recycling industry emission factors used compared to the virgin production emission factors.

X4 is negatively correlated to cluster 7. NRE_Fuel oil in cluster 7 is thus increasing when X4 is at low level. The reason is mainly the oil combustion in the recycling industry. In the virgin production modules, no oil is combusted. The HC emission which increase when X4 decrease has the same origin, the oil combustion in the recycling industry.

X1 co-varies with the parameter RE_Bark and is has a positive correlation with Cluster 7. The co-variation with RE_Bark is easily explained when comparing the data sets used in the model. The specific data (X1 low) used has a higher bark consumption than the generic data (X1 high). The correlation with cluster 7 is mainly explained by the differences in oil combustion in the generic and specific data modules. The oil combustion is higher in the generic data set, which also affects the HC emissions in cluster 7 as discussed above.

X5, which is mainly explained by component three, is viewed in the second projection. The projection shows a clear co-variation between X5 and P_Paper. X5 at a high level (liquid paper) results in a lesser amount of produced paper compared to X5 at a low level, which is obvious. More recycled cardboard can be produced when recycling paper cardboard waste than liquid cardboard waste

As mentioned before the greatest variance in data is explained by the first principal component. The X-parameter mainly explained by component one is X4 (choice of technique). X4 is considered to be the most important X-parameter in this data matrix. The second most important is X5 (composition of paper packaging waste).

3.2 Partial Least Square Model

The same data material was used in these evaluations except that only one response parameter from each cluster was used. They were considered to represent the parameters in the defined clusters.

Data matrix 1

According to the cluster formation in the PCA-model four response parameters were chosen to be included in the model. The parameters were G_SO_2 (Cluster 1), NRE_NatG (Cluster 2), G_CO (Cluster 3).

The parameter X3 was removed from the data matrix since its VIP-values (showing the importance and the degree of explanation) were low for all Y-parameters. A parameter with low VIP-values are not significant to the variance in Y and are preferably removed. The resulting PLS-model consisted of two components explaining 97,3 % of the variance in the Y-parameters. The prediction ability was 94,2 % showing the model quality to be good.

The remaining variance in data, not explained by the model, is called the residual variance. It is represented by residuals. When validating a PLS model the residuals have to be studied to detect if the model may be improved further or not. The residuals were plotted in a normal probability plot and observed values from the experiments were plotted against predicted values calculated by the model. For the selected Y-variables both types of plots gave satisfactory results. This implies that the quality of the model was good and did not need any improvement..

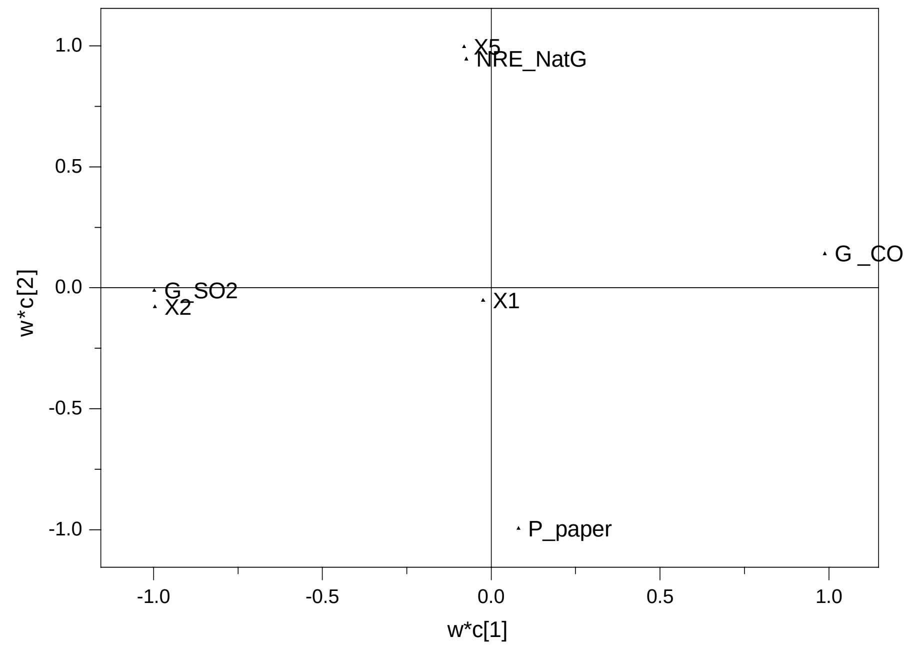


Figure 9. Scatterplot of the 1st and 2nd principal components showing the results of the PLS.

The scatterplot in figure 9 gives an idea about how the X-parameters affect the Y-parameters. G_SO_2 (Cluster 1) and G_CO both co-varies with X2 but in opposite ways. X5 is positively correlated with NRE_Natgas and negatively correlated with P_Paper, which was indicated in the PCA. To be able to more precisely define the impact of the X-parameters their coefficients were plotted for each Y-parameter used in the PLS model (figure 10).

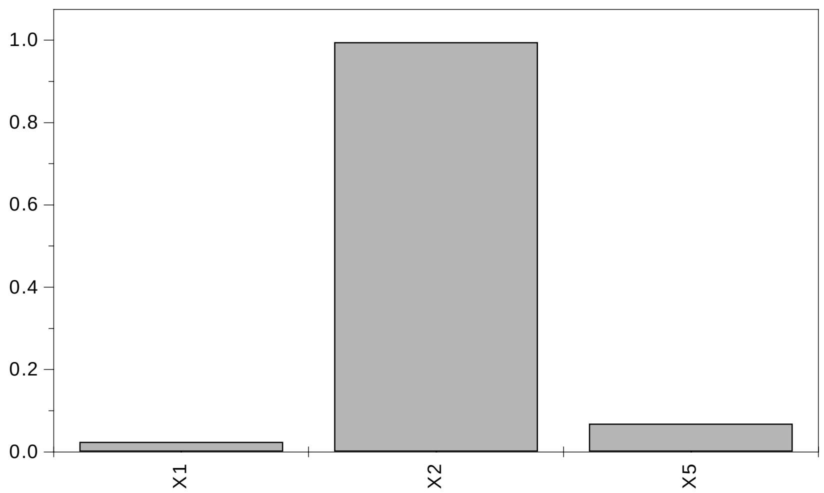


Figure 10. The coefficients of the X-variables for G_SO_2 (cluster 1).

Cluster 1, represented by G_SO_2, is mainly affected by X2. The contribution from the other X-parameters are negligible. X2 is positively correlated with the parameters in cluster 1. X2 at a high level (resource oil) results in a high emission of SO_2 and thus high values on the parameters in Cluster 1.

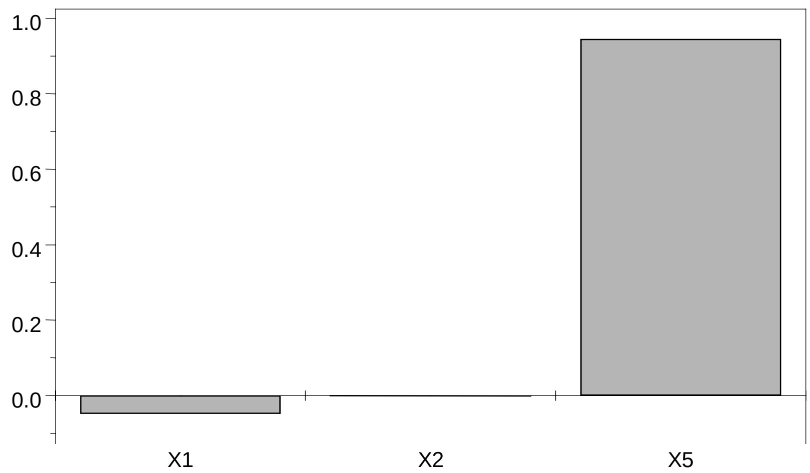


Figure 11. The coefficients of the X-variables for NRE_NatG (cluster 2).

