JRC 85582

ISSN 1977-5296

ESARDA

Bulletin

NUMBER 18
JULY 1990

13th ESARDA Symposium

14 - 16 May 1991, Avignon, France

The 13th ESARDA Symposium will be held in the historical Palace of the Popes in Avignon. This is a town in the south east of France near the confluence of the rivers Rhone and Durance. It belonged to the popes up to 1791, when it was definitively incorporated into France.

Avignon is a historical town which hosted the popes in the middle ages. At the beginning of the 14th century, insecurity was souverain in Italy, particularly in Rome. This forced the pontiffs to seek refuge in a more hospitable land.

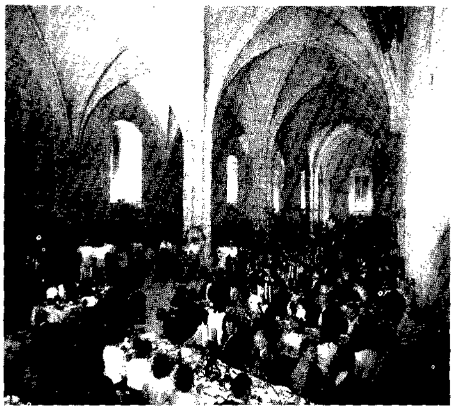
Therefore, Clement V who was souverain pontiff from 1305 to 1314, decided to settle down in Avignon, a town near to a countryside owned by the Holy See since 1274. He lived there in the Dominican Monastery. His successor, pope John XXII (1316-1334), who formerly was bishop of Avignon, settled in the episcopal palace and started the work of restoring and enlarging. This palace became big and powerful but was considered insufficient by Benedict XII (1334-1342), who demolished and reconstructed all the parts which are now called the Old Palace.

Clement VI (1342-1352) added the big buildings at the south and west sides which are now known as New Palace. (Note that the region is now proud of its excellent wine which is, in fact, called with the name of the new palace of the popes, namely "Chateauneuf du Pape"). Innocent VI (1352-1362) completed and reinforced the work of Clement VI. Then, Urban V (1362-1370) built the so-called "Roma" in the east gardens. Finally Gregory IX (1370-1378), receiving the supplications of St. Catherine of Siena, decided to bring the Holy See back to Rome in 1376.

After the departure of the popes, the palace has been the residence of the Vatican legates up to the French revolution. The palace was then used as military barracks up to 1906. Now, the palace is restored and reconstruction work is still going on.

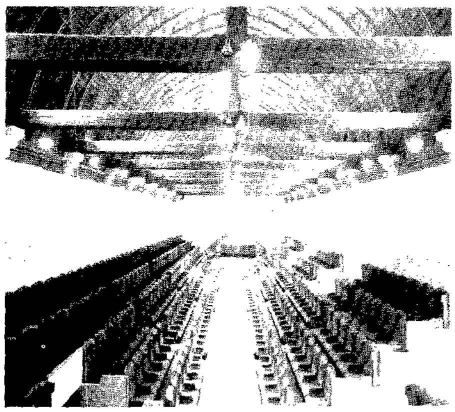
A visit to the palace is remarkably interesting. The town of Avignon in the heart of the Provence, dominates from its towers the valley of the Rhone. The surroundings of Avignon are very charming. The ESARDA Symposium could not choose a more attractive place. We will have there centuries of history escorting our work when sitting in the Room of the Conclave.

L. Stanchi



Reception room

General view of the Palace of the Popes



Room of the Conclave

View from the Rhone

Computer Applications in Optical Surveillance

F. Sorel

CEC, JRC-Ispra
Institute for Systems Engineering and Informatics

Introduction

Optical surveillance will play in the future a more important role in safeguards especially in long term storage and inaccessible areas of nuclear material. The most used surveillance device is the TV camera which delivers a video picture, but also other devices such as laser could be included in the category of optical surveillance. The TV picture is easy to review by visual inspection, but it is difficult to characterize in a formalized way the performance for safeguards objectives. Furthermore, the increasing amount of surveillance data (presently several million images per year) will require a higher automation process by computer either for on site data reduction or for assistance during review phase. Similar problems arise for the verification of seals presenting their identity under the form of pictures.

