Evaluation of assessment tools of concept maps from teaching in sustainable development

***Jonas Sjöblom**, department of Applied Mechanics, division of Combustion*
Chalmers University of technology, SE-412 96 Gothenburg, Sweden.
Telephone: +46-31-772 1389, e-mail: Jonas.sjoblom@chalmers.se

This report was a part of a pedagogical project by Jonas Sjöblom 2013. Published on-line August 2014 in CPL as document 20899, available at https://publications.lib.chalmers.se/publication/200899

Table of Contents

Abstract

The use of concept maps have been evaluated and analysed in conjunction to other Teaching and learning activities (TLA's) and to examination results. The results showed how the different aspects of sustainability can be represented in a plot from multivariate analysis of this data. From interpretation of these plots, suggestions could be made on what TLA's supports different type of sustainability aspects and a data-driven development of the course was enabled. However, the interpretation was based on that concept maps do convey information about student learning and this depends on a reliable methodology. In this study the evaluation method of concept maps itself is scrutinized and aspects of reliability and validity is analysed.

The results show that the applied methodology is both reliable and valid to a certain extent. Reflections and recommendations are given for future work when combining concept map data in conjunction to other student data.

Introduction

Education for sustainable development (ESD)

The global society is facing big challenges in the coming years. Although the living conditions in general have been greatly improved for people in most parts of the world over the last century (e.g. quantified in terms of Human Development Indicators, HDI [UNDP 2011]) , there are still many aspects of human well-being that need to be developed further. On a global level, "sustainable development" (SD) is the term used to describe development that leads to fulfilment of basic human needs for both living and for coming generations [Brundtland 1987]. Even though the definition of sustainable development is discussed, there is today more or less consensus that current development is far from sustainable, and that SD is the overarching goal for the present society. This implies drastic changes from current activities concerning not only environmental and economic issues, but also concerning social and institutional aspects. Since the current world situation is a result of how humans have been thinking and acting over a long time period, manifested in locked in by different societal structures, the role of education is becoming increasingly important. "Education for sustainable development" (ESD) is highlighted by the UN as one of the major focus areas in addressing sustainable development by the proclamation of the "Decade for Sustainable Development 2005-2014 [UNESCO 1995].

ESD puts new demands on both professionals and on individuals. It emphasizes the interconnectivity of different systems in the world and tries to build competencies that help in understanding the consequences of different actions. ESD is also about selection, development and implementation of solutions that have a larger potential to contribute to sustainable development. Competencies that are often mentioned are systemic thinking, critical and creative thinking, ability to shift perspectives and anticipatory thinking. These competencies are sometimes summarized as action competence or change agency. ESD is still hampered by a lack of good descriptions of the required competencies and studies on which teaching and learning activities (TLAs) that provide these competencies. Furthermore, there is a lack of methods for assessment of ESD competencies, which is needed for both examination of students and for determining the efficiency of the TLAs. Thus, in order to be able

to be able to constructively align intended learning outcomes, TLAs and assessment, development is needed in all these areas.

Concept maps and ESD

A concept map (CM) is a graph of nodes representing concepts and labelled lines denoting the relations between a pair of nodes [Ruiz-Primo 1996]. The concept maps builds on cognitive theory proposed by Ausbel and Deese and the term concept map was coined by Novak [Novak 1984]. They described concept maps as hierarchical maps of concepts (or terms) within a subject domain, with superordinate concepts at the top and subordinate levels below. Later studies have regarded the CM to be less hierarchical and more of a (non-hierarchical) network of cognitive knowledge structures.

As an assessment tool, the CM is considered as a "procedure to measure the structure of a student's declarative knowledge" [Ruiz-Primo 1996]. (Note the use or the word "assessment" and not "test" to clarify that a test of student knowledge requires several pieces of information and CM is one of them.) The assessment is a combination of three (independent) components:

- The task
- The response format
- The scoring system

All these three components need to be considered in assessing CM's. Concepts maps have been more frequently used as instructional tools than for assessing learning. In the paper by Ruiz-Primo [Ruiz-Primo 1996], the existing literature on the use of concept maps is assessment of learning was reviewed. They found big variations in the task, the format and in the scoring system. They concluded that issues of reliability and validity needs to be addressed when assessing concept maps.

In this study, these issues are addressed whining the practical possibilities for the course and conclusions about reliability and validity are drawn in order to increase the confidence of the findings about the use of concept maps in ESD.

Master course in Global Chemical Sustainability

The course "Global Chemical Sustainability" (KBT140) in study is a 7.5 hec masters course at the Masters Programme of Innovative and Sustainable Chemical engineering (MPISC) and has been given for more than 10 years. Besides lectures It has 8 different activities (TLA) and has a split examination (50% written exam, 50% TLAs). The course is well "aligned" [Biggs 2007] where the 6 course objectives are assessed with 6 dedicated questions on the written exam.

From the course PM:

Purpose: This course provides the students with an understanding of the effects on sustainability of the actions of an engineer and with means to identify appropriate changes.

Learning outcomes: After successfully completing the course, the students should be able to:

1. Describe the function and state of the natural systems in the world
2. Explain how the human society, in particular chemical and chemical engineering industry, affects and depends on natural systems
3. Describe international and regional work on environment and sustainable development on governmental and non-governmental level

4. Identify appropriate tools and strategies for sustainable development in society, in particular industry
5. Analyze sectors of society, in particular chemical and chemical engineering industry, and formulate appropriate strategies for sustainable development
6. Describe the importance of including different stakeholders and perspectives in discussions on sustainable technology development

The course also includes experiences of using skills from earlier courses, e.g.

- Oral and written presentation
- Working in team
- Project work
- Critical review

The necessary information from course PM as a background for this study can be found in Appendix A.

Concept maps have been used at the first occasion as well as at the final occasion. The task was formulated as writing a concept map (CM) (no concepts given) on a blank space (only the concept "sustainability" written in the center) during 15 minutes. The data has been analysed by the university of Catalan [Segalas 2008; Segalàs 2010] where CM results (pre- and post-TLA) from 10 different European universities were compared and analysed. However, the data analysis was based on averaging the data over a whole course and a coupling to individual TLA and examination results was not possible.

Concept map evaluation

The evaluation methodology was developed previously [Segalas 2008; Segalàs 2010] and consisted of counting all concepts, categorising each concept to one of ten categories and counting the number of (valid) links between concepts belonging to different categories. Two different category taxonomies were used in this study:

Table 1. Taxonomy used for categorizing each concept in the analysis.

No	Category	No (4)	Category (4)
1.	**Environment**	1.	**Environmental**
2.	**Resources**		
3.	**Social impact**	2.	**Social**
4.	**Values**		
5.	**Future**		
6.	**Unbalances**		
7.	**Technology**	3.	**Economic**
8.	**Economy**		
9.	**Education**	4.	**Institutional**
10.	**Stakeholders**		

The concepts were categorised to the 10-cat taxonomy and the categorisation according to the 4-cat taxonomy was made by summing the concepts according to the table above.

Multivariate data analysis (MVDA)

Multivariate data analysis is a method to analyse large amounts of data. In contrast to most other methods (like multivariate linear regression analysis), this method is suitable to handle co-linearity, i.e. when there are correlations in the data. By projecting the data on a low-dimensional hyper plane (called latent variables) an approximation of the original data is obtained but with better possibilities of interpretation and data mining is obtained. A short introduction of MVDA is presented in appendix B (an extract from my thesis [Sjöblom 2009]). In a MVDA model, a matrix X is decomposed into two new sets of vectors called scores vectors (T) and loading vectors (P)

$$X = TP' \qquad (1)$$

The nice feature of this model is that is has fewer dimensions that the original model. The matrix X in this study has typically 40-70 columns (variables) which will make data analysis if performed one variable at the time very time consuming. Furthermore, due to the inherent correlation structure, the statistical analysis of any inter-relation will be difficult. The resulting scores (T) and loading (P) matrices on the other hand will have fewer dimensions (in this study typically 2-4 columns) and it thus easier to analyse by scatter plots.

Simple example using only CM data

A Principal Component Analysis (PCA) model was made for illustrative purpose. Here only CM data was used and the data set contained 108 rows (students) and 14 variables (before & after, relative abundance of categories according to the 4-category taxonomy, the number of concepts, the number of covered categories and the number of intercategory links). 2 components were calculated and this model could explain 36% of the variation in the data.