The parameters in Cluster 2 represented by NRE_NatG mainly depends on X5 (figure 11). X2 does not affect NRE_NatG at all while X1 shows some effect but it is negligible. X5 is positive correlated with Cluster 2 thus X5 at a high level results in a large consumption of natural gas and high values on the other parameters in Cluster 2. A high value on X5 means that liquid cardboard waste is treated. As stated in the PCA, the liquid cardboard treatment requires more electricity which for example results in an higher consumption of natural gas.

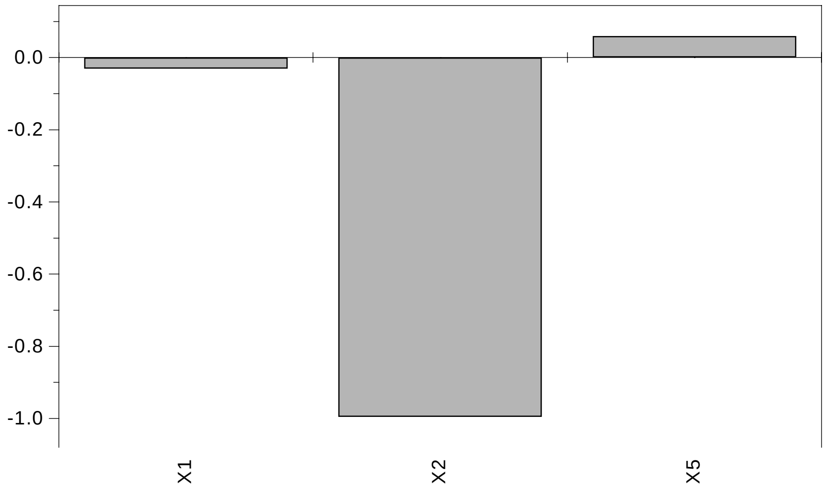


Figure 12. The coefficients of the X-variables for G_CO (cluster 3).

The most significant X-variable to the parameters in cluster 3, represented by G_CO, is X2 (figure 12). The effects from X1 and X5 are negligible. The Y-parameters in this cluster receives high values when X2 is at low level (resource biofuel). This is explained when studying the LCA-modules as done in the PCA. Biofuel combustion leads to higher CO emissions.

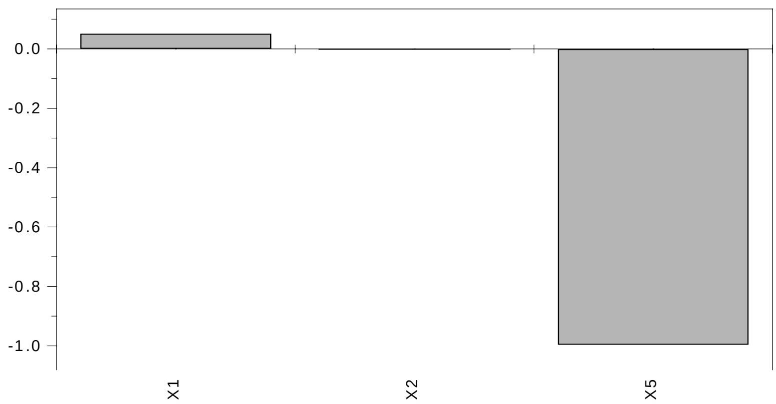


Figure 13. The coefficients of the X-variables for P_paper.

The variation of the parameter P_paper is of course solely explained by X5 (figure 13).

Data matrix 2

Based on the result of the cluster formation in the PCA-model one parameter was chosen from each cluster (except for cluster four from which two parameters were chosen). The chosen parameters were: NRE_NatG (Cluster 6), G_SO_2, G_CO (Cluster 4), and W_Ash (Cluster 5). No parameter from cluster 7 was evaluated, since it contains only two parameters and is close to cluster 4.

During the PLS-modelling it was found that interaction effects between the X-parameters played an important role to the variance in Y. When two X-parameters are at high level at the same time and they together result in an increase of the studied parameter the X-parameters are said to have an interaction effect. These interaction effects were added to the data set as additional X-parameters. All possible interaction combinations were evaluated. Two out of three interaction effects showed to be of importance, X1*X4 (C1*2) and X4*X5 (C2*3). The addition of interaction effects resulted in an improved model.

The resulting PLS model consisted of two PLS components explaining 96,3 % of the variance in Y. The model had a prediction ability of 86,8 %. To further validate the quality the residuals were plotted in a normal probability plot and observed values from the experiments were plotted against predicted values calculated by the model. Both types of plots gave satisfactory results.

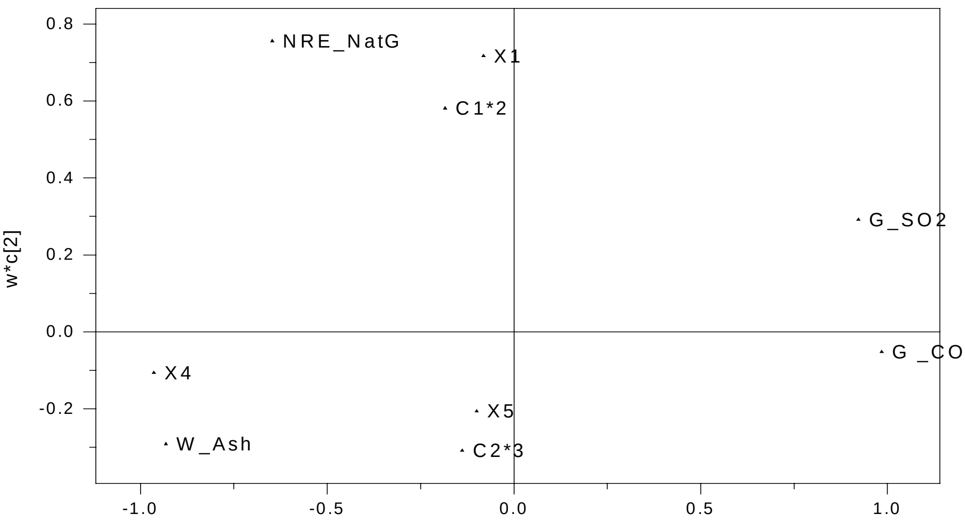


Figure 14. Scatterplot of the 1st and 2nd principal components showing the results of the PLS.

The scatterplot (figure 14) gives an idea of how the X-parameters affect the Y-parameters. X4 co-varies with W_Ash, G_SO_2 and G_CO. The interaction effects between X4 and X5 co-varies with X5. X1 is positively correlated to the interaction effects of X1 and X4 and NRE_NatG. The relationships can be further studied in coefficient plots.

