



2nd International Symposium on PLS and Related Methods

Capri Palace, Island of Capri (Naples, Italy)

October 1st-3rd, 2001

organised by

Department of Mathematics and Statistics of the University of Naples "Federico II" – Italy

HEC School of Management – France

CISIA-CERESTA – France

with the scientific sponsorship of

IFCS (International Federation of Classification Societies)

SBI (Statistics in Business and Industry) Committee of ISI (International Statistical Institute)

IASC (International Association for Statistical Computing)

SIS (Italian Statistical Society)

OBJECTIVES OF THE SYMPOSIUM

The PLS methods (Partial Least Squares or Projection on Latent Structures) were introduced in 1983 by Svante Wold and his colleagues. Since then, many developments and several applications have been proposed, mainly in the field of chemometrics, cosmetics, petrol and agro-alimentary industry as well as in industries where production is run through continuous processes.

Nowadays, these methods of statistical data analysis are widely used, in particular when there are a lot of variables and few statistical units.

PLS'01 is an exceptional Symposium where, after the first edition of PLS'99 in Jouy-en-Josas (France), the most outstanding experts in PLS methods, from all over the world, will meet.

During the three days, aiming at both studying and formation of young researchers, the best experts on the topic will present the state-of-art on the PLS methods as well as the most recent developments and the most challenging advances demanded for the future. The PLS techniques will be compared to other methods and several applications to real data will be shown.

The Symposium will provide each participant with the necessary formation as complete as possible from both the theoretical and the application viewpoints. Furthermore, it will give the chance to personally meet the most specialised people in the field of interest as well as the real users of PLS methods in different fields.

Briefly speaking, the International Symposium PLS'01 pursues the following objectives:

- Making the state-of-art of PLS methods, PLS approach and any related methods
- Presenting the panorama of the most outstanding reference scientists connected to PLS
- Exhibiting all existing commercial and academic software for PLS
- Exploring numerous real applications of PLS in different industries

WHOM IS THE SYMPOSIUM PLS'01 ADDRESSED TO?

The Symposium is addressed to people who are interested in either the methodological aspects or the application potentialities of the PLS methods. Therefore, it is very suitable for actual and potential real users such as engineers and decision-makers working in the most different fields: chemical industries, pharmaceuticals, cosmetics, mining, paper industries, agro-alimentary companies. But, it also regards the people who are responsible for the top management and marketing researches.

These three days of intensive formation, will give the participants a complete overview, both didactic and pragmatic, for approaching decisional analysis with PLS techniques. The best solution for the following problems will be presented:

- Multivariate modelling with independent factors
- Perform the multivariate modelling of a process
- Improvement and optimisation of industrial production performances (agro-alimentary, cosmetics, chemical, petrol etc.)
- Quantification of relations between chemical structures and biological activities, etc.
- Evaluation of Customer Satisfaction and related issues

At the same time, the Symposium will represent a forum of discussion for the scientists working on the theoretical and methodological issues related to PLS.

The International Symposium PLS'01 will gather all those people who are already using PLS methods at any level as well as all potential users.

STEERING COMMITTEE:

Carlo Lauro (<i>Chairman</i>)	University of Naples "Federico II" - Italy
Michel Tenenhaus (<i>Co-Chairman</i>)	Groupe HEC - France
A. Morineau (<i>Co-Chairman</i>)	CISIA-CERESTA - France
Vincenzo Esposito (<i>Scientific Secretary</i>)	University of Naples "Federico II" - Italy
Bovas Abraham	University of Waterloo, Ontario - Canada
Sijmen de Jong	Unilever Research, Vlaardingen - The Netherlands
Nouna Kettaneh-Wold	UMETRI, Umeå - Sweden
Harald Martens	Norwegian U.Sci.Techn./Techn.U.of Denmark
D.L. Massart	ULB, Brussels - Belgium
Tormod Næs	Matforsk, As - Norway
Sung Hyun Park	Seoul National University – Republic of Korea
Svante Wold	University of Umeå, - Sweden