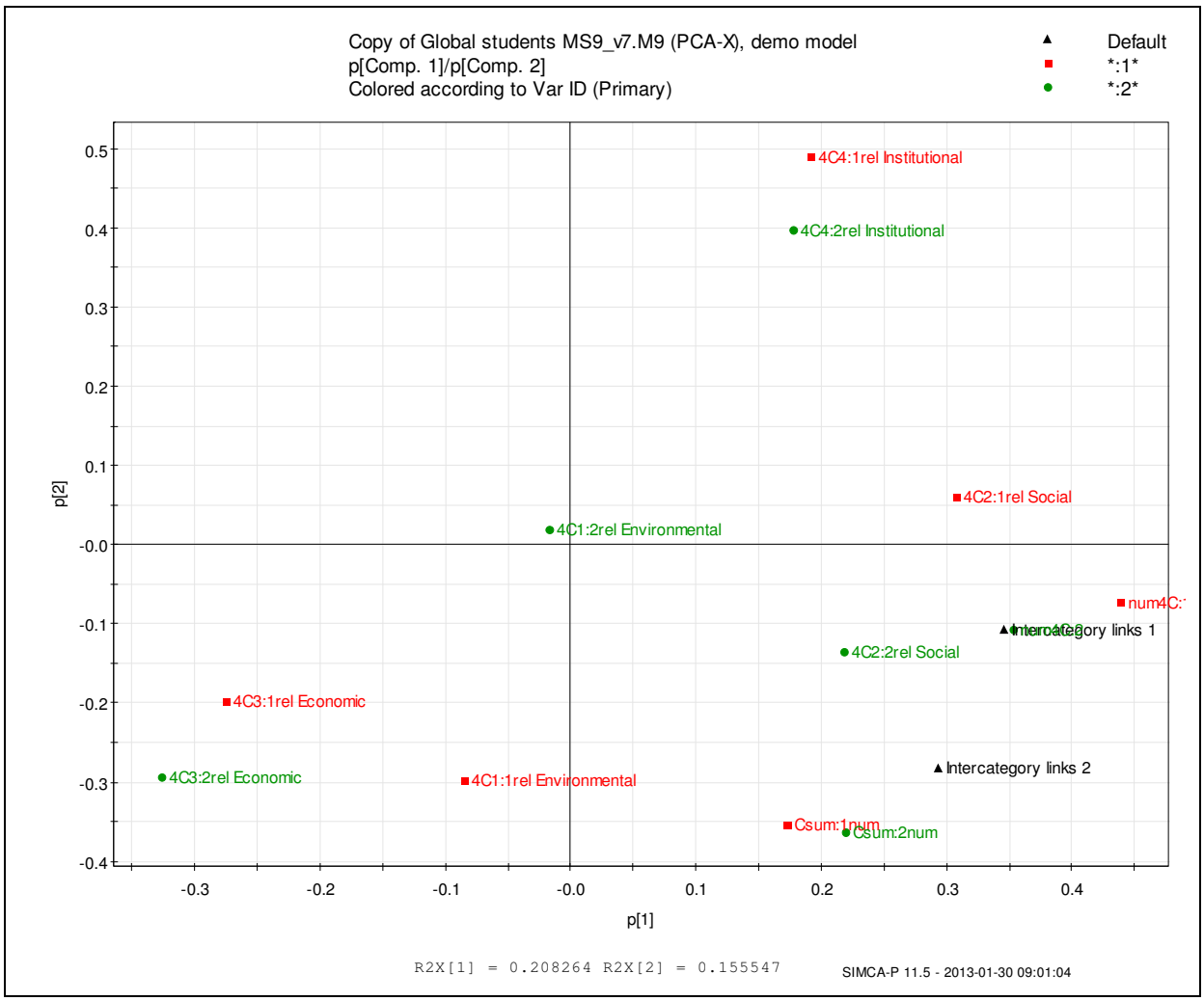


Figure 1 Example of a loading scatter plot of the first two components in a PCA model. Items indicated with "1" correspond to pre-course and "2" corresponds to CM after the course.

The loadings plot shows the linear combinations of the original variables that makes the new (linearly independent) variables (components). The interpretation is done such as objects that are close are "similar" and objects that forms an angle are not similar (un-correlated). Here the number of intercategory links is similar (correlated) to the relative amounts of concepts belonging to the "social" category. This is because the social category is less frequent and the students that put concepts related to social aspects are more likely to have more inter-category links.

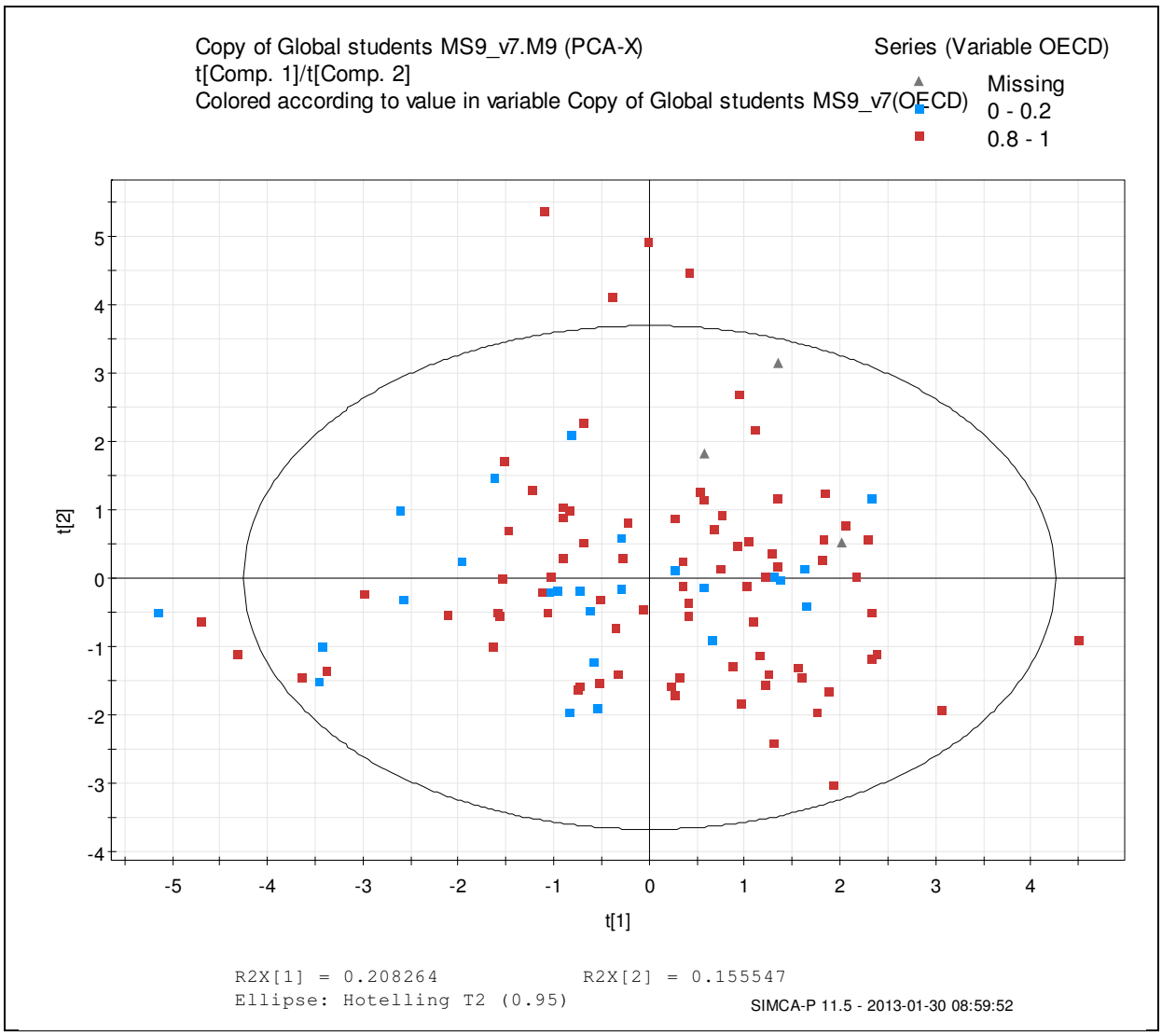


Figure 2. Example of a scores scatter plot of the first two components in a PCA model. The colours represent student origin.

The scores are the new variables and is the projection of each student data onto the principal component (defined by the loadings plot shown above). In the scores scatter plot, each object corresponds to one student. The interpretation is similar to the loadings plot, i.e. objects that are close in the plot are similar (has high correlation) and object that forms an angle are not similar (un-correlated. In this figure, the colors indicate whether the student comes from a OECD country or not. We can see a small trend that OECD country is situated in the bottom right quadrant. The connection to the loading plot above gives the conclusion (based on correlation!) that student from OECD countries perform better (e.g. have more intercategory links, which is considered to show a more complex understanding of SD) than other students.

Previous studies to couple SD competencies with TLA's

The student data (incl. student home country info) from TLA, CM and examination was arranged according to Figure 3

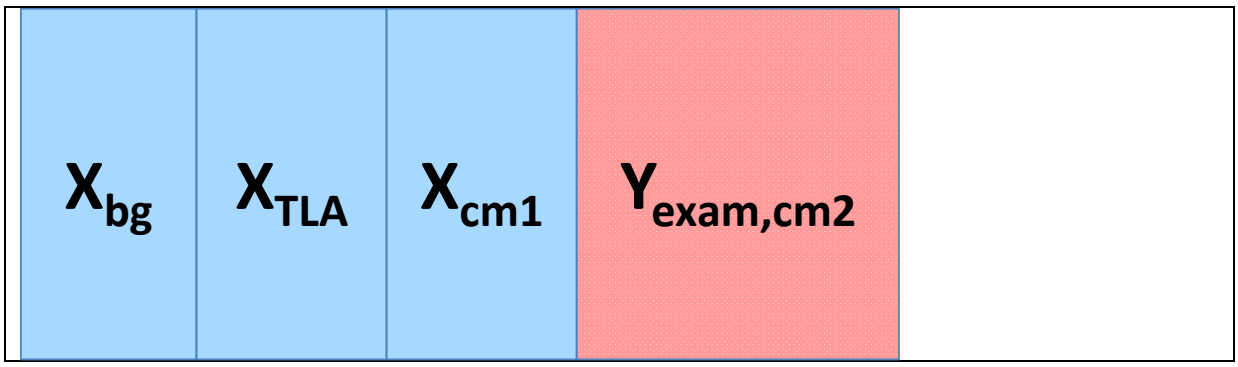


Figure 3. Data structure in the study. Each row represents one student, each column represents info for student background, TLA, exam and CM.

Different types of MVDA models were analysed:

- Principal Component Analysis (PCA): No specific regressor variable ("y") was assigned, all variable was treated as "x" and the correlation structure was revealed.
- Partial Least Squares (PLS): result variables (examination and CM from the last occasion) was used as "y" and the other variable were treated as input variables. Here the analysis shows which factors are most correlated with learning outcomes (as measured/defined by "y").