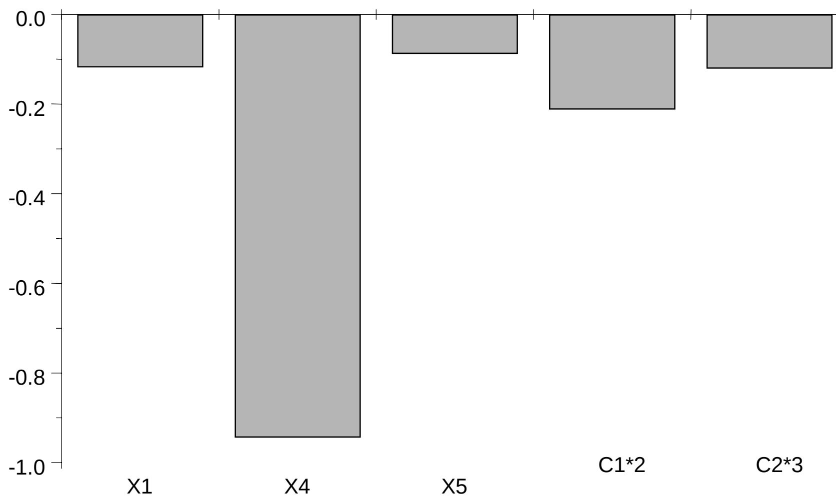


Figure 15. The coefficients of the X-variables for G_CO (cluster 4).

G_CO representing cluster 4 is mainly affected by X4 (choice of technique) (figure 15). As stated in the PCA-evaluation, X4 at a low level, material recycling, results in higher combustion related emissions than X4 at a high level. It is also affected by the other X-parameters and the interaction effects. X4 at a high level (energy recovery) results in low values of the parameters in cluster 4.

X1 at a low level, specific data, increases the CO emissions. The reason for this is that the specific data modules (both for virgin cardboard production and recycling) consume more biofuel and less oil than the generic data modules, which result in a higher CO-emission.

X5 at a low level, cardboard waste, also increases the CO emissions. This depends partly on he fact that when recycling cardboard waste, different fuels are used in the specific and in the generic data module, which result in different CO emissions.

This fact is even stronger when X4 has its high value, material recycling, which might be the reason for the influence of the interaction effect between X4 and X5.

The interaction effect between X1 and X4 might be due to the fact that the choice between specific and generic data is important in the incineration alternative (X4 = 1). Specific data for CO emissions are higher than the generic data for these emissions. In the material recycling case (X4 = -1), on the other hand, the choice between generic and specific data has no influence at all.

The interaction effect between X4 (technique) and X5 (composition of paper packaging waste) may be explained in a similar way. In the material recycling case (X4 = -1) the SO_2 emissions are hardly affected by the composition, whereas in the incineration case (X4 = 1) the SO_2 emissions are increased, when pure cardboard is incinerated.

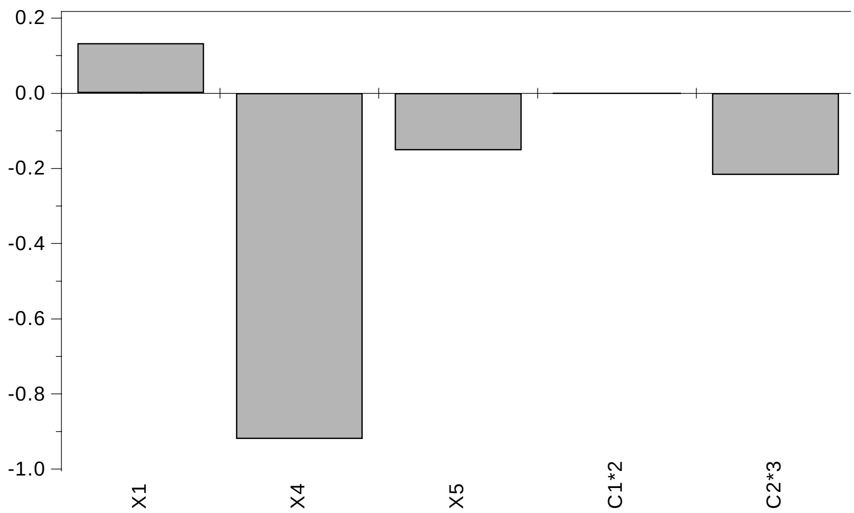


Figure 16. The coefficients of the X-variables for G_SO_2 (cluster 4).

Cluster 4 is represented by both G_SO_2 and G_CO. The coefficient for G_SO_2 are similar to G_CO. Detected differences between the coefficients are concerning X1*X4 and X1, see figures 15 and 16. As mentioned for G_CO the parameters in cluster 4 receive low values when energy recovery is used. The reason why the X1 influence on SO_2 differs from its influence on CO is due to the fact that the data choice affects the emission factors for SO_2 and CO differently. In this case, the specific data includes more biofuel combustion which give rise to higher CO-emissions. The generic data modules include more oil combustion and thus, higher SO_2 emissions.

The absence of influence of the interaction between X1 and X4 on SO_2 is explained through the emission factors. After the choice of material recycling or energy recovery, the choice between specific or generic data does not have a large influence, thus no interaction effects can be found.

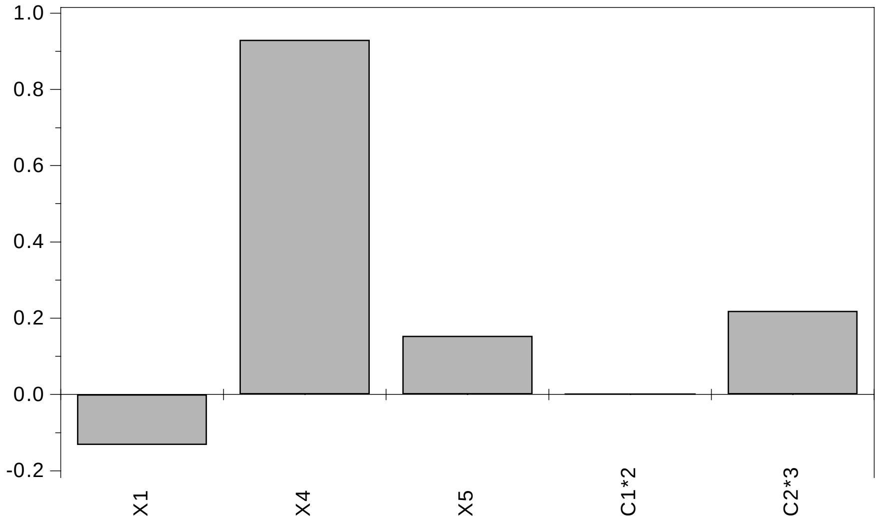


Figure 17. The coefficients of the X-variables for W_Ash (cluster 5).

Also cluster 5 is affected by X4 but in an opposite way compared to cluster 4.

X4 has a positive correlation with the parameters in Cluster 5 and vice versa (figure 17). The reason for the correlation is discussed in the PCA analysis, the reason is mainly due to the incineration process when combusting the paper packaging waste. The type of input data, has a small influence on W_Ash and the parameters in cluster 5. Again, the answer is to be found in the fuel composition. In the specific data modules, more biofuel is combusted and thus more ash is produced. Most of the other parameters in cluster 5 can be explained in a similar way, for example R_Biom. The other parameters do not follow the same patterns as W_Ash.

A high value on X5, liquid cardboard waste, increases the ash production, both when material recycling or incinerating the cardboard. The largest amount ash is produced when the liquid cardboard is energy recovered, which explains the influence of the interaction effect between X4 and X5.