NATIONAL ORGANISING COMMITTEE:

Luigi D'Ambra (<i>Chairman</i>)	University of Naples "Federico II"
Carlo Lauro (<i>Co-Chairman</i>)	University of Naples "Federico II"
Vincenzo Esposito (<i>Scientific Secretary</i>)	University of Naples "Federico II"
Simona Balbi	University of Naples "Federico II"
Cristina Davino	University of Naples "Federico II"
Michele Gallo	University of Naples "Federico II"
Gianfranco Galmacci	University of Perugia
Marco Gherghi	University of Naples "Federico II"
Francesco Mola	University of Cagliari
Francesco Palumbo	University of Macerata
Antonio Perna	University of Naples "Federico II"
Corrado Provasi	University of Padova
Alfredo Rizzi	University "La Sapienza" of Rome
Germana Scepi	University of Naples "Federico II"
Roberta Siciliano	University of Naples "Federico II"
Rosanna Verde	Second University of Naples - Capua
Maurizio Vichi	University "La Sapienza" of Rome

STRUCTURE OF THE SYMPOSIUM:

The scientific programme will comprise:

- invited lectures
- contributed papers
- specialised sessions
- software exhibitions

The specialised sessions will be mostly devoted to specific areas of research within PLS or to real problems actually faced, for instance, in food industry, statistical process control procedures, and chemometrics where PLS and related methods are proved to help in getting a solution. These sessions shall comprise three speakers and may also be proposed and sponsored by private companies thus regarding their specific problems.

CALL FOR PAPERS - INSTRUCTIONS FOR AUTHORS

CONTRIBUTED PAPERS

Prospective authors are invited to submit an extended abstract (2 pages) of their paper to Dr. Vincenzo Esposito **by not later than May 8th, 2001.**

The contents of the abstract should clearly contain the following information about the paper:

- Relationship with literature
- Original contributions with main formulas
- Methodological, computational and/or application relevance
- Main references

The authors will be informed about the paper acceptance **by June 8th, 2001.**

The final paper (maximum 12 pages) has to reach Dr. Vincenzo Esposito **by July 16th, 2001.**

SPECIALISED SESSIONS

Prospective organisers of specialised sessions are invited to propose a basic topic and a maximum of three speakers' names with contributions titles and extended abstracts to Dr. Vincenzo Esposito **by not later than May 8th, 2001.**

The papers forming an accepted specialised session (final paper: maximum 12 pages) must reach Dr. Vincenzo Esposito **by not later than July 16th, 2001.**

PROCEEDINGS

Proceedings with all presented papers will be available on site and printed by CISIA-CERESTA. The insertion in these Proceedings will be ensured only for authors registered **by not later than July 16th, 2001.**

Moreover, after the Symposium, a number of selected papers will appear on special issues of the Italian Journal of Applied Statistics and the Journal of Chemometrics and Intelligent Laboratory Systems.

SOFTWARE EXHIBITION

Companies or individual researchers wishing to participate in the Software Exhibition should send their proposals to Dr. Vincenzo Esposito **by not later than May 30th, 2001.**

IMPORTANT DEADLINES FOR AUTHORS AND PARTICIPANTS

May 8th, 2001:	submission of abstracts and proposals of specialised sessions
May 30th, 2001:	submission of proposals for Software Exhibitions
June 8th, 2001:	notification of paper acceptance
July 16th, 2001:	submission of final version of the papers (contributed, invited, and specialised) and last useful day for registration
July 21st, 2001:	last useful day for Hotel Reservation
October 1-3, 2001:	PLS'01 SYMPOSIUM in Anacapri at Capri Palace Conference Centre.