The input data could be used to explain some of the differences in course performance and the correlation between the different course activities and the examination results could be displayed. The impact of how the groups were composed (mainly mix of nationalities) on the evolution of conceptual maps was analyzed. By using concept maps as a tool to monitor ESD competences on a higher level, valuable feedback for further course development was obtained. The detailed results was presented at the KUL2012 conference [Sjöblom 2012] and will be submitted in a full-length format in 2013.

Since this study as a new combination of different methods (use of concept maps in conjunction to MVDA) some concern about the reliability and validity of the method should be addressed.

Objectives

The aim is to assess the validity and reliability issues from the method used. The definitions and assessment methods will follow [Ruiz-Primo 1996] (and also ev. [Rice 1998]) and here the following issues will be covered:

- Reliability:
  - **inter-rater reliability**: Does it matter which teacher is correcting the concept map?
- Validity:
  - **Content validity:** Use of "experts" to infer that the mind maps are reflecting what we want to measure, in this case student cognitive structure of SD.
  - **Concurrent validity:** Use of TLA to infer the correlations and interferences between TLA's and CM in order to evaluate the course in specific and ESD in general
- Sensitivity
  - **Instructional sensitivity:** Does teaching improve the quality of concept maps? (Self-evident, but still should be addressed!). Moreover, does it matter how the task and the format of the CM are defined?

- **Known group differences:** How sensitive to known differences (gender, country of origin) is the analysis and does it help our understanding of student learning?

The objective in this study is to verify that the combination of using CM, TLA's and MVDA as an assessment method of student performance is reliable so that the interpretations and conclusion regarding course improvements is considered useful based on collected data ("empirical evidence").

"Experimental"

Inter-rater reliability

Here the CM data from 2008 was used and 46 concept maps (pre-course only!) together with other student data were used.

Different types of inter-rater reliability can be considered. One could consider the correlation coefficient for each and every variable (i.e. #concepts in each category and # intercategory links) and the correlation was generally very high, on average 88%. A better way to represent this correlation is via a loading plot, see Figure 4. Here the data set for the two evaluators are placed "side-by-side", i.e. each row corresponds to the same student and it is by inspecting the loading plot (i.e. similarities between variables) that the two evaluators can be compared.

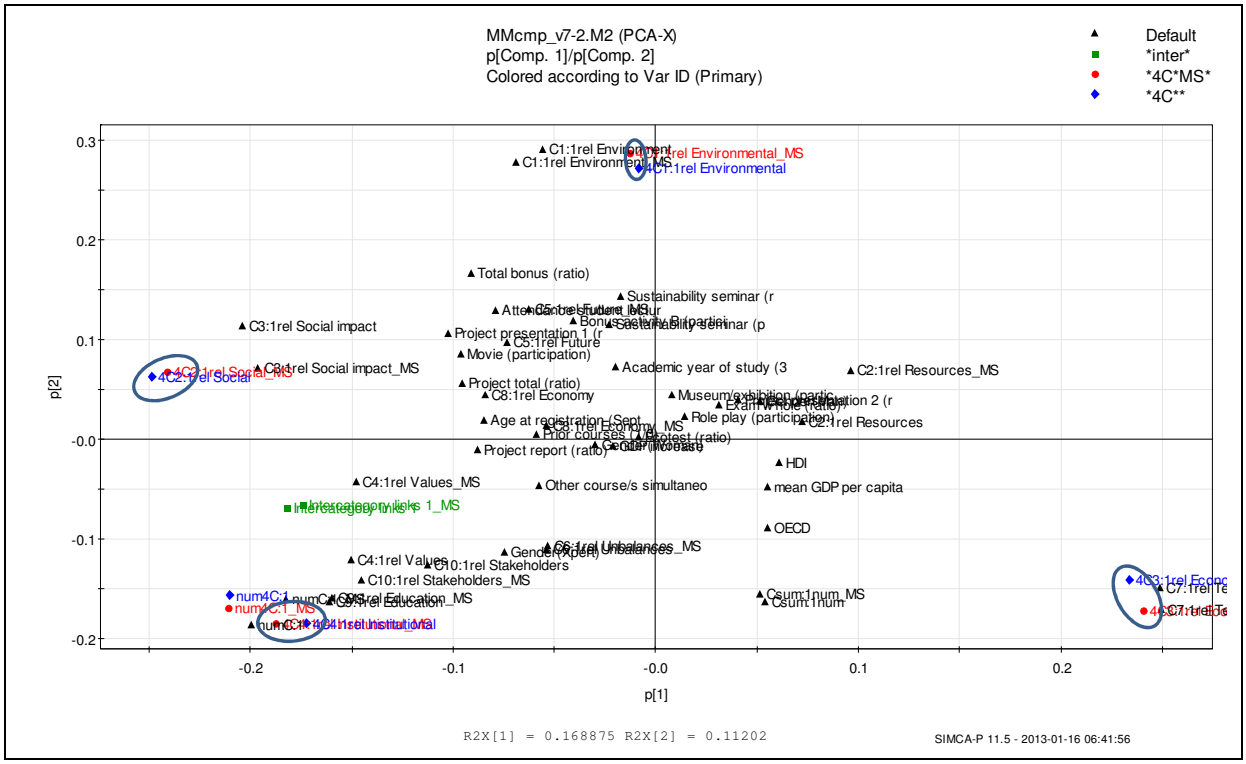


Figure 4. Loading plot of the first two components (describing 28% of the total variance of the data). PCA model had 4 components (R2=63%). The four categories are high-lightened as an example.

In Figure 4, the loading coefficients for CM-data for the two different evaluators are displayed. It is clear that the coefficients for the different evaluators are close, i.e. they have similar latent-variable structure.

A different approach would be to put the data "top on top" instead of "side by side", see Figure 5. In this mode, the variables are the same (e.g. relative distribution of categories) but every student is contributing with two rows, one for each evaluator. Here the scores plot is used to analyse similarities between evaluations. The same students evaluated by different evaluators are marked as different classes in Figure 5.

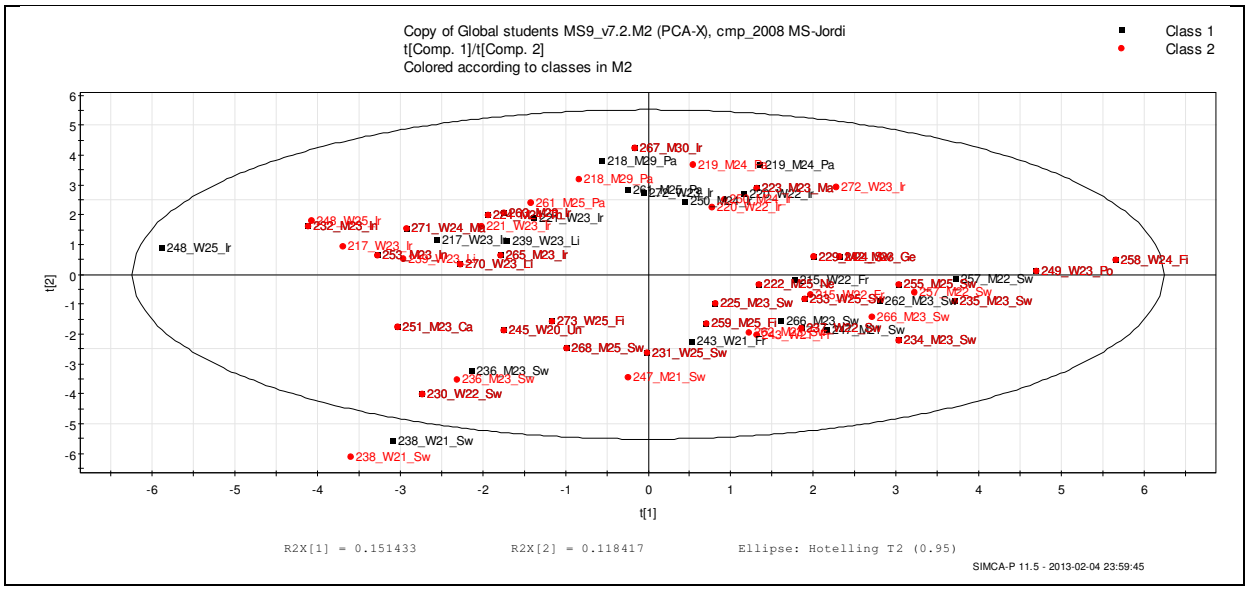
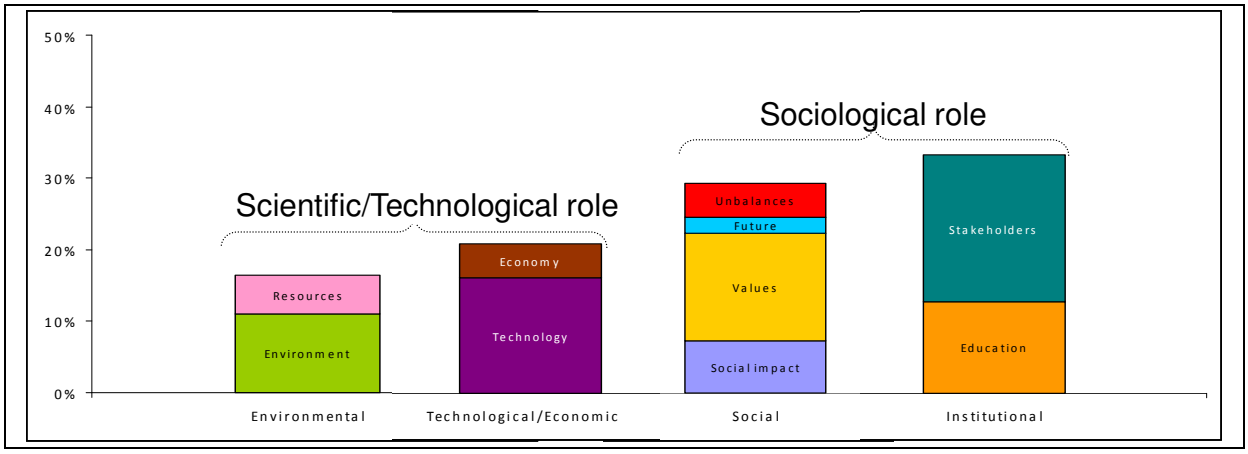


Figure 5. Scores plot from a PCA model "stacked on top". The two components explain 27% (PCA model had 6 comp, R2X=60%)

From Figure 5, it is clear that (again the two evaluators are giving similar results. This approach (top-on-top) is also useful for validation studies as will be shown below.