Caution must be exercised when extrapolating the PLS model for W_ash to the other parameters of cluster 5. E.g. if we extrapolate the model to R_biom we would predict a small increase of the consumption of biomass for the production of virgin paper when X5 is increased, i.e. when we increase the share of liquid cardboard in the packaging waste. In reality the opposite is true, since liquid cardboard is a composite material of cardboard and plastics. The mathematical reason for this discrepancy between model and reality may be, that R_biom is zero irrespective of the value of X5, when X4 = -1, i.e. when material recycling is used as a disposal method. The behaviour of R_biom is not well represented by the behaviour of the variable W_ash, although the PCA analysis assigns them to the same cluster.

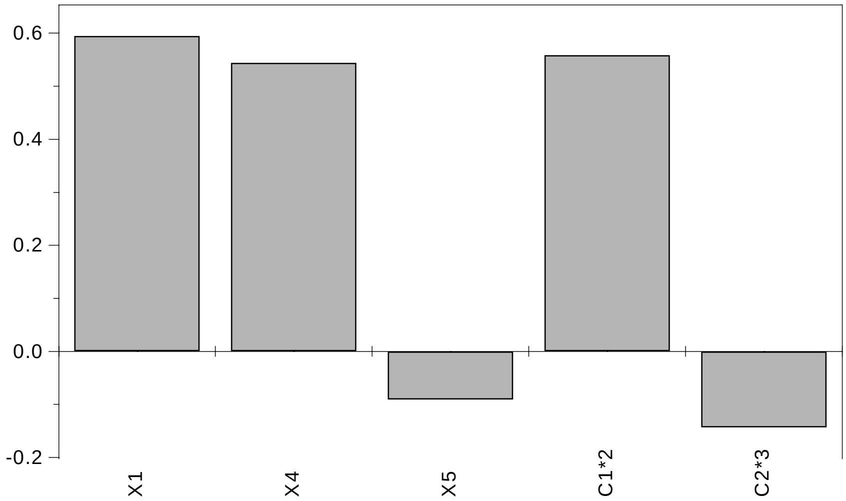


Figure 18. The coefficients of the X-variables for NRE_NatG (cluster 6).

NRE_NatG is affected by X1 and X4 to about the same degree. The interaction effect between X1 and X4 is also important to be able to describe the variance of the variables in cluster 6 (figure 18). X1 and X4 at high levels are positive correlated with cluster 6.

All parameters in cluster 6 are directly electricity related, except N_2O. When studying the modules, the only module that produce N_2O is the electricity production, which in this case makes N_2O to directly electricity related too. X1 at a high level, generic data increases the electricity consumption, and thus increase the parameters in cluster 6. A high value on X4, incineration, increase the electricity consumption, and also the parameters in cluster 6. The large influence from the interaction effect between X1 and X4 is due to the large increase in electricity consumption when both X1 and X4 have their high values.

3.3 Uncertainty Analysis

When evaluating results it is always important to know the uncertainty in data to know how valid the drawn conclusions are. In this case a number of different ways to determine the uncertainty is studied. To evaluate the uncertainty replicates of two chosen experiments were made. Replicates were made on experiment 1 and 3. These experiments were chosen since they are identical apart from choice of technique (X4). X4 was considered to be the most important parameter to study in this matter.

Three Y-parameters were chosen for the evaluation, G_SO_2, G_NO_X and AQ_COD. All other Y-parameters except these three were deleted from the matrix. The emission factors in the LCA-model, concerning the mentioned substances, were given realistic uncertainty intervals. For SO_2, the emission factors was given the uncertainty ± 20% and indirectly related emission factors ±10%. For NO_x, the uncertainty was ± 30% and ± 10%, respectively. For COD, the uncertainty was ± 50% and ± 10%, respectively . By using the LCA-software KCL-ECO a minimum and a maximum value on each emission were calculated with Monte Carlo simulations, with regard to a 95 % confidence interval. The calculated values were added to the data matrix. The replicates were named high respectively low.

3.3.1 Multivariate evaluation

The spread of the Y-parameters were studied in a PCY. A PCY is similar to a PCA. In a PCY the variance of the Y-parameters are studied instead of the variance of all parameters, like in a PCA.

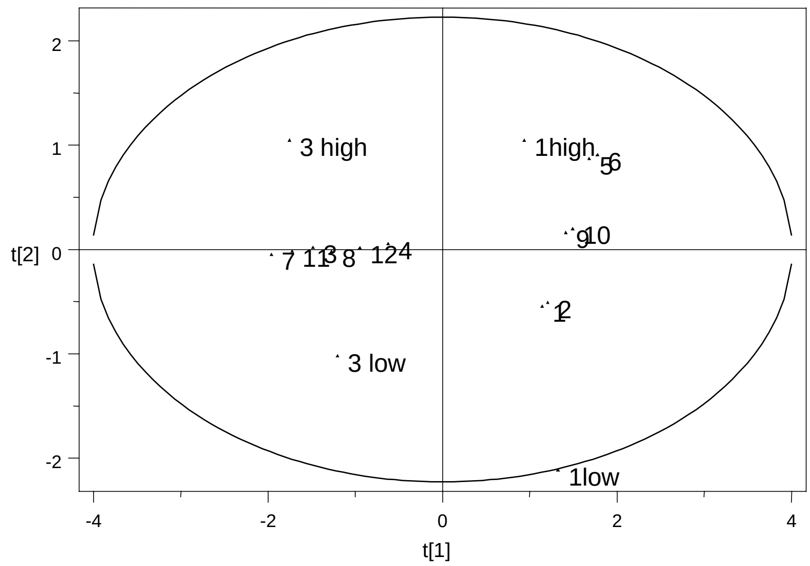


Figure 19. A PCY plot showing the spread of the observations according to the variance of the Y-parameters.

The results (figure 19) show that the spread of the replicates is larger than the spread between the remaining observations. The desired result is that the replicates should vary less than separate experiments. This result implies that uncertainty in data is too large to be able to draw conclusions to which cluster a certain parameter belongs. More accurate statistical methods to describe the uncertainty in data are studied below.

3.3.2 Evaluation of variance

The variance of the replicates for each Y-parameters were calculated. The calculated values were compared with variance values of the original experiments. Data of the variance are shown in table 12.

Table 12. Results from variance calculations.

Variance	**G_NO_x**	**G_SO_2**	**AQ_COD**
Experiments	$1.63*10^{-7}$	$1.22*10^{-7}$	$5.78*10^{-5}$
Replicates exp. 1	$7.66*10^{-7}$	$3.38*10^{-8}$	$4.82*10^{-6}$
Replicates exp. 3	$3.78*10^{-7}$	$7.8*10^{-9}$	$1.23*10^{-6}$

For the data quality to be adequate for a multivariate evaluation the variance within the replicates should be less than the variance within the original experiments. This is valid for G_SO_2 and AQ_COD. For G_NO_X the variance within the replicates is slightly larger than the variance within the original experiments. Although the variance within the replicates is less than within the experiments it is questionable if the assumed data quality is satisfactory.

3.3.3 t-test

A t-test was performed to determine if the difference in Y-data, for material recycling (A) and energy recovery (B), is a coincidence due to poor data quality. The result is given in probability values telling the probability of a coincidence. The principles for t-tests are found in Box et. al (1978).

Results were produced for each of the three evaluated emissions, NO_x, SO_2 and COD. The probability that the difference between case A and B occurs by a coincidence are shown in table 13.