REGISTRATION AND ACCOMODATION

VENUE

The Capri Palace is a Conference Centre annexed with a luxury hotel where painstaking attention to detail and the ancient art of good living go hand in hand with a family tradition and the highest standards of service. The Centre is often the site of major international conferences. It is situated in Anacapri, the most exclusive part of the Island of Capri in the Bay of Naples, at a point standing 300 metres above sea level. Well known in the world for the hospitality of its people, Anacapri offers a unique mixture of natural beauty and historic ruins and museums. From the celebrated Blue Grotto, with its marvellous nuances of blue, to the Church of San Michele which boasts a magnificent 18th century ceramic pavement portraying the Garden of Eden. Moreover, an amazing wealth of enchanting panoramic sites together with the modernity of the architecture blend in harmoniously with the tranquil and relaxing nature of the surrounding country-side. An ambience of refined elegance where everything seems to be designed with comfort and discretion in mind.

SOCIAL EVENTS

The social programme will include the following events:

- Welcoming reception on the Roof Garden of the Capri Palace;
- Excursions to the most attractive sites of the Island of Capri;
- Gala Dinner on Tuesday, October 2nd, 2001.

REGISTRATION FEES AND PROCEDURES

1.400.000 ItL. (people from private organisations)

600.000 ItL. (people from University or Public Institutions)

A 10% discount will be offered to IFCS-IASC-ISI-SIS members

300.000 ItL. (Phd Students upon presentation of official certification)

The registration fees include:

- welcoming reception;
- attendance at the symposium;
- proceedings and symposium materials;
- coffee-breaks;
- an excursion.

Authors of accepted papers need to be registered by not later than **July 16th, 2001** in order to have their papers published in the Conference Proceedings.

All payments must be made, free of all charges for the beneficiary, by following the methods of payment outlined on the registration form. Filled registration forms must reach Dr. Vincenzo Esposito by **not later than July 16th, 2001**.

Accompanying persons may choose to enjoy some services and social events of the Symposium by booking and paying for them directly on-the-spot at least one day ahead of the event.

ACCOMMODATION

Due to the high-season time and the limited number of rooms available on the Island of Capri within each price bracket, hotels will be allocated on a **first-come first-served basis**, i.e. a prompt return of the registration form including payment of the hotel deposit will probably secure the hotel class you require.

Hotels in Anacapri	Category	Type of Room	Daily Rates
Capri Palace (site of the Conference) Tel.: +39 081 837 3800 Fax: +39 081 837 3191 e-mail: reservations@capri-palace.com	*****L	Double (2 persons) Double - Single use	ItL. 400.000 ItL. 350.000
Caesar Augustus Tel.: +39 081 837 1421 Fax: +39 081 837 1444 e-mail: c.augustus@capri.it	****	Double (2 persons) Double - Single use	ItL. 400.000 ItL. 350.000
San Michele Tel.: +39 081 837 1427 Fax: +39 081 837 1442 e-mail: smichele@capri.it	***	Double (2 persons) Double - Single use	ItL. 260.000 ItL. 220.000
Casa Caprile Tel.: +39 081 837 3948 Fax: +39 081 837 3949 e-mail: info@casacaprile.com	***	Double (2 persons) Double - Single use	ItL. 250.000 ItL. 200.000
Biancamaria Tel.: +39 081 837 1000 Fax: +39 081 837 2060	**	Double (2 persons) Double - Single use	ItL. 250.000 ItL. 200.000
Hotel downtown Capri			
La Palma Tel.: +39 081 837 0133 Fax: +39 081 837 6370 e-mail: htl@capri.it	****	Double (2 persons) Double - Single use	ItL. 400.000 ItL. 350.000

Hotel rates include bed, breakfast, taxes and service. A one night deposit is necessary for the hotel reservation.

With respect to lunches, they will be served at the Capri Palace. The price for each lunch is 75.000 ItL. per person (including mineral water, wine and coffee) and will be paid on-the-spot. Please remember to **book your lunches** when filling in the Hotel Reservation Form

Filled hotel reservation forms must reach the chosen hotel by **not later than July 21st, 2001**:

Hotel deposits are non-refundable in case of cancellation or no-show.