Content validity

It is useful to use experts to judge the representativeness of concepts and the accuracy of maps within the subject domain (here domain of Sustainable Development). A group of experts in SD have contributed with data that have been used for content validity [Segalàs 2010]. The main finding is that experts tend to have a more even distribution of categories whereas students (before the course) tend to consider SD with mainly environmental issues. This has also been the foundation when interpreting the loading plots for concurrent validity as well for instruction sensitivity (see below). The connection between traditional analysis and MVDA is illustrated in the figures below.



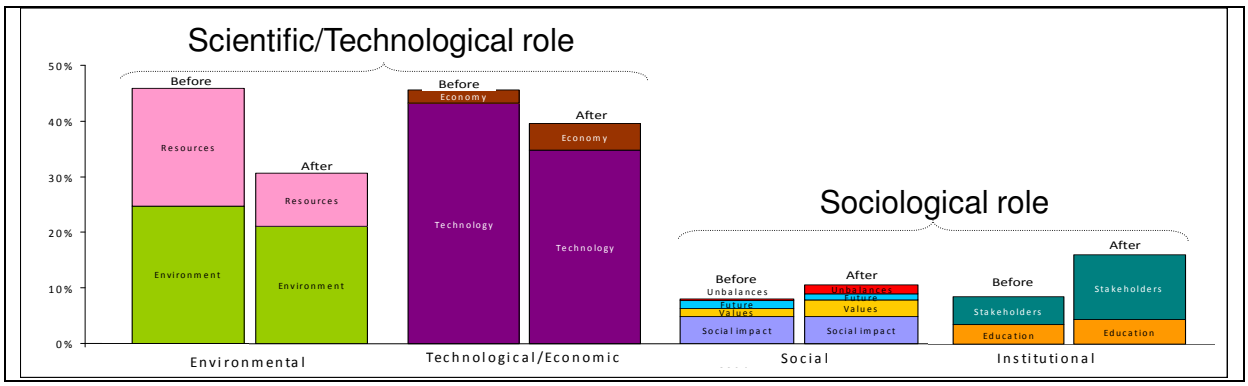


Figure 6. Bar charts illustrating the distribution of categories from experts (top panel) and students (bottom panel).

In Figure 6, it is shown how the distribution "after" tends to approach the more even distribution of the experts after the course. A PCA model was made to illustrate the same trend.

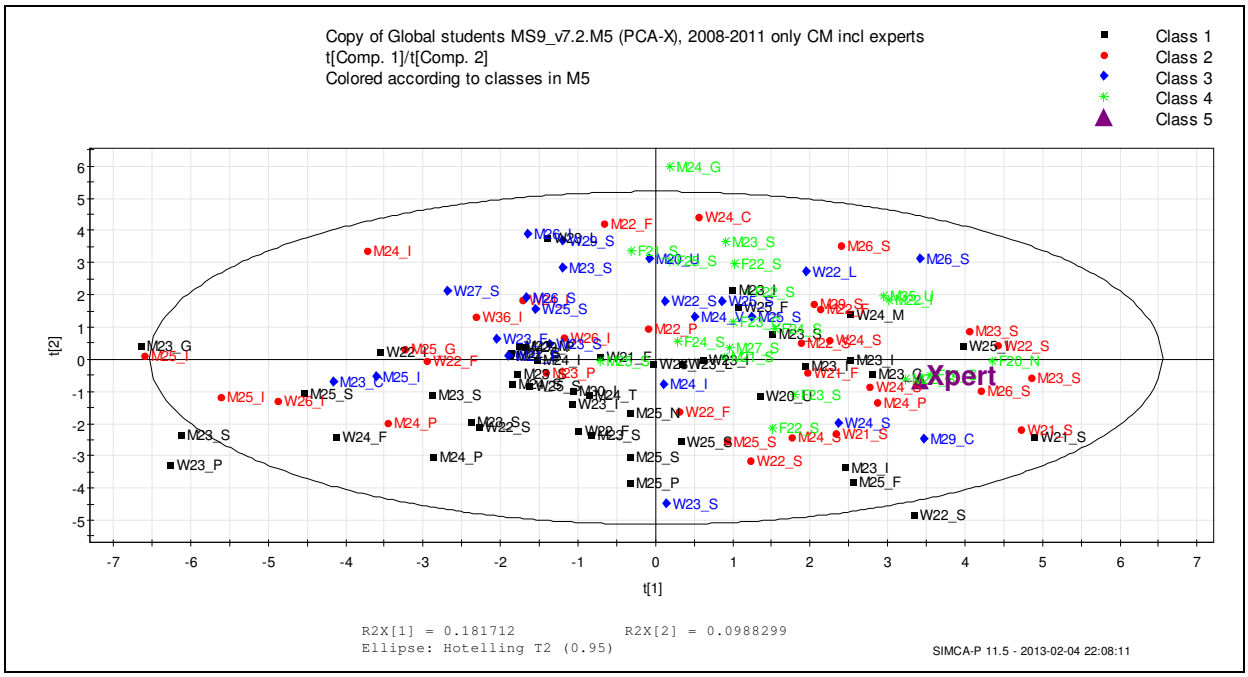


Figure 7. A scores scatter plot where a different year has one class and the "expert" is class 5 situated in the right side.

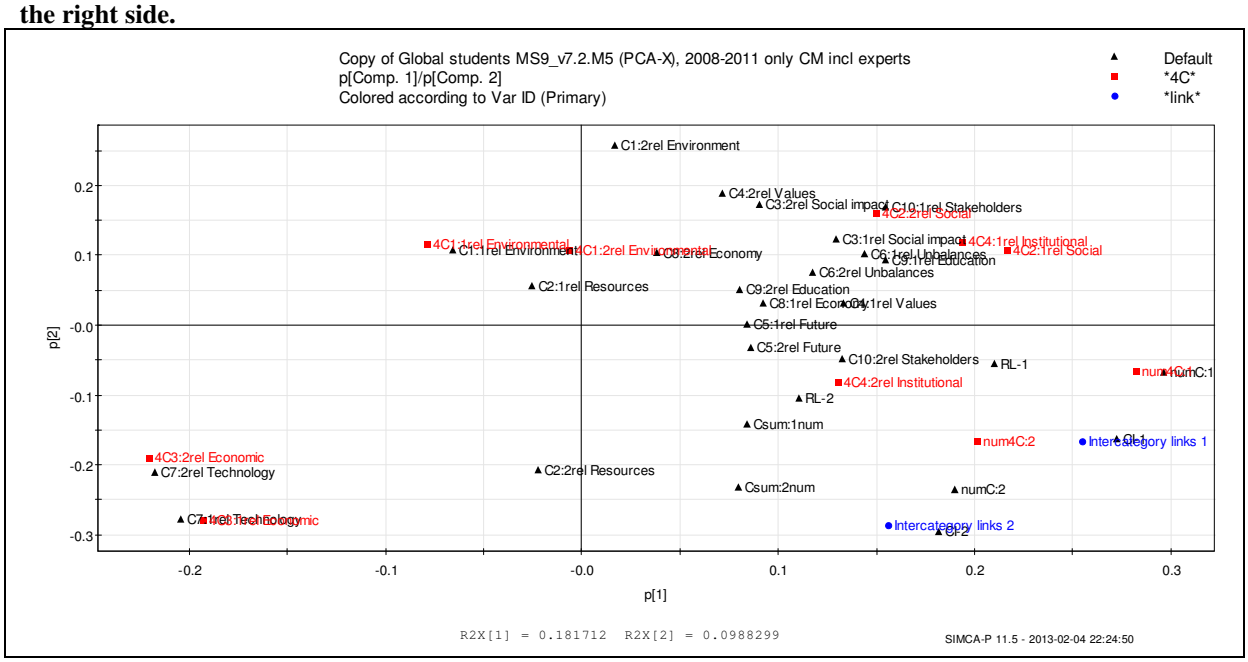


Figure 8. A loading scatter plot for the same components as the scores scatter plot above. The 4-category taxonomy is high-lightened in red and the intercategory links is high-lightened in blue.

The interpretation from Figure 7 and Figure 8 is that the Expert is located on the right hand side and in the loading plot the right hand side is represented by categories of the "sociological role" (as displayed in Figure 6 above). Moreover (not shown in the bar charts) the intercategory links is correlated with sociological categories. This implies that students (and experts) that have a high portion of sociological categories also has many intercategory links, which is an indication of better understanding of SD.

Concurrent validity and Known group differences

Concurrent validity is the measure on correlation between CM and other TLA's and examination results. By analysing the correlation structure, insight to what parts of the course works well and suggestion for improvements can be made. By analysing the correlation structure of known group differences, insights and interesting observations may be done that may further improve the teaching of SD. In order to address these issues (concurrent validity and known group differences) a PCA model was made that included student background info, TLA results, examination results as well as CM data (only 4-category taxonomy for the sake of plot clarity). (This model had 3 components explaining 34% of the variation in the data. Only the two first components is displayed here).