This article presents some computer applications in video data reduction, integration of multisensor system, automatic verification of seal pictures and laser surveillance.

Video Data Reduction or Review Aid

From the technical point of view on-line data reduction or review aid are similar : in the case of data reduction the computer processes directly the TV signal before recording on videotape; in the case of a review aid the computer receives the signal from the videorecorder. The method of image processing will be the same, but the image capture from a recorder presents more problems due to the signal jitter especially from time lapse recorders.

The JRC developed several image processing systems for the detection of scene changes in sequences of video pictures. These systems process a standard video signal; they have been designed for general purpose applications and are not depending on the recording system or the type of videorecorder. The first generation proved the validity of a detection method based on the cross-correlation of predefined areas of interest /1,2/. The system detected scene changes in selected picture windows compared to a reference image and stored these pictures separately. It was, however, not fast enough for processing pictures in video real-time. A second generation based on the polyline method has been developed for real-time applications /3,4/.

The image processing equipment comprises a personal computer with a plug-in image processing board for digitizing and storing the video signal in a local image memory. The memory capacity of one board amounts to four images with a resolution of 512x512 pixels each. The computer can then process the image on pixel basis according to the developed application software.

The method for detecting scene changes is based on a correlation technique of polyline profiles between the grabbed and a reference image. If the correlation factors of a specified set of polylines drop below their thresholds, an alarm is triggered and the reference image is updated. This technique is rather insensitive to homogeneous variations in illumination and to noise of electronic equipment. The set of polylines, from one to sixteen, is created by the system operator using an interactive graphic editor. This set is stored on disk and recalled for successive routine use.

In case of alarm the computer inserts in the video image the concerned polylines in red color, updates the alarm table and provides an external action, e.g. record the video picture sequence on tape. The possibility of defining a detection mask on several polylines reduces the false alarm rate.

A special software module has been developed to correct for vertical or horizontal image instabilities caused by film cameras or recorders /5/. Recently some new features have been added concerning alternative detection methods to the correlation and the option of programmable timed mask sequences. This option makes the detection process more selective because it distinguishes between objects moving in different directions or at different speeds.