FOR ADDITIONAL INFORMATION

Dr. Vincenzo ESPOSITO

Dipartimento di Matematica e Statistica
Università "Federico II" di Napoli
Facoltà di Economia
via Cintia, Monte Sant'Angelo
80126 Napoli - ITALY

tel.: +39 081 675112
fax: +39 081 675113
e-mail: binici@unina.it
Web-site: <http://www.dms.unina.it/PLS2001.html>

INVITED LECTURES

Bovas ABRAHAM

University of Waterloo, Ontario, Canada

Dimensionality Reduction Approach to Multivariate Prediction

We consider dimensionality reduction methods used for prediction, such as reduced rank regression (RRR), principal components regression (PCR), and partial least squares (PLS). We show how it is possible to obtain intermediate solutions by estimating simultaneously the latent variables for the predictors and for the responses. We obtain a continuum of solutions that goes from reduced rank regression to principal component regression via maximum likelihood and least squares estimation. Different solutions are compared using simulated and real data.

Bovas Abraham is the Director of the Institute for Improvement in Quality and Productivity and has been a consultant with the Institute since its inception. He has given seminars across North America, Europe, Australia, and the Far East. Bovas is also a professor in the Department of Statistics and Actuarial Science at the University of Waterloo. He received his Ph.D. from the University of Wisconsin, Madison, U.S.A.

His main areas of interest include Quality Improvement, and the management and implementation of statistical procedures such as Designed Experiments, SPC, and Time Series Analysis. He is co-author of the book "Statistical Methods for Forecasting" and the editor of a volume "Quality Improvement Through Statistical Methods".

Bovas is a Fellow of the American Society for Quality, a Fellow of the Royal Statistical Society, a Fellow of the American Statistical Association, and a member of the International Statistical Institute, International Environmetrics Society and the Statistical Society of Canada.

Garmt DIJKSTERHUIS, Harald MARTENS and Magni MARTENS

KVL, Royal Veterinary and Agricultural University, Department of Dairy and Food Science, Sensory Science Group, Rolighedsvej 30, DK 1958 Frederiksberg C., Denmark

GP-PLSR: An integrated technique for analysing sensory science data

GP-PLSR (Generalised Procrustes-Partial Least Squares Regression), as we coined it, is a combination of Generalised Procrustes matching of data matrices (GPA) and Partial Least Squares Regression (PLSR). The combination is such that the Procrustes rotations inherent in the GPA are guided by an external data set by means of incorporating a PLSR step in the iterative GPA process. The goal is to obtain a group average which is representative of:

1. The agreement among the K data sets in the \mathbf{X}
2. The structure in \mathbf{Y} as far as it accords with $\bar{\mathbf{X}}$

The above mentioned 1 can be referred to as a 'democratic process', and 2 to side-conditions that also must be taken into account. It can so happen that an individual data set \mathbf{X}_k is off, with respect to the other sets in \mathbf{X} but in accordance with the external information in \mathbf{Y} . In that case this \mathbf{X}_k should get the opportunity to make itself heard. This will be accomplished through a PLSR step in the algorithm (PLSR($\bar{\mathbf{X}}$, \mathbf{Y})). Analogously, when the majority of sets \mathbf{X}_k , $k=1, \dots, K$, are in strong agreement, they will together mainly shape $\bar{\mathbf{X}}$, and the effect of the PLSR($\bar{\mathbf{X}}$, \mathbf{Y}) should be outweighed by them.

Garmt Dijksterhuis is a sensory psychologist and methodologist. He studied psychology of perception at the University of Utrecht and wrote his Ph.D. dissertation at the department of Data Theory at the University of Leiden, in the Netherlands.

Garmt is one of the founders of the sensometric society, and he takes part in the scientific and organising committees of these and other conferences. He has written or co-authored over sixty publications. Garmt taught courses in sensory science and related topics in various countries and has been a visiting scientist at several universities and research institutes.