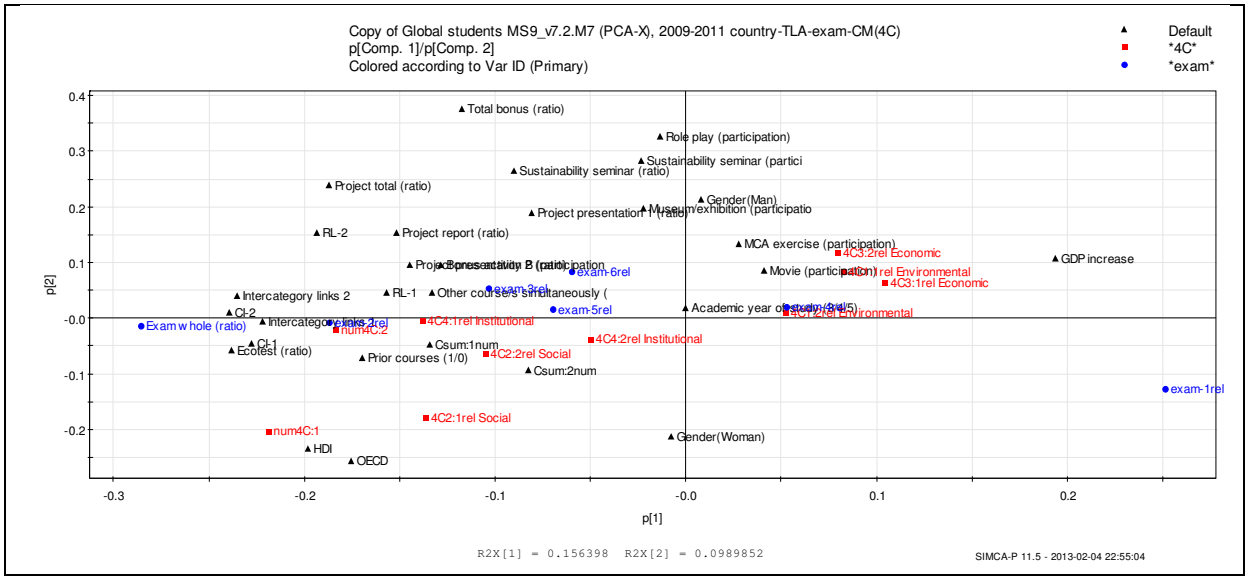


Figure 9. A loading plot of a PCA model that includes student background, TLA's, examination results and CM data. The CM categories are marked in red, examination results in blue.

From Figure 9 a number of conclusions can be drawn:

- Student background correlation indicates no gender correlation with examination results nor with the number of intercategory links. Students from "developed" countries (indicated by e.g. high HDI) tends to perform better than student from countries with high GDP increase (i.e. coming from a low GDP).
- Of the TLA's that has the higher correlation with the "desired outcomes" (high examination results and many intercategory links) is the project (not surprisingly). Also the ecotest (diagnostic test at the very beginning) and "prior courses" is highly correlated. This indicates that students with a good background is favoured to get better results. This is not very surprising since SD competencies should require long lasting reflections on SD.
- The examination results correlated with intercategory links, which is natural. The relative distribution of the examination questions (which was linked to the learning objectives)

showed that the different learning goals correlated with the CM data. E.g. objective no1 (related to environmental issues) is most correlated to environmental categories.

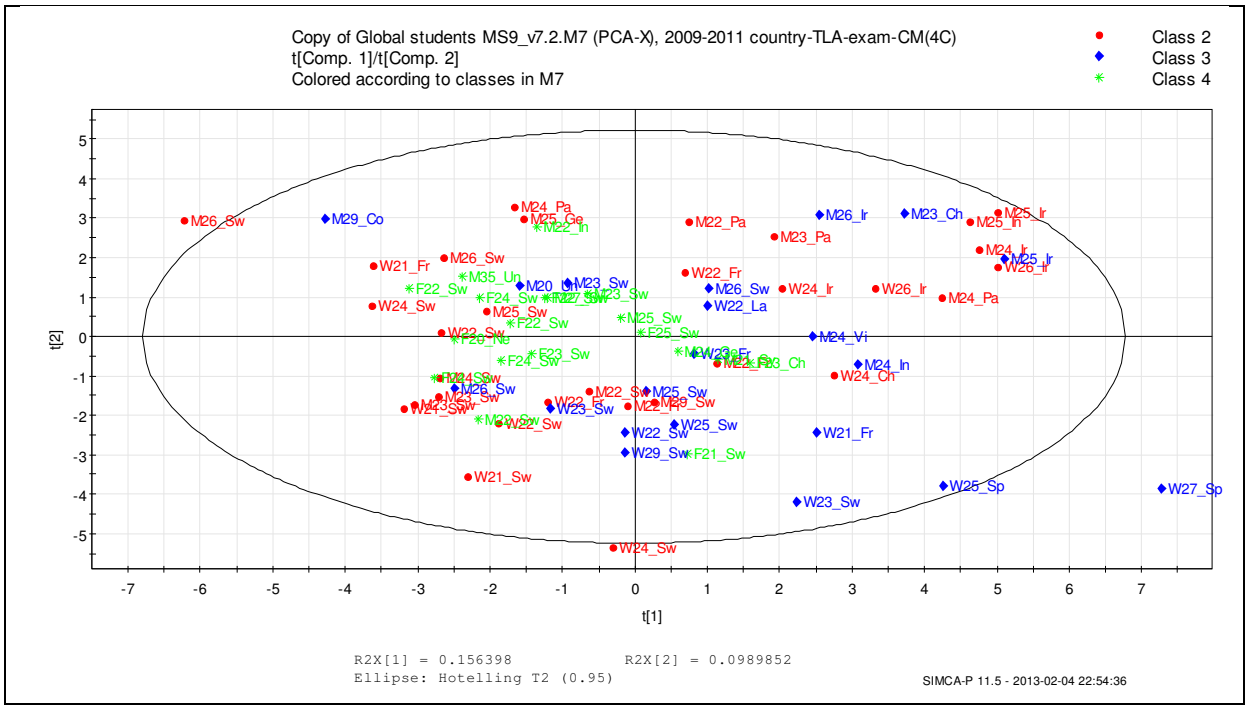


Figure 10. A scores scatter plot where each year is masked as one class. The lables also indicated gender, age and nationality.

From the scores plot in Figure 10 some interferences could be made:

- The labels of nationality corresponds to the loading plot results, i.e. western countries are situated to the left (correlated to high course results) whereas men from middle east is situated in the upper right corner.
- A small trend between years could be seen where 2011 tends towards the left (and thus increased results). The spread in the scores plot is also smaller.

The analysis from the figures above is merely an example of what can be done when combining data of different sources using MVDA.

Instructional sensitivity

The change in CM upon TLA's is very clear, eg from Figure 6 (bar plots). One similar exercise can be made by making a PCA of pre-test CM data only and use the post test as a test set. A validation case (including instructional sensitivity) was made during 2012. This year the number of student was small and three sets of concept maps were made. Two CM as "usual" pre- and post course and a third using different "tasks". In this case the instruction was to make a CM by using the first CM as a base. Also the time available was one week instead of 15 min. The PCA model from Figure 5 was used as a calibration model and data from 2012 was used as a test set.

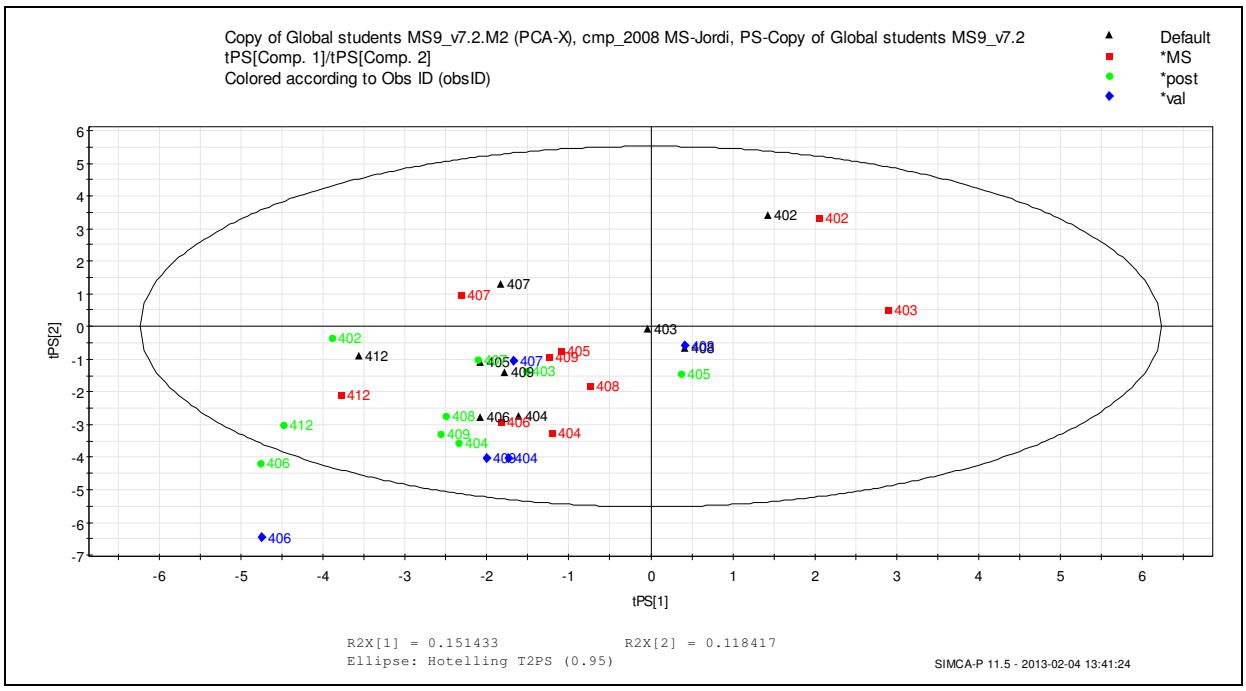


Figure 11 A scores scatter plot where pre-test is marked in red& black (two different evaluators). Post-test data is marked with green and the extra CM is marked in blue.