Table 13. Probability values.

Emission	**G_NO_x**	**G_SO_2**	**AQ_COD**
Probability	198/1000	5/1000	2/1000

For our purposes, the data quality seems to be less good for NO_x compared to the others. The results from SO_2 and COD may be considered as sufficient to determine that the difference between A and B is not a coincidence. For these parameters it is a 5 and 2 % probability resp., that the emission data resulting from case A and B differ by a coincidence.

4. Interpretation

4.1 Data quality check

Only a limited sensitivity analysis with assumed uncertainties could be performed within the frame of the project. Nevertheless the analysis suggests, that rather stringent data quality criteria are required, if statistically significant conclusions are to be drawn. The sensitivity analysis was performed on a very clear-cut comparison between two scenarios, where only one variable (the most interesting parameter) was changed, and where the data quality was average to good (section 3). Even so, for one out of three studied standard emission parameters no statistically significant dependence on the studied variable could be verified. The observed increase of NO_x emissions when changing from material recycling to incineration could be a coincidence, despite the fact that the uncertainty of the NO_x-emission data are only 20 % or less, and despite the fact that all other changes than the change of technology from material recycling to incineration have been ruled out. (The ratio $(NO_x)_{incineration}/(NO_x)_{material\ recycling}$ is about 1.5 in our scenarios. For the other two emission parameters studied in the sensitivity analysis, emissions of SO_2 and COD, the ratios incineration/material recycling are 0.7 and 0.8 respectively. For these two parameters the probability that the difference is a pure coincidence is very low however).

4.2 Drawing conclusions

Keeping in mind, that it is actually doubtful whether or not the observed differences between the calculated scenarios are coincidental or actually caused by the changes of the independent variables, and that a sensitivity analysis for each studied parameter would be warranted, we will try the procedure of drawing conclusions. Since the primary question is the difference between material recycling and incineration of paper packaging waste, the first step is to list those response parameters which are solely or at least mainly dependent on the independent variable X4 “choice of technology”. These response parameters have been identified by the PCA and PLS analyses and are grouped in the clusters 4 and 5. Table 14 gives the list.

In addition to the parameters of table 14, the parameters of cluster 6 and 7 also depend on the variable X4 (see table 11 and figure 18). Cluster 6 describes electricity use, and the PLS analysis shows, that a change from material recycling to incineration increases the consumption of electricity (figure 18). However, at least in our case study, the amount of electricity also depends on the choice of data, generic or specific.

Cluster 7 is mainly fuel oil consumption. No PLS analysis has been performed for cluster 7, but the PCA shows, that the cluster is negatively correlated to X4, i.e. material recycling increases the consumption of fuel oil (see page 27). The oil consumption is, however, also dependent on the choice of data, like the electricity consumption.

It follows from the description of the multivariate results (see page 15) that the interpretation must be made with caution. The fact that the PCA analysis assigns a group of parameters to the same cluster does not necessarily mean that the parameters will behave in the same way in all respects. Since we have in this study restricted the PLS analysis to one parameter per cluster, we will only consider the dominating effects found in the analysis.

Table 14. Response parameters dependent on the choice of technology.

Response parameter	**Effect of changing the technology from material recycling to incineration, i.e. of changing X4 from –1 to +1** **- = decrease, + = increase,** **0 = no effect.**	**Response parameter also dependent on**
E_H_Bio	- if X2 = -1 0 if X2 = +1	X2, heat prod.
RE_Biofuel	- if X2 = -1 0 if X2 = +1	X2, heat prod.
R_Water	0 if X5 = -1 - if X5 = 0 or +1	X5, waste comp.
G_CO	- if X2 = -1 + if X2 = +1	X2, heat prod.
G_CH_4	-	
G_HCl	- if X2 = -1 + if X2 = +1	X2, heat prod.
G_Dust	- if X2 = ±1	X2, heat prod.
G_SO_2	- if X2 = ±1	X2, heat prod.
AQ_TSS	0 if X5 = -1 - if X5 = 0 or +1	X5, waste comp.
AQ_BOD	-	
AQ_COD	- for all X5	X5, waste composition
W_Reject	- for all X5	X5, waste composition
E_H_energy	+	
NRE_diesel	+ if X2 = ±1	X2, heat prod.
R_Biom	+	
G_CO_2	- if X5 = -1, X2 = -1 + if X5 = 0 or +1, X2 = -1 - if X2 = +1 for all X5	X2, heat prod, X5, waste comp
G_NO_x	+	
W_Ash	+ if X2 = ±1	X2, heat prod.

It is obvious from table 14 that the first conclusion must be, that there is no such thing as an unambiguous environmental effect of a change of the waste packaging disposal technology. There are only five environmental parameters, G_CH_4, AQ_BOD, E_H_energy, R_Biom and G_NO_x, which are unequivocally dependent on the variable X4 "choice of technology". The effect on those other parameters, which are also functions of X4, is also dependent on the variables X2 "heat production" and/or X5 "composition of the packaging waste". In table 14 we have indicated how the variables X2 and X5 influence the effect of the variable X4 on the parameters. The influence of X2 has been deduced by comparing results from material recycling with replacement heat production from oil (matrix 1) and results from incineration (matrix 2).

Table 14 could for instance be used to conclude, that a change of disposal technology for waste packagings from incineration to material recycling **may** decrease emissions of carbon dioxide, **provided that a biofuel is used to produce replacement heat, and provided that the packaging material contains at least 50 % liquid cardboard**. If one of these two conditions is not fulfilled, the exchange of incineration for material recycling will increase the emissions of carbon dioxide.

The second conclusion to be drawn from table 14 is that aggregation of emissions to impact categories may obscure the influence of a system change. Taking as an example the impact category global warming potential, GWP, there are two greenhouse gases in table 14, which are affected by the variable X4, namely methane and carbon dioxide. We have already concluded, that conditions may be found, under which introduction of material recycling will reduce the emission of carbon dioxide. The emission of methane will unambiguously increase, if incineration is replaced by material recycling. A practitioner using the impact category GWP may well report no significant change as a consequence of a change of disposal technology. In reality the character of the emissions of greenhouse gases may have been fundamentally changed.

The same reasoning may be applied to acidifying emissions. Table 14 contains to acidifying gases, sulphur dioxide and nitrogen oxides, which belong to different clusters.

4.3 Comparison with an alternative interpretation procedure

The case selected for this study, namely material recycling versus incineration of paper packaging waste, has been the object of several investigations, with apparently conflicting and confusing results. The problem how to draw conclusions from these results has been addressed before. We can thus compare the result of our interpretation with an earlier attempt.

Finnveden and Ekvall (Finnveden and Ekvall 1997) used a simpler approach than multivariate analysis. They compiled results from12 studies of recycling versus incineration. By comparing the assumptions of the different studies they identified three variables besides the technology choice recycling/incineration, namely alternative heat production (our X2), electricity production technology (not studied by us), and heat source at the paper mills. They also identified 12 parameters (resource consumptions and emissions), which had been used by most of the investigators.

The method of analysis was a simple, qualitative approach. A table with the 12 selected response parameters was constructed. For each parameter and each case study the authors made an entry, indicating in which case, recycling or incineration, the impact was lower. Data uncertainty and the magnitude of the difference between recycling and incineration were disregarded. An excerpt of the table is shown below as table 15.