Currently he is employed as an associate professor at the Sensory Science Group of the Department of Dairy and Food Science at the Royal Veterinary and Agricultural University, Copenhagen, Denmark. His main research interests are perception and appreciation of food, and sensometrics.

Michele FORINA

DICTFA, University of Genova, Italy

Weighting predictors and objects in PLS

The flexibility of PLS algorithm can be used to assign suitable weights to predictors or to objects or to both predictors and objects. Iterative weighting by using regression coefficients (predictors), or residuals and leverages (objects), can produce refined regression models with better predictive ability and stability. By increasing the number of iteration a steady state is attained, where useless predictors are cancelled and a very economical model is obtained.

Michele Forina is Professor of Analytical Chemistry at the University of Genova, Italy. He was one of the first Italian chemometricians, with special interest in the application of Chemometrics to food problems. In 1983 he was the local organizer of the NATO ASI on Chemometrics in Cosenza, where he was appointed President of the International Chemometrics Society, the second after the co-founder B.Kowalski.

Inge HELLAND

University of Oslo, Norway

Rotational symmetry, model reduction and optimality of prediction from the PLS population model

We start by referring to some recent results indicating that the ordinary PLS1 algorithm itself cannot give optimal predictions in any ordinary statistical sense of the word. Then we show how a population model for this algorithm can be motivated and derived by letting the amount of data tend to infinity. The purpose of the talk is to show that this population model has some well-defined optimality properties connected to prediction. To define these properties we first remark that almost all known prediction methods behave as we expect under the group of rotations, so this is a natural symmetry group to look at in general. The theoretical best possible prediction methods under rotational symmetry can then be defined in principle. Considering the freedom of choice we are left with from this when doing model reduction, we are led in a natural way to the PLS population model. Prediction methods arising from this are complicated, but can be studied.

*Inge Helland was born in 1947. He got married in 1973 and has two children.
Education: Statistics; Cand. real, University of Bergen, 1973; Dr. philos, University of Oslo, 1980.
Lecturer/ Associate Professor, Agricultural University of Norway, 1978-83.
Professor, Agricultural University of Norway, 1983-1995.
Professor, University of Oslo, 1996-present.*

Sijmen de JONG

Unilever Research Laboratorium, Vlaardingen, Netherlands

(Direct) Orthogonal Signal Correction: comparison of old and new algorithms

Occasionally the first few PLS factors in multivariate calibration account for much X-variance and little y-variance. Orthogonal Signal Correction (OSC) has been proposed by Wold to remove such factors beforehand with the hope to obtain a more parsimonious and better performing calibration model with the pre-processed data. Several alternative versions of the OSC algorithm have been proposed. The

results obtained with these methods are of varying quality, so one cannot point to a 'winner'. We will discuss and compare the algorithms from a theoretical perspective. None of them simultaneously satisfies the two criteria sought for: exact y-orthogonality and maximum X-variance accounted for. We present a new algorithm, Direct Orthogonal Signal Correction (DOSC), a non-iterative algorithm that does find the solution satisfying the two criteria. This theoretical nicety does not translate into consistent better results, however. At the time of writing it appears that giving up the strict orthogonality requirement may lead to better predictions. Possible reasons for this behaviour will be discussed.

Sijmen de Jong joined Unilever Research as a chemist. In the early 80s he diverted his attention to statistics and chemometrics with a special interest in multivariate analysis (pattern recognition, multivariate calibration, PLS). Application areas span a wide range, covering a.o. analytical chemistry, microbiology, sensory analysis, QSAR, and, lately, high throughput screening. He has co-authored over 40 papers and one book on chemometrics. Currently he leads the Data Science skillbase.