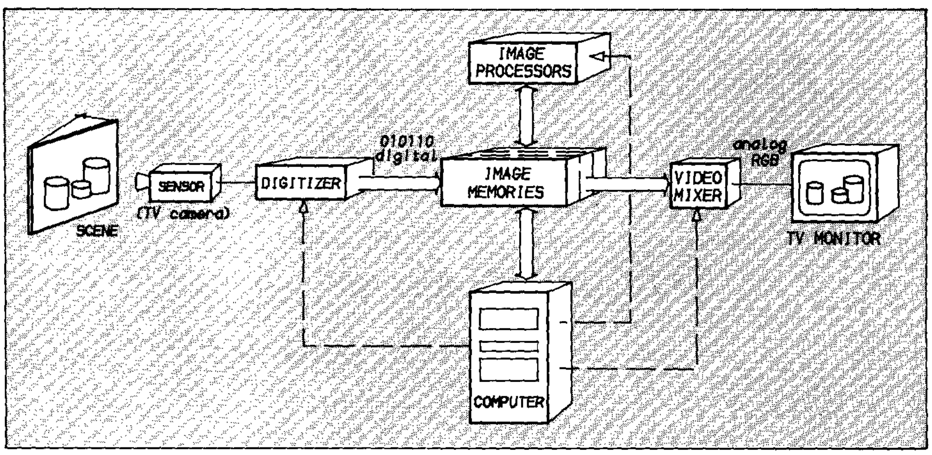


FIG. 1. *General configuration of an image processing system*

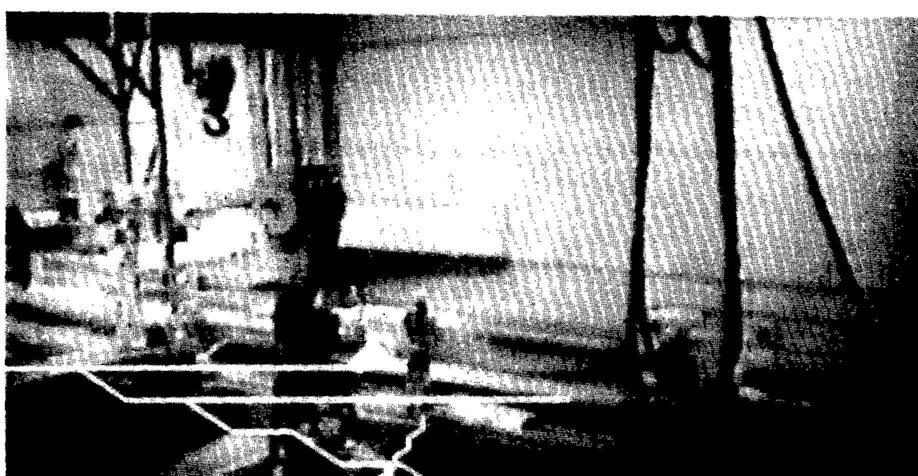


FIG. 2. *Display of a video picture with insertion of polylines*

Integration in Video Surveillance

The complexity in recording surveillance data is increasing with multicamera systems and the requirement of monitoring different sensor signals for assuring better reliability and data evaluation. JRC developed in collaboration with Euratom Safeguards Directorate a modular integrated system in which all functions are computer controlled. This system, called CAVIS (Computer Aided Video Surveillance), assists inspectors with system operation and trained staff at headquarter with a more detailed data analysis or maintenance plan /6,7,8/. The main features are failure detection in system components, alarm table storage, remote control by computer of connected units and efficient data evaluation.

The development effort in JRC-Ispra was dedicated to the control unit having in mind to be as much as possible independent of external device hardware. The first prototype accepts signals from analog and digital sensors and includes two processors for maintaining the operational control of video recording during the data evaluation phase.

The software module for self-surveillance monitors continuously all system parameters like supply voltages, temperatures of electronic equipment and read-after-write check on video recorders. In case of abnormal operation the corresponding conditions are recorded in the alarm tables.

Since the timing of the video multiplexer and videorecorders are computer controlled, it is possible to select by keyboard the required system configuration for different applications, e.g. cycle time for acquisition, number of recorded frames per camera, etc.

The data evaluation software provides the necessary tools for the retrieval, analysis and visualization of recorded data, e.g. alarms in a selected time period or error history for reliability figure on certain system components. The software dialog guides the inspector through all steps of procedure using only the computer keyboard, even for video review. It prepares the final report on printer and copies the stored data on floppy disk for an eventual more detailed analysis on the headquarter computer.

The open system design allows a future integration of other C/S devices like motion detectors or electronic seals.

Automatic Verification of Seal Pictures

Some seals, e.g. E-type or Cobra, present their identity under the form of pictures and the verification is based on image comparison. The computer technology required for the automation of this process is the same than that used in video surveillance and therefore this application has been included in the article. JRC-Ispra developed several computer vision systems for the automation of the verification technique for E-metal seals. The seal picture includes random distributed solder taps and scratches which identify the seal.

In the past, the seal pictures have been stored on slides before installation and the verification was done by a visual comparison of the observed seal picture with the reference image on slide. Both pictures were displayed on a TV monitor for an easier comparison. This procedure was man-power intensive considering the high number of seals (about 18,000 seals per year at Luxembourg) to be identified and verified.

For achieving a valuable automation two main problems have to be solved: the first of an efficient image archive with high capacity and fast retrieval, the second one concerning the automatic recognition of the seal pattern by computer.

The chosen solution for the picture archive consisted of a digital optical disk with removable cartridges which can be written only once and easily read by computer during the verification phase /9/. The computer configuration is similar to that described in the chapter of video data reduction. The seal picture taken by a TV camera is digitized in the image memory in which the computer adds the seal number and is then transferred to the optical disk. The capacity of one cartridge amounts to 1 gigabyte per side which corresponds to 12,800 pictures in the used data format /10/.

Several programs have been implemented for guiding the system operator through the different steps of the verification procedure /11/. Two complete systems have been realized and tested at Euratom Safeguards Directorate in Luxembourg for the archival and retrieval of the E-seal pictures.