Table 15. Inventory results from the Skara case of Finnveden et al (1994) as interpreted by Finnveden and Ekvall (Finnveden and Ekvall 1997). (Our case X1 = -1 “specific data” corresponds to Skara Biofuel/Oil, Swed. av., Unspec.). An “R” indicates lower impacts from recycling, an “I” lower impacts from incineration, “0” no difference. A gap denotes that the parameter was not included.

Case	Skara			
Alternative energy source	Biofuel	Biofuel	Biofuel	Oil
Electricity production	Swed. av.	Coal	Swed. av.	Swed. av.
Heat source at mills	Unspec.	Unspec.	Biofuel	Unspec.
Biomass	R	R	R	R
Fossil fuels	I	R	R	I
Hydro- and nuclear power	R		R	R
Total renewable energy	R	R	R	R
Total non-renewable energy	0	R	R	I
Total energy	R	R	R	R
CO_2	I	R	R	I
SO_2	I	0	R	I
NO_x	R	R	R	R
Dust	R	R	R	I
COD	I	I	I	I
Solid waste	I	I	I	I

By inspection of table 15 in its entirety Finnveden and Ekvall arrived at the following conclusions:

1. The environmental effects of waste packaging disposal is influenced by several variables besides the technology choice recycling/incineration, notably by the choice of the alternative energy source (to replace recycled paper).

2. The effect of a change recycling/incineration on the following parameters are not dependent on the other variables studied: Consumption of biomass, electricity demand, total energy consumption.

3. Transportation does not influence the effect of a change of technology recycling/incineration.

These conclusions are mainly in agreement with our findings (compare with table 14 and page 38), although we have identified a few more parameters, the effect upon which are not dependent on other variables, e.g. emission of CH_4, and emission of SO_2. This discrepancy may be due to the fact that our set of independent variables is different from that of Finnveden and Ekvall. Total energy consumption is not calculated as a parameter in our study.

In the paper of Finnveden and Ekvall several more conclusions than the three ones given above are formulated in words. Basically these conclusions describe the effects of the change recycling/incineration on 8 parameters. Some examples of their conclusions are given below (their numbering):

4. *The use of fossil fuels and associated emissions of CO_2 increase with increased recycling, when fossil fuels are assumed to be the alternative heat source.*
6. *When biofuel is the alternative heat source, the emissions of CO_2 follow the same pattern as the use of fossil fuels.*
9. *Emissions of SO_2 follow emissions of CO_2 in most cases.*
10. *Emissions of NO_x are in most cases lower in the case of recycling.*
11. *Emissions of dust and particulates are in most cases lower in the case of recycling. (They are however not always included). The exceptions are some of the studies in which oil is assumed to be the alternative energy source.*
12. *Emissions of COD decreases in most cases with increased recycling.*

This description would in principle correspond to our table 14.

The main innovation in our approach is the conceptual-model starting point, which enables us to study 36 cases with 29 response parameters under 5 pre-selected conditions (X-variables), all from an inventory, which is more limited than the inventory, on which the procedure of Finnveden and Ekvall is based. In principle it should be possible to perform a study of packaging waste treatment without an inventory of any real waste treatment system at all, if the goal is to find out, under which conditions a change of technology will influence selected impact parameters in a desirable way. The mathematical method makes it possible to sort out those parameters which are not affected by a any selected X-variable, and to describe the influence of that X-variable on the other parameters in a tabulated form, with defined conditions of validity.

5. Conclusions

- The interpretation procedure suggested in the introductory survey, i.e. construction of a conceptual model, sensitivity and uncertainty analysis with multivariate methods, and conclusions based on the results of principal component analysis and partial least-square models, can give easily surveyable descriptions of complicated decision-making situations, where the environmental effects of technology changes depend on several pre-conditions.

- A systematic structuring of methodological choices and the use of factorial experimental designs to organise scenario calculations can minimise the necessary inventory work.

- Even good-quality inventory data may be insufficient to draw statistically significant conclusions. Monte Carlo simulations in combination with multivariate evaluation and other statistical tests can determine, whether or not observed differences between two cases are significant, even if the numerical difference is only 20 to 30 %.

6. References

Baumann, H. (1996), ”LCA Use in Swedish Industry”, Int. J. LCA, 1(3), 122 – 126.

Bjuggren, C., M. Almemark, and J. Röttorp (1998), ”Using Chemometrics as a Way to Evaluate Results from Life-Cycle-Analysis Models – A Pre-Study”, Report, Swedish Waste Research Council, to be published.

Box, G. E. P., W.G. Hunter, and J.S. Hunter (1978), ”Statistics for Experimenters. An Introduction to Design, Data Analysis and Model Building”, John Wiley & Sons, New York.

Chatfield, C., and A. J. Collins (1980), ”Introduction to Multivariate Analysis”, Chapman & Hall.

Finnveden, G. and T. Ekvall (1997), “On the Usefulness of LCA in Decision-Making – The Case of Recycling vs. Incineration of Paper”, 5th LCA Case Studies Symposium SETAC-Europe 1997.

Finnveden, G., B. Steen, and J.-O. Sundqvist (1994), ”Kretslopp av pappersförpackningar: materialåtervinning eller energiåtervinning? En miljöstudie baserad på fem verkliga fall”. IVL report B 1128.

Geladi, P. and B.R. Kowalski (1986), ”Partial Least-Squares Regression – A Tutorial”, Analytica Chim. Acta, 185, 1 – 17.

Kennedy, D. J., D. C. Montgomery, and B. H. Quay (1996), ”Stochastic Environmental Life Cycle Assessment Modelling”, Int. J. LCA, 1(4), 199 – 207.

Kennedy, D. J., D. C. Montgomery, D. A. Rollier, and J. B. Keats (1997), ”Assessing Input Data Uncertainty in Life Cycle Assessment Inventory Models”, Int. J. LCA, 2(4), 229 – 239.

ISO, the International Organisation for Standardisation (1997), ”Environmental Management – Life cycle assessment – Principles and framework”, ISO 14040:1997.

Le Téno, J. F. (1997), ”How to Produce Useful Results with LCA: The Equity Model for Building Products”, 2nd International Conf. Buildings and the Environment, Paris, p.125 – 132.

Society of Environmental Toxicology and Chemistry (1994), ”Life-Cycle Assessment Data Quality: A Conceptual Framework”, SETAC and the SETAC Foundation for Environmental Education, Pensacola, Florida.

U.S. Environmental Protection Agency (1992), ”Guidelines for Assessing the Quality of Life-Cycle Inventory Data”, EPA 530-R-95-010, Office of Solid Waste.

Weidema, W. P., and M. S. Wesnoes (1995), ”Data Quality Management for Life Cycle Inventories – An Example of Using Data Quality Indicators”, 2nd SETAC World Congress, Vancouver, Canada.

Appendix

Factorial design of the experiments and calculated response parameters for the two data matrices used to in the sensitivity analysis of the system “Disposal of paper packaging waste”.