Nouna KETTANEH-WOLD¹ and Svante WOLD²

¹Umetrics Inc., Kinnelon, NJ, USA,

²Research Group for Chemometrics, Umeå University, S-901 87 Umeå, Sweden

PCA and PLS with very large data sets

Chemometrics was started around 30 years ago to cope with the rapidly increasing volumes of data produced in chemical laboratories. A multivariate approach based on projections – PCA and PLS – was developed that adequately solved many of the problems at hand. However, with the further increase in the size of our data sets seen today in all fields of science and technology, we start to see inadequacies in our multivariate methods, both in their efficiency and interpretability.

Starting from a few examples of complicated problems seen in RD&P (research, development, and production), possible extensions and generalizations of the existing multivariate projection methods – PCA and PLS – will be discussed. Criteria such as scalability of methods to increasing size of problems and data, increasing sophistication in the handling of noise and non-linearities, interpretability of results, and relative simplicity of use, will be held as important.

The discussion will be made from a perspective of the evolution of scientific methodology as driven by new technology, e.g., computers and graphical displays, and the need to answer some always reoccurring and basic questions, and

constrained by the limitations of the human brain, i.e., our ability to understand and interpret scientific and data analytic results.

Nouna Kettaneh-Wold is owner and president of MDS Inc, and Umetrics Inc, Kinnelon, NJ, USA, and major owner of Umetrics, Umetrics AB, Umeå, Sweden, companies that consult with industry on experimental design and multivariate modelling, and develop appropriate software. She is main designer of the MODDE and SIMCA software of UMETRI. She has published several papers in the chemometrics literature on the design of experiments, mixture modelling, and process modelling.

Svante Wold is professor of chemometrics at the University of Umeå. He has published 305 papers in chemometrics and applications (multivariate modelling, design, and analysis in chemistry, structure-property and structure-activity relationship, multivariate dynamic modelling of chemical processes). He has developed several new algorithms as SIMCA, PLS regressions; hierarchical PCA and PLS, non-linear PLS and multivariate process modelling and monitoring. Professor Wold received the Arrhenius medal, the Torbern Bergman medal and the Herman Wold medal from the Swedish Chemical Society, the Nordbankens Scientific Price and the Eastern Analytical Symposium's NIR award. He is a member of the Swedish Academy of Engineering Sciences. He has been the chairman of the 1994 Gordon Conference on Statistics in Chemistry and Chemical Engineering.

Øyvind LANGSRUD and Tormod NÆS

Matforsk, As - Norway

Optimised score plot by principal components of predictions

A common problem in statistics/chemometrics is to relate two data matrices (X and Y) to each other, with the purpose of either prediction or interpretation. Usually one is interested in understanding which directions in Y-space that can be predicted by which directions in X-space. Several methods exist for this, for instance PLS regression and canonical correlation. The present paper presents a new plot for visualising the relationship between X and Y. The plot can be used for any regression method.

Professor Tormod Næs is a principal research scientist at the Norwegian Food Research Institute and professor in statistics at the University of Oslo. He also works part time as a private consultant in statistics. In 1984 he obtained a Ph.D. in statistics from the University of Oslo. He is on the editorial board of Journal of Chemometrics, Food Quality and Preference and Journal of Near-infrared Spectroscopy. He has co-authored two books and co-edited two, in the field of multivariate analysis and analysis of variance. He has published papers in statistics, spectroscopy, food science and sensory analysis.

Øyvind Langsrud is a research scientist at MATFORSK (Norwegian Food Research Institute). He has a Ph.D. in statistics and has published papers in Journal of Chemometrics and Computational Statistics and Data Analysis. A main research interest is how to handle multiple (collinear) responses in designed experiments (See <http://www.matforsk.no/ola/index.html>).