From the scores plot it can be seen a shift towards the down-left corner. This direction is also the direction for intercategory links in the loading plot (not shown). Hence the course is indeed making a change on the CM. The third CM (marked "val") is also displayed in Figure 11. Here it can be seen that the scores also tends towards the lower left corner. It can thus be concluded that increased time and/or more instructions enhances the CM performance.

Conclusions

A general comment on MVDA models is needed here: Note that all conclusions from the models are based on correlations! Correlation is NOT the same as causality, but it can still be useful for various purposes:

- Critical reflection is enhanced by looking and analysing the different plots
- Interesting correlations can be discovered and analysed further in order to confirm (or reject) the findings.
- The results are based on real data and not based on intuition and feeling, which sometimes can be hindering further actions towards improved teaching.

In this report the aspects of reliability and validity have been addressed and the conclusions are:

- The use of concepts maps to assess learning of sustainable development is both reliable and valid to the extent analyzed here.
- The use of MVDA as a tool to visualize the large amount of data and to analyze the correlation structure has been demonstrated useful.
- The content of this report could serve as a thorough assessment of the methods used (CM and TLA in conjunction with MVDA) in order to reflect on improved teaching in sustainable development in engineering education.

Appendix A: Extract from course PM

This extract is from the course 2011, but the content has been the same.

Purpose: This course provides the students with an understanding of the effects on sustainability of the actions of an engineer and with means to identify appropriate changes.
Learning outcomes: After successfully completing the course, the students should be able to:

- Describe the function and state of the natural systems in the world
- Explain how the human society, in particular chemical and chemical engineering industry, affects and depends on natural systems
- Describe international and regional work on environment and sustainable development on governmental and non-governmental level
- Identify appropriate tools and strategies for sustainable development in society, in particular industry
- Analyze sectors of society, in particular chemical and chemical engineering industry, and formulate appropriate strategies for sustainable development
- Describe the importance of including different stakeholders and perspectives in discussions on sustainable technology development

The course also includes experiences of using skills from earlier courses, e.g.

- Oral and written presentation
- Working in team
- Project work
- Critical review

Examination: The course is examined by collecting scores from different activities (max 100p). To pass the course you have to reach a minimum combined score of 50 points from different activities in the course (grade 3 from 50p, grade 4 from 70p and grade 5 from 85p). Points can be achieved by participating in different exercises during the course and by participating in a group project and in the written examination in the end of the course. Scoring system (maximum points for the different activities):

- Preknowledge test 6p *Movie 4p
- Museum 4p *Sustainability seminar 6p
- Project 20p *Stakeholder dialogue exercise 4p
- Multicriteria analysis exercise 4p *Bonus activity 2p
- Written examination 50p (to get any points at all from the examination you must get at least half the total examination score)

The written examination is a closed-book exam (you cannot bring the course material), but you can bring a non-technical dictionary if you wish. Calculators are not allowed. Answers are to be written in English.

Detailed information on some course activities:

Project on sustainable chemical engineering: problem-based, open-ended student team project including oral and written presentation. The main purpose of the project is to analyse different technical solutions in the context of sustainable development, looking at the present situation but also on the future potential in relation to anticipated changes. Each topic represents an important function in present society that we need to find a solution for. Each sub-topic presents an alternative (potentially sustainable) solution for fulfillment of the function. The sub-topics are first studied in smaller groups, and subsequently compared in the combined larger group. Participation in the sustainable chemical engineering project may give you 20 points adding to the final score for the course.

Topic	Sub-group
ELECTRICITY	Coal power with CCS

	4th generation nuclear power
VEHICLE FUELS	RME biodiesel
	Tall oil biodiesel
FOAM PACKAGING	Mycobond®
	PLA
METAL DEGREASING	Limonene
	EnSolv®
PROTEIN SOURCES	Artificial meat
	Soy protein

The topics should be studied from a global point of view, but local and specific cases can serve as examples. In the first phase, chemical and chemical engineering issues should be prioritized, but a broad view on sustainability is needed. A systemic view and a life cycle perspective on the topics are important. The following aspects might be important (but there may also be other important issues):

- Technology:
  - Processes
  - Limitations set by technical systems
  - Level and rate of process development; optimization opportunities
  - Working conditions and safety aspects
- Natural resources:
  - Raw materials – renewable or non-renewable, scarce or abundant
  - Accessibility and transports
  - Competition with other uses
  - Land area restrictions
  - Impacts from extraction/harvesting on local environment and population
- Waste and emissions
  - Types; from all life cycle phases
  - Options: preventive measures or end-of-pipe technologies
  - Opportunities for reuse, recycling, energy recovery
  - Trends
  - Societal strategies, policy instruments
  - Environmental impacts
- Products(s) or service(s) provided
  - Fulfillment of human needs; necessity
  -
  -
  - Function and quality
  - Additives and contaminants
  - Alternatives
- Society, global context and long-term perspective
  - Market and trends
  - Opportunities for developed and developing countries
  - Equity in terms of economic, social and environmental impacts
  - Role in a future sustainable society – cradle-to-cradle, necessary changes

Each sub-group consists of three to five students and should be as mixed as possible regarding the participants‟ countries of origin and background, to enhance a broad-minded and far-sighted discussion. […]

Preknowledge test: The preknowledge test (diagnostic) is a test on knowledge that students are expected to have at the start of the course, and is intended to motivate course participants to freshen up required preknowledge. The test is a multiple choice type test that can give you up to 6

points (depending on your performance) and will be given on Tuesday **Sept 6**. You can prepare for the test by reading the material in the textbook that is designated "preknowledge" in the reading instructions.

Al Gore Movie exercise: Paramount Classics and Participant Productions present a film directed by Davis Guggenheim, AN INCONVENIENT TRUTH. Featuring Al Gore, the film is produced by Laurie David, Lawrence Bender and Scott Z. Burns. Jeff Skoll and Davis Guggenheim are the executive producers and the co-producer is Leslie Chilcott.

> *"Humanity is sitting on a ticking time bomb. If the vast majority of the world's scientists are right, we have just ten years to avert a major catastrophe that could send our entire planet into a tail-spin of epic destruction involving extreme weather, floods, droughts, epidemics and killer heat waves beyond anything we have ever experienced. AN INCONVENIENT TRUTH offers a passionate and inspirational look at one man's fervent crusade to halt global warming's deadly progress in its tracks by exposing the myths and misconceptions that surround it. That man is former Vice President Al Gore, who, in the wake of defeat in the 2000 election, re-set the course of his life to focus on a last-ditch, all-out effort to help save the planet from irrevocable change."*

You need to sign up for participation in the Movie exercise no later than Wed **Sep 7** at 10am on the course home page. Before seeing the movie, you need to study what different people have said about the movie, positive and negative remarks about the movie's purpose, quality, potential bias or skew towards a certain position etc. You report this by handing in a filled in 'Voices on the Al Gore Movie' form when you come to see the movie. You will find this form on the course home page.

After watching the movie together, we will discuss different sustainability topics relating to the movie (e.g. the role of science and role models, responsibility, uncertainties, myths and visions, lobbyism, globalization etc). Watching the movie and participating in the discussion is optional and gives you 4p adding to the total score in the course.

Museum visit: In order to broaden the sustainability concept beyond traditional engineering issues, an opportunity to visit the World Culture Museum is given (http://www.varldskulturmuseet.se/smvk/jsp/polopoly.jsp?d=126&l=en_US). At the museum, students will be asked to reflect upon the implications for sustainable development of globalization and the contents in one of the exhibitions. A short presentation for students who have reflected on other topics is also required. This exercise gives you the opportunity to get a better understanding of the complexity and diversity of sustainable development, and a chance to share your reflections with a small group of students. The museum visit is optional, and can give you a maximum of 4p adding to your final score for the course. For students that wish to attend, we meet in the museum lobby at 1.30pm sharp on Thursday **Sept 22**.

Sustainability Seminar: At a seminar, groups of students meet to discuss different sustainability aspects and to comment on essays on the topics prepared by the other students. Eventually, the essays are finalized based on the reviews and discussions and handed in. In this way, the students penetrate different aspects of sustainability and get input in their thinking from other students. The focus in this exercise is primarily on social sustainability. Writing an essay in English is also practiced.