Datamatrix 1

experiment	X1	X2	X3	X5	E_H_bio	E_H_energy	E_H_oil	P_paper	NRE_Coal	NRE_diesel	NRE_Fueloil	NRE_NatG	NRE_Peat	NRE_Uran	RE_Biofuel
1	-1	-1	-1	-1	16.7	0	0	0.75	0.0542	0.46699	1.15672	0.0126015	0.0060975	2.168	17.831
2	1	-1	-1	-1	16.7	0	0	0.75	0.0478459	0.467061	1.21805	0.0111242	0.00538267	1.91384	17.1368
3	-1	1	-1	-1	0	0	16.7	0.75	0.0542	0.116157	19.5267	0.0126015	0.0060975	2.168	1.13098
4	1	1	-1	-1	0	0	16.7	0.75	0.0478459	0.116228	19.5881	0.0111242	0.00538267	1.91384	0.436827
5	-1	-1	1	-1	16.7	0	0	0.75	0.0542	0.66099	1.15672	0.0126015	0.0060975	2.168	17.831
6	1	-1	1	-1	16.7	0	0	0.75	0.0478459	0.661061	1.21805	0.0111242	0.00538267	1.91384	17.1368
7	-1	1	1	-1	0	0	16.7	0.75	0.0542	0.310157	19.5267	0.0126015	0.0060975	2.168	1.13098
8	1	1	1	-1	0	0	16.7	0.75	0.0478459	0.310228	19.5881	0.0111242	0.00538267	1.91384	0.436827
17	-1	-1	-1	1	19.4	0	0	0.6	0.0641252	0.53011	0.57361	0.0149091	0.00721408	2.56501	19.4721
18	1	-1	-1	1	19.4	0	0	0.6	0.0678111	0.52986	1.72101	0.0157661	0.00762875	2.71244	19.8593
19	-1	1	-1	1	0	0	19.4	0.6	0.0641252	0.122555	21.9136	0.0149091	0.00721408	2.56501	0.0721408
20	1	1	-1	1	0	0	19.4	0.6	0.0678111	0.122305	23.061	0.0157661	0.00762875	2.71244	0.459288
21	-1	-1	1	1	19.4	0	0	0.6	0.0641252	0.72411	0.57361	0.0149091	0.00721408	2.56501	19.4721
22	1	-1	1	1	19.4	0	0	0.6	0.0678111	0.72386	1.72101	0.0157661	0.00762875	2.71244	19.8593
23	-1	1	1	1	0	0	19.4	0.6	0.0641252	0.316555	21.9136	0.0149091	0.00721408	2.56501	0.0721408
24	1	1	1	1	0	0	19.4	0.6	0.0678111	0.316305	23.061	0.0157661	0.00762875	2.71244	0.459288
33	-1	-1	-1	0	18.05	0	0	0.675	0.0591626	0.49855	0.865165	0.0137553	0.00665579	2.3665	18.6516
34	1	-1	-1	0	18.05	0	0	0.675	0.0578285	0.49846	1.46953	0.0134451	0.00650571	2.31314	18.4981
35	-1	1	-1	0	0	0	18.05	0.675	0.0591626	0.119356	20.7202	0.0137553	0.00665579	2.3665	0.601558
36	1	1	-1	0	0	0	18.05	0.675	0.0578285	0.119266	21.3245	0.0134451	0.00650571	2.31314	0.448057
37	-1	-1	1	0	18.05	0	0	0.675	0.0591626	0.69255	0.865165	0.0137553	0.00665579	2.3665	18.6516
38	1	-1	1	0	18.05	0	0	0.675	0.0578285	0.69246	1.46953	0.0134451	0.00650571	2.31314	18.4981
39	-1	1	1	0	0	0	18.05	0.675	0.0591626	0.313356	20.7202	0.0137553	0.00665579	2.3665	0.601558
40	1	1	1	0	0	0	18.05	0.675	0.0578285	0.313266	21.3245	0.0134451	0.00650571	2.31314	0.448057

RE_Bark	RE_Hydrop	R_Biom	R_Water	G_CH	G_CH4	G_CO	G_CO2	G_HCl	G_N2O	G_NOX	G_Dust
0	0.63685	0	0	2.08E-05	0.0370002	0.0178978	0.127564	0.0010662	5.69E-07	0.00338662	0.000301366
0.605	0.56219	0	0	2.19E-05	0.0370002	0.0174534	0.132672	0.00102498	5.02E-07	0.00338668	0.000295786
0	0.63685	0	0	0.00032135	0.0370002	0.00130796	1.43232	6.42E-05	5.69E-07	0.00298565	0.000549181
0.605	0.56219	0	0	0.000322463	0.0370002	0.00086358	1.43743	2.30E-05	5.02E-07	0.00298571	0.000543601
0	0.63685	0	0	2.08E-05	0.0370002	0.0178997	0.142813	0.0010662	5.69E-07	0.00363882	0.000303306
0.605	0.56219	0	0	2.19E-05	0.0370002	0.0174553	0.14792	0.00102498	5.02E-07	0.00363888	0.000297726
0	0.63685	0	0	0.00032135	0.0370002	0.0013099	1.44757	6.42E-05	5.69E-07	0.00323785	0.000551121
0.605	0.56219	0	0	0.000322463	0.0370002	0.00086552	1.45268	2.30E-05	5.02E-07	0.00323791	0.000545541
2.388	0.753471	0	7.03	1.02E-05	0.0592002	0.0204949	0.0857215	0.001164	6.73E-07	0.00387911	0.000315853
0.605	0.796781	0	7.03	3.09E-05	0.0592002	0.0201797	0.177144	0.00118698	7.12E-07	0.00395234	0.000351493
2.388	0.753471	0	7.03	0.000359442	0.0592002	0.00122292	1.60143	0	6.73E-07	0.00341331	0.000603734
0.605	0.796781	0	7.03	0.00038009	0.0592002	0.000907674	1.69285	2.30E-05	7.12E-07	0.00348653	0.000639373
2.388	0.753471	0	7.03	1.02E-05	0.0592002	0.0204969	0.10097	0.001164	6.73E-07	0.00413131	0.000317793
0.605	0.796781	0	7.03	3.09E-05	0.0592002	0.0201816	0.192392	0.00118698	7.12E-07	0.00420454	0.000353433
2.388	0.753471	0	7.03	0.000359442	0.0592002	0.00122486	1.61668	0	6.73E-07	0.00366551	0.000605674
0.605	0.796781	0	7.03	0.00038009	0.0592002	0.000909614	1.7081	2.30E-05	7.12E-07	0.00373873	0.000641313
1.194	0.695161	0	3.515	1.55E-05	0.0481002	0.0191964	0.106643	0.0011151	6.21E-07	0.00363287	0.00030861
0.605	0.679485	0	3.515	2.64E-05	0.0481002	0.0188165	0.154908	0.00110598	6.07E-07	0.00366951	0.000323639
1.194	0.695161	0	3.515	0.000340396	0.0481002	0.00126544	1.51688	3.21E-05	6.21E-07	0.00319948	0.000576457
0.605	0.679485	0	3.515	0.000351276	0.0481002	0.000885627	1.56514	2.30E-05	6.07E-07	0.00323612	0.000591487
1.194	0.695161	0	3.515	1.55E-05	0.0481002	0.0191983	0.121891	0.0011151	6.21E-07	0.00388507	0.00031055
0.605	0.679485	0	3.515	2.64E-05	0.0481002	0.0188185	0.170156	0.00110598	6.07E-07	0.00392171	0.000325579
1.194	0.695161	0	3.515	0.000340396	0.0481002	0.00126738	1.53213	3.21E-05	6.21E-07	0.00345168	0.000578397
0.605	0.679485	0	3.515	0.000351276	0.0481002	0.000887567	1.58039	2.30E-05	6.07E-07	0.00348832	0.000593427