Harald MARTENS¹ and Magni MARTENS²

¹ Professor of chemometrics at section for physical chemistry, Norwegian University of Science and Technology & at dept. of biotechnology, Technical University of Denmark

² Professor of sensory science at section for sensory science, Royal Veterinary and Agricultural University, Denmark

Multivariate Analysis of Quality

The present talk presents our approach to PLSR-based data analysis (Martens & Martens 2000), and explains the statistical and didactic principles involved. It gives an overview of how ONE SINGLE METHOD (multivariate soft modelling by cross-validated weighted, rank-stabilized Partial Least Squares Regression) may be used for MANY DIFFERENT data analysis purposes, ranging from purely explorative factor analytical inspection of a data table, via interdisciplinary two-block regression, multivariate calibration and prediction and classification / discrimination, to the analysis of effects in designed experiments.

Thereby, the researcher can do safe, understandable and effective analysis of his or her own data, both inductively and deductively, without having to master several methods and several software systems, and without alienating mathematical/ statistical theory. The new extended jack-knifing method visualises the uncertainty of the results directly. The application examples, concerning the definition and assessment of quality, will be shown.

Reference: Martens, H. and Martens, M. (2000), Multivariate Analysis of Quality. An Introduction. 420 pp; J. Wiley & Sons Ltd, Chichester UK

Harald Martens, Dr. techn., Professor of chemometrics at section for physical chemistry, Norwegian University of Science and Technology & at dept. of biotechnology, Technical University of Denmark. Born 1946, M.Sc. industrial biochemistry NTNU Trondheim 1971, Dr.techn. chemometrics NTNU Trondheim 1985. More than 100 publications & patents on data analysis in chemistry, signal processing and sensory science. Recipient of the first Chemometrics Prize from the Norwegian Chemical Society/ Chemometrics Work Group 1993, for 15 years of chemometrics work. Honorary member of Chemometrics Work Group of the Norwegian Chemical Society from 1994. Member of the Norwegian Academy of Technical Sciences from 1998. Herman Wold gold medal 1999, Chemometrics section of Swedish Chemical Society. Chemometrics Award 1999, Eastern Analytical Symposium (USA), sponsored by Galactic Industries, Inc. Honorary issue of J. Chemometrics (in press, 2000).

D.L. MASSART

ChemoAC, Farmaceutisch Instituut, Vrije Universiteit Brussel, Laarbeeklaan 103, B 1090 Brussel, Belgium

Diagnostics for application of PLS in multivariate calibration

Optimal practical application of PLS for multivariate calibration requires that diagnostics should be available for detecting problems that can affect the quality of the result. Such diagnostics are needed for instance for the following problems:

- is the calibration set representative for the problem studied?
- how to detect the presence of inhomogeneities in the calibration set (clustering, outliers)
- how to detect the presence of outliers and inliers in prediction
- how to detect non linearities?

Diagnostics will be described and, where relevant, remedies will be discussed.

Professor dr D.L. Massart is head of the department of analytical chemistry at the Vrije Universiteit Brussel. His research concerns the practical application of chemometrics, including PLS, but also many other chemometric topics. He is author or co-author of several books on Chemometrics, the latest one being the Handbook of Chemometrics and Qualimetrics published by Elsevier. He is editor in chief of Chemometrics and Intelligent Laboratory Systems.

Aloke PHATAK

CSIRO Mathematical and Information Sciences, Private Bag 5, Wembley, WA 6913, Australia

PLS, Conjugate Gradients, and Lanczos Methods

The connection between PLS and Lanczos methods for approximating the extreme eigenvalues of a symmetric matrix has long been known. Less well known, however, is that PLS is in fact identical to a common implementation of the conjugate gradient algorithm for solving the normal equations. In this talk, I will outline the links between PLS, Lanczos, and conjugate gradients, and show that not only do they shed more light on PLS, they also provide a useful framework for classifying the many algorithms that exist for univariate PLS. More importantly, however, these connections provide us with simpler proofs of two important properties of PLS: that it shrinks and that it 'fits better' than PCR.

Dr Aloke Phatak is a senior research scientist with CSIRO Mathematical and Information Sciences. Although he began his career working on solid rocket propellants, he is now an applied statistician who consults in the areas of industrial statistics and chemometrics. He also carries out research in applied statistics and teaches industrial short courses.