Sustainability is an extremely versatile concept. One objective of this course is for every student to have a basic understanding of the challenges of sustainability. The sustainability seminar may help you to faster grasp this complexity, especially in terms of social sustainability. Each student will prepare a written discussion on one topic within the concept of sustainability. This text will be peer-reviewed and discussed in seminar groups with five students in each group, where each student has worked with a different topic within the concept of sustainability. The written discussion must in the end be adjusted and finalized according to the feedback received during the seminar. Subjects have to be chosen no later than Friday **Sept 2** at 1 pm by signing up on lists at the course notice board. Nationalities must be divided between the groups as far as possible – the

different cultural backgrounds are an important part in the learning process for sustainability! Note that this might result in that you can not have your preferred choice of topic.
The sustainability seminar is optional, and can give you a maximum of 6p adding to your final score for the course.
Subjects for the sustainability seminar:

- Education for sustainable development
- Poverty and health and sustainable development
- Terrorism and sustainable development
- Gender equality and sustainable development
- Child labor and sustainable development

Within each subject, you can freely choose a specialization according to your own interests, but make sure that this exercise will broaden your view and not only deal with issues that are familiar to you. A global and broad approach is preferred over a narrow and local one.
Prepare a text with your own words, discussing the sustainability implications for your area, and on how your subject should be taken into account in order to reach a sustainable society. You must refer to at least one scientific text of your choice from available literature sources (you can use a text from the internet, but make sure it is of proper quality, e.g. a United Nations report, and not a mere Wikipedia definition!). Your text and arguments must be relevant and easy to read and your

argumentation possible to follow also for persons without deep knowledge on your topic. However, it must be on the level of master students! Since all participants in the seminar will be working with the concept of sustainability, this concept should not be described in detail.
The Header must contain: the text "Sustainability Seminar 2011", the group number of the seminar group, the sub-topic discussed, and the author name. The header must be written in Times 10, aligned right. The body of your text should be written in Times 10. The title heading should be Bold Times 14. If you use sub-headings, they should be Bold Times 10. The title heading should be placed three lines under the header information. Your discussion should have a length of preferably two (minimum one, maximum three) pages. References should be placed on an additional page. In the finalized text, this page must also contain the headings "Name of author", and "This text has been peer-reviewed by" followed by the names of the reviewing members in the seminar group. In scientific situations it is very important to make clear references! You should use numbered references. In the text, references shall occur in straight brackets [ ]. In the reference list, full bibliographic references shall be given for each reference number. The referred material shall be easy to find using the given information!
Send out copies for peer review and review the texts of the other members in the group. Your draft version should be finalized at the latest Friday **Sept 16.** You must send a copy of your text by e-mail to each of the other members in your seminar group. If you do not receive the texts from your group members on time, you must contact them yourself – they may have tried to send the text. Before the seminar, you must review the texts of the other members of your seminar group. Use the peer-review report cards that can be found on the course web-page (one for each of the texts from the other authors). Peer reviewing is the way that is normally used to guarantee high quality in scientific contexts. Since research deals with new information, there is no answer book! Instead, other researchers (peers) study the suggested text and comment on it critically and suggest improvements in argumentation, clarity, language, etc. In this exercise, you will train your ability to study new areas, identify weaknesses and errors in other students‟ argumentation, and suggest improvements. Read each of the texts of the other group members and give suggestions on improvements, directly in the text and on the peer-review report card according to the instructions. All major suggestions must be written on the report card. Write down at least one question for each text to discuss at the seminar.
On Tuesday **Sept 20**, 15^{15} –17^{00}, you will meet with your seminar group in rooms that will be announced later. Each of the group members will give a short (maximum 5 min) summary of his or her topic and lead a short (about 10 min) discussion on this topic in order to get feedback from the other students. When all the topics have been discussed, the peer-review report cards attached to the text (with comments) shall be handed over to the corresponding author.
You shall now finalize your text, taking into account the peer-review comments and the discussion at the seminar. You may need to correct minor errors and clarify things in your discussion that were misunderstood by your peers. At the latest Friday **Sept 23** at 1 pm, you must hand in a word file containing your report on the course home page and you must also hand in a paper copy of your finalized text together with all the review cards that the other students in your group gave you during the seminar. Make sure these are properly attached to the text. There is a box for hand-ins on the table below the course notice board at the department.
You may collect a maximum of 6 points adding to your final score in the course by fully participating in the Sustainability Seminar [hand out your draft text on time, participate in the

seminar and hand in your finalized text on time]. If you have not revised your text taking into account the peer-review comments, you will get a reduction of your score with 1 point for each relevant comment you have neglected. If the teachers find severe errors in contents, use of references and citations, clearness or arguments, that have not been pinpointed by any of the reviewers, all members in the group will have a reduction of their score with 1 point. If parts of your text are discovered to be copied word-by-word from another source without proper citation, your participation in the sustainability seminar will be cancelled and you will be reported to the university administration.

Stakeholder dialogue exercise: In this two-step exercise, students enhance their understanding of the importance of including different stakeholder perspectives in discussions on sustainability issues, e.g. in technology development, by analyzing important stakeholders and preparing for a dialogue in which they play a different role than they are used to and argue for the standpoint that they represent. Participation in the full two-step exercise will give you 4p adding to your final score for the course. You have to participate in the full exercise to get any points at all.

The topic for the stakeholder dialogue exercise is related to the topic of the projects. This work can therefore provide valuable input to the project work. If you are doing the project, you should participate in the stakeholder dialogue group that handles your topic. If you are not doing the project, you should participate in the stakeholder dialogue group that has the lowest number of participants. Groups will be formed at the Stakeholder analysis exercise that takes place on Thursday **Sept 8** at 3-5pm. The groups assign Stakeholder dialogue roles to each group member in the end of the stakeholder analysis exercise. Students prepare for the Stakeholder dialogue according to their assigned role, and they hand in a filled in Stakeholder dialogue form on the course home page no later than Friday **Sept 16**. In the afternoon on the same day, the Stakeholder dialogue will take place according to the following schedule: Group 1: 12.30-13.20; Group 2: 13.25-14.15; Group 3: 14.30-15.20; Group 4: 15.25-16.15; Group 5; 16.30-17.20.

In order to receive your bonus points, you must participate in both exercises, make preparations according to your role, and hand in the filled in form before the deadline.

Multicriteria analysis exercise: In this exercise, you will be performing a multicriteria analysis together with a group of other students. The purpose is to illustrate the importance of having the right information, of stakeholder involvement and transparency in decision-making for sustainability. In order to provide some depth to the exercise and also to provide some input to the work in the projects, the exercise will focus on the topics of the projects. This work can therefore provide valuable input to the project work. If you are participating in the projects, you will participate in the exercise group that is dealing with your topic. If you are not doing the project, you may participate in the exercise group that has the lowest number of participants. Groups will be formed at the exercise. Since you need to prepare for the exercise, you must sign up for participation on the course home page no later than **Sep 23** at 1pm.

Before the exercise, you must prepare by trying to find out information about important sustainability considerations in relation to your topic and hand in a participant sheet in which you have pointed out different aspects that you think, based on your reading about the topic, should be considered in choosing the most sustainable option. The participant sheet is found on the course home page. It should be filed in and brought to the exercise. This input will be used by the groups at the exercise. The multicriteria analysis is performed in groups on **Sep 27** at 3-5pm. Presentation and discussion is also made at the exercise.

Bonus activity of your choice: To make it possible for you to go into details on a topic of your choice, it will be possible for you to do a bonus activity. The activity must be accepted by the course leader beforehand. After doing the activity, you write a 1-page review and hand it in by e-mail to the course leaders. If the text fulfills the basic requirements on text quality and content, you will be given 2p adding to the final score for the course. Students are also encouraged to recommend interesting lectures or other activities (on environmental science and sustainable development issues) to the course leaders. Some relevant activities will be advertised on the course home page. Only one bonus activity will give you bonus points but you can of course do as many of them as you like. Some suggested activities:

Internet movies:

- The Next Big Thing, Round Table Discussions On Science, The Vega Science Trust: Endless Energy - Is it really time to stop using fossil fuels? What real alternatives are there? http://www.vega.org.uk/video/programme/8
- The Next Big Thing, Round Table Discussions On Science, The Vega Science Trust: GM Foods - Safe? Is GM a potential saviour of Mankind or a disaster waiting to happen? http://www.vega.org.uk/video/programme/2

Books:

- Ishmael – An adventure of the mind and spirit by Daniel Quinn
- Happiness – Lessons from a new science by Richard Layard
- Cradle to cradle – Remaking the way we make things by William McDonough and Michael Braungart
- The end of poverty – How we can make it happen in our lifetime by Jeffrey Sachs

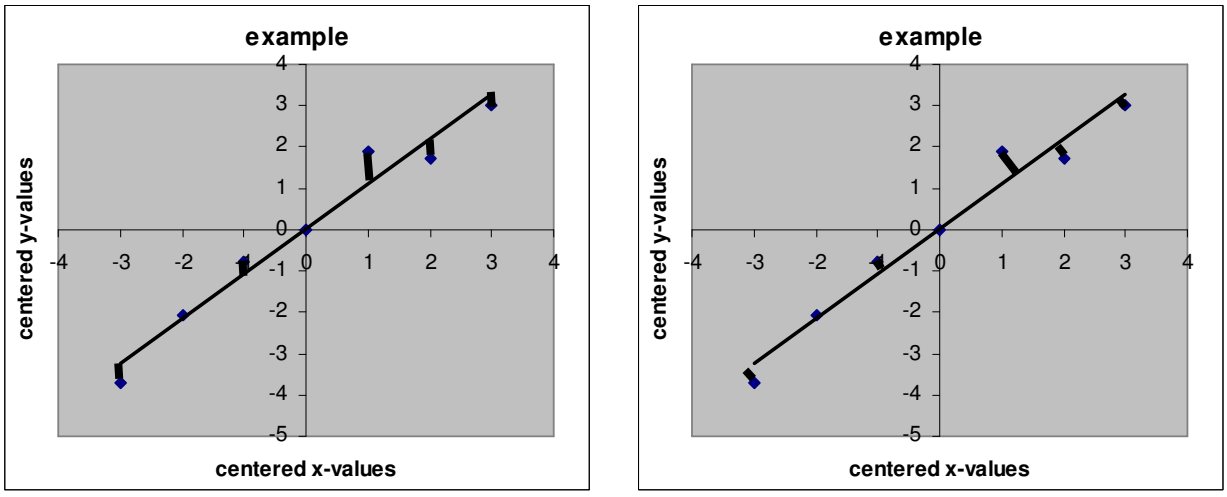
Appendix B

Multivariate analysis, Latent variable models

The multivariate methods described below (PCA, PCR, PLS) and similar methods are called "Latent variable" methods, because the nature of the low-dimensional hyper plane can be regarded as "latent variables". These methods are very useful in complex systems with many correlated variables and observations.