G_SO2	AQ_BOD	AQ_COD	AQ_TSS	W_Ash	W_Reject
0.00101546	0.0053	0.0139766	0.0005	0.0107466	0.25
0.00103104	0.00345	0.0143767	0.009	0.0125024	0.25
0.00705534	0.0053	0.0139929	0.0005	0.00086017	0.25
0.00707092	0.00345	0.014393	0.009	0.00261604	0.25
0.0010368	0.0053	0.0139768	0.0005	0.0107466	0.25
0.00105238	0.00345	0.0143769	0.009	0.0125024	0.25
0.00707668	0.0053	0.0139931	0.0005	0.00086017	0.25
0.00709226	0.00345	0.0143932	0.009	0.00261604	0.25
0.000872839	0.003089	0.0280821	0.0325	0.0203329	0.4
0.00130792	0.005739	0.0204832	0.009	0.014158	0.4
0.00788922	0.003089	0.028101	0.0325	0.00884812	0.4
0.00832431	0.005739	0.0205021	0.009	0.00267317	0.4
0.000894179	0.003089	0.0280823	0.0325	0.0203329	0.4
0.00132926	0.005739	0.0204834	0.009	0.014158	0.4
0.00791056	0.003089	0.0281012	0.0325	0.00884812	0.4
0.00834565	0.005739	0.0205023	0.009	0.00267317	0.4
0.00094415	0.0041945	0.0210293	0.0165	0.0155397	0.325
0.00116948	0.0045945	0.0174299	0.009	0.0133302	0.325
0.00747228	0.0041945	0.0210469	0.0165	0.00485414	0.325
0.00769761	0.0045945	0.0174475	0.009	0.00264461	0.325
0.00096549	0.0041945	0.0210295	0.0165	0.0155397	0.325
0.00119082	0.0045945	0.0174301	0.009	0.0133302	0.325
0.00749362	0.0041945	0.0210471	0.0165	0.00485414	0.325
0.00771895	0.0045945	0.0174477	0.009	0.00264461	0.325

experiment	X1	X4	X5	E_H_bio	E_H_energy	E_H_oil	P_paper	NRE_Coal	NRE_diesel	NRE_Fueloil	NRE_NatG	NRE_Peat	NRE_Uran	RE_Biofuel
1	-1	-1	-1	16.7	0	0	0.75	0.0542	0.46699	1.15672	0.0126015	0.0060975	2.168	17.831
2	1	-1	-1	16.7	0	0	0.75	0.0478459	0.467061	1.21805	0.0111242	0.00538267	1.91384	17.1368
9	-1	1	-1	0	16.7	0	0.75	0.05892	0.716827	0.094272	0.0136989	0.0066285	2.3568	8.52628
10	1	1	-1	0	16.7	0	0.75	0.2034	0.701794	0.87294	0.0472905	0.0228825	8.136	5.20882
17	-1	-1	1	19.4	0	0	0.6	0.0641252	0.53011	0.57361	0.0149091	0.00721408	2.56501	19.4721
18	1	-1	1	19.4	0	0	0.6	0.0678111	0.52986	1.72101	0.0157661	0.00762875	2.71244	19.8593
25	-1	1	1	0	19.4	0	0.6	0.047136	0.576431	0.0754176	0.0109591	0.0053028	1.88544	6.82103
26	1	1	1	0	19.4	0	0.6	0.16272	0.564405	0.698352	0.0378324	0.018306	6.5088	4.16706
33	-1	-1	0	18.05	0	0	0.675	0.0591626	0.49855	0.865165	0.0137553	0.00665579	2.3665	18.6516
34	1	-1	0	18.05	0	0	0.675	0.0578285	0.49846	1.46953	0.0134451	0.00650571	2.31314	18.4981
41	-1	1	0	0	18.05	0	0.675	0.053028	0.646629	0.0848448	0.012329	0.00596565	2.12112	7.67366
42	1	1	0	0	18.05	0	0.675	0.18306	0.633099	0.785646	0.0425615	0.0205942	7.3224	4.68794

Datamatrix 2

RE_Bark	RE_Hydrop	R_Biom	R_Water	G_CH	G_CH4	G_CO	G_CO2	G_HCl	G_N2O	G_NOX	G_Dust	G_SO2	AQ_BOD
0	0.63685	0	0	2.08E-05	0.037	0.017898	0.127564	0.001066	5.69E-07	0.003387	0.000301	0.001015	0.0053
0.605	0.56219	0	0	2.19E-05	0.037	0.017453	0.132672	0.001025	5.02E-07	0.003387	0.000296	0.001031	0.00345
1.509	0.69231	0.75	0	1.62E-06	0.00222	0.010182	0.066715	0.000518	6.19E-07	0.004405	0.000193	0.000465	0.002268
0.75	2.38995	0.75	0	1.54E-05	0.002776	0.006408	0.122949	0.000309	2.14E-06	0.004168	0.000161	0.000588	0.001148
2.388	0.753471	0	7.03	1.02E-05	0.0592	0.020495	0.085722	0.001164	6.73E-07	0.003879	0.000316	0.000873	0.003089
0.605	0.796781	0	7.03	3.09E-05	0.0592	0.02018	0.177144	0.001187	7.12E-07	0.003952	0.000351	0.001308	0.005739
1.2072	0.553848	0.6	0	1.30E-06	0.001776	0.008426	0.637905	0.000406	4.95E-07	0.004492	0.000178	0.000292	0.001814
0.6	1.91196	0.6	0	1.24E-05	0.002221	0.005408	0.682893	0.000239	1.71E-06	0.004302	0.000153	0.00039	0.000918
1.194	0.695161	0	3.515	1.55E-05	0.0481	0.019196	0.106643	0.001115	6.21E-07	0.003633	0.000309	0.000944	0.004195
0.605	0.679485	0	3.515	2.64E-05	0.0481	0.018817	0.154908	0.001106	6.07E-07	0.00367	0.000324	0.001169	0.004595
1.3581	0.623079	0.675	0	1.46E-06	0.001998	0.009304	0.35231	0.000462	5.57E-07	0.004449	0.000185	0.000379	0.002041
0.675	2.15096	0.675	0	1.39E-05	0.002498	0.005908	0.402921	0.000274	1.92E-06	0.004235	0.000157	0.000489	0.001033

AQ_COD	AQ_TSS	W_Ash	W_Reject
0.013977	0.0005	0.010747	0.25
0.014377	0.009	0.012502	0.25
0.004994	0.000488	0.050603	0.015
0.010124	0.000998	0.046043	0.01875
0.028082	0.0325	0.020333	0.4
0.020483	0.009	0.014158	0.4
0.003995	0.00039	0.097482	0.012
0.0081	0.000798	0.093834	0.015
0.021029	0.0165	0.01554	0.325
0.01743	0.009	0.01333	0.325
0.004495	0.000439	0.074043	0.0135
0.009112	0.000898	0.069939	0.016875

IVL Svenska Miljöinstitutet AB

IVL Swedish Environmental Research Institute Ltd

Box 210 60, SE-100 31 Stockholm
Hälsingegatan 43, Stockholm
Tel: +46 8 598 563 00
Fax: +46 8 598 563 90

Box 470 86, SE-402 58 Göteborg
Dagjämningsgatan 1, Göteborg
Tel: +46 31 725 62 00
Fax: +46 31 725 62 90

Aneboda, SE-360 30 Lammhult
Aneboda, Lammhult
Tel: +46 472 26 20 75
Fax: +46 472 26 20 04

www.ivl.se