The next important step concerned the automatic recognition of the seal pattern by computer. The investigated method is based on the correlation in zones of the observed and the reference image. The whole picture is divided into 64 correlation zones which gives a good balance between detection capability and stability over long periods. A configuration with many

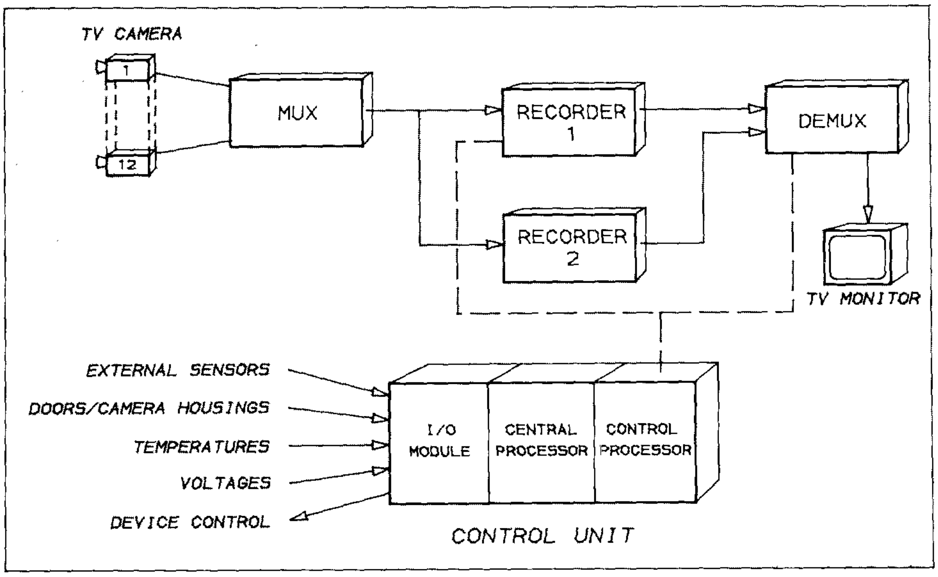


FIG. 3. *CAVIS 1 configuration*

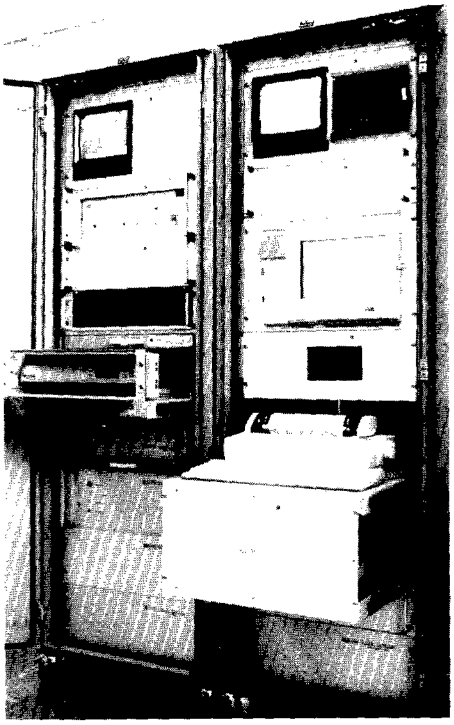


FIG. 4. *Front view of CAVIS 1*

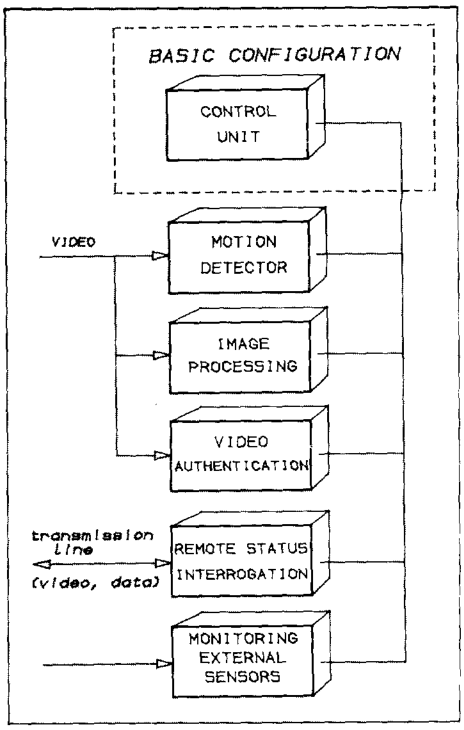


FIG. 5. *Possible integration in CAVIS*