Principal Component Analysis, PCA

One of the assumptions for linear regression analysis is that the x-variables are exactly known. This may mostly be true enough (at least compared to the uncertainty in measuring the y-variable). Quite often though there is an interest in handling uncertainties in **X** as well. This was first analyzed by Pearson in 1901 [Pearson 1901]. The concept has been developed a lot since then and a model type that corresponds to Pearson's study is called Principal Component Analysis (PCA). The difference between a linear regression situation and a PCA model is that the residual to be minimized is not the "vertical" distance but the distance orthogonal to the line (the model). This is depicted in the figure below:



a) "normal" linear regression.: the residual is the vertical distance between the data and the model.

b) PCA: the residual is the distance between the data and the model perpendicular (orthogonal) to the model.

Figure 12. Differences in how the residual is defined between standard linear regression and a PCA model.

In the linear regression case, the model is y=kx and the only parameter is k.

In the PCA case, the model assumes errors in both x and y and the model therefore is

$$\begin{bmatrix} x & y \end{bmatrix} = [X] = TP' \qquad (2)$$

T is called the score matrix and consists of the values along the model plane (in the example above, the scores correspond to the values along the line shown in Figure 12b.) The score matrix is the new approximation of the original matrix X, but using fewer dimensions.

P is the "model" and consists of the linear combinations of the original variables that are used to project on to the model plane. Here we have two "parameters" p(1) and p(2) in the vector P. The loading matrix P is orthogonal and normalized to the size of one (orthonormal), *i.e.*;

$$P'P = I \qquad (3)$$

The PCA example above can be extended to many more variables and many more observations but works out the same way. We get a loading matrix P that will be used to project the original matrix X onto a low-dimensional plane, T.

The main advantages of PCA (and other LV methods) are

- It handles errors in x and y.
- It handles co-linear variables.
- It produces models that have components that are orthogonal.

For more details about PCA, see *e.g.* [Martens 1989; Eriksson 2001]

Principal Components Regression, PCR

After making a PCA on a set of x-data, we have the situation where we no longer have correlation between the variables. One way to proceed then is to make a multivariate linear regression but using the scores T instead of the matrix X. the model then becomes:

$$\hat{\mathbf{y}} = \mathbf{Tb} \qquad (4)$$

Where T is the score matrix from a PCA model of X, T=XP. This method will not be further discussed but serves as a "bridge" to the PLS method in the next section.

Partial Least Squares, PLS

The Partial Least Squares method (PLS) or "Projections to Latent Structures" as it sometimes is called is a regression method very similar to the standard multivariate linear regression and the PCR case described above. It uses two separate models for X and Y and then tries to find the correlation between these two models. The model now becomes:

$$\hat{\mathbf{y}}_i - \bar{\mathbf{y}} = (\mathbf{x}_i - \bar{\mathbf{x}})'\mathbf{b} \qquad (5)$$

Where **b** is the regression vector given by:

$$\mathbf{b} = \mathbf{W}(\mathbf{P}'\mathbf{W})^{-1}\mathbf{c} \qquad (6)$$

W, P, and **c** are loading vectors, *i.e.* linear combinations of the original x and y variables. A geometrical picture is given below:

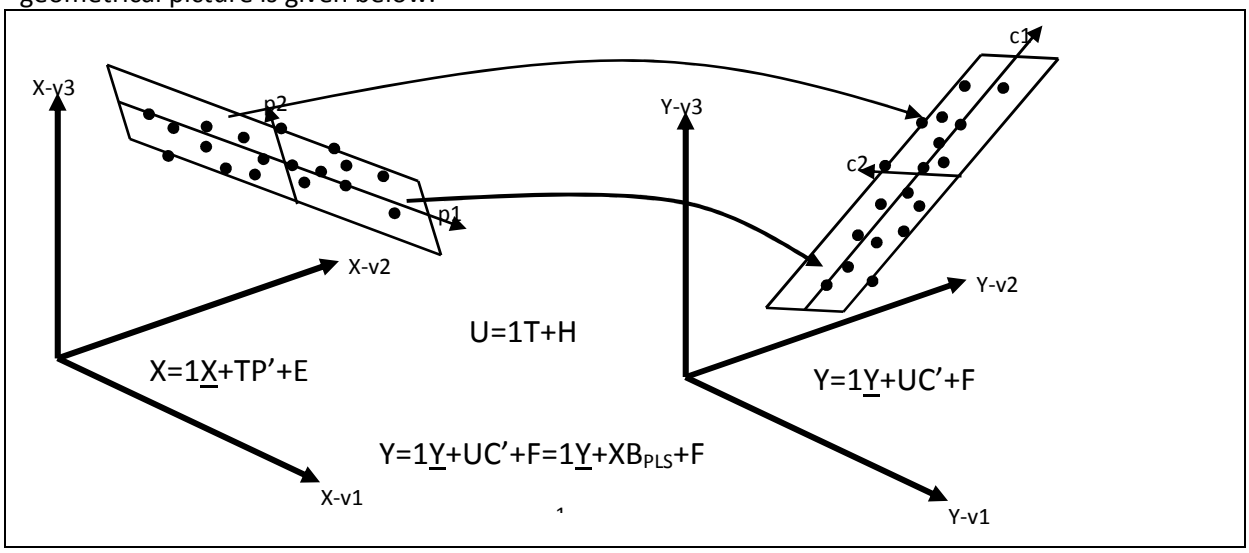


Figure 13. Geometrical description of the PLS model.

Since the **X** and **Y** matrices are "stripped" for every component, the corresponding PLS loadings (**W**) are associated with the corresponding, "stripped" matrix X. In order to make the interpretations more clear, $\mathbf{W}^*$ is used ($\mathbf{W}^*=\mathbf{W}(\mathbf{P'W})^{-1}$). W* can now be compared to the variables in X. Furthermore, this is used in paper II. For further details about calculations and algorithms see [Martens 1989].

Comments about latent variable models and projection methods

The phenomenon of projections is not new itself. Plato describes in his book "The republic" how projections of an unobservable reality is taken for the truth and anyone questioning it will be discredited [Lewi 2004]. The same story goes in the novel "Flatland" by E. A. Abbott [Abbott 1884]. In this chapter only the numerical methods of PCA, PCR and PLS are mentioned. There are a large

number of similar models and similar algorithms that all produce latent variable types of models. The concept of Latent Variables as a frame work for multivariable modelling have been well described by Burnham [Burnham 1996; Burnham 1999]. Another similar method is Factor Analysis which also is used during ALS.

References

Abbott, E. A. (1884). Flatland, A romance of many dimensions.

Biggs, J. and C. Tang (2007). Teaching for Quality Learning at University : What the Student Does, Open University Press

Brundtland, G. H. (1987). Report of the World Commission on Environment and Development: Our Common Future Oslo, United Nations General Assembly.

Burnham, A. J., J. F. MacGregor, et al. (1999). "Latent variable multivariate regression modeling." Chemometrics and Intelligent Laboratory Systems **48**(2): 167-180.

Burnham, A. J., R. Viveros, et al. (1996). "Frameworks for latent variable multivariate regression." Journal of Chemometrics **10**(1): 31-45.

Eriksson, L., E. Johansson, et al. (2001). Multi- and Megavariate Data analysis, Priciples and Applications. Umeå, Umetrics AB.

Lewi, P. J. (2004). "From data to knowledge to more data. Where is the portal to progress?" Chemometrics and Intelligent Laboratory Systems **73**(2): 167-168.

Martens, H. and T. Naes (1989). Multivariate Calibration. New York, Wiley.

Novak, J. D. and D. R. Gowin (1984). Learning how to learn. New York, Cambridge Press.

Pearson, K. (1901). "On Lines and Planes of Closest Fit to Systems of Points in Space." Philosophical Magazine **6**(2): 559-572.

Rice, D. C., J. M. Ryan, et al. (1998). "Using Concept maps to Assess Student Learning in the Science Classroom: Must different Methods Compete?" Journal of Research in Science Teaching **35**(10): 1103-1127.

Ruiz-Primo, M. A. and R. J. Shavelson (1996). "Problems and Issues in the use of Concept maps in Science Assessment." Journal of Research in Science Teaching **33**(6): 569-600.

Segalas, J., D. Ferrer-Balas, et al. (2008). "Conceptual maps: Measuring learning processes of engineering students concerning sustainable development." European Journal of Engineering Education **33**(3): 297-306.

Segalàs, J., D. Ferrer-Balas, et al. (2010). "What do engineering students learn in sustainability courses? The effect of the pedagogical approach." Journal of Cleaner Production **18**(3): 275-284.

Sjöblom, J. (2009). Parameter estimation in heterogeneous catalysis. Department of Chemical and Biological engineering. Gothenburg, Chalmers university of technology. **PhD:** 82.

Sjöblom, J., M. Svanström, et al. (2012). Assessing learning for sustainable development using concept maps and multivariate data analysis. Konferens om Undervisning och Lärande (KUL). Chalmers Lindholmen.

UNDP. (2011). "Human Development Index (HDI)." Retrieved 2011-10-26, from http://hdr.undp.org/en/statistics/hdi/.

UNESCO. (1995). "Education for Sustainable Development (ESD)." Retrieved 2013-01-24, from http://www.unesco.org/new/en/education/themes/leading-the-international-agenda/education-for-sustainable-development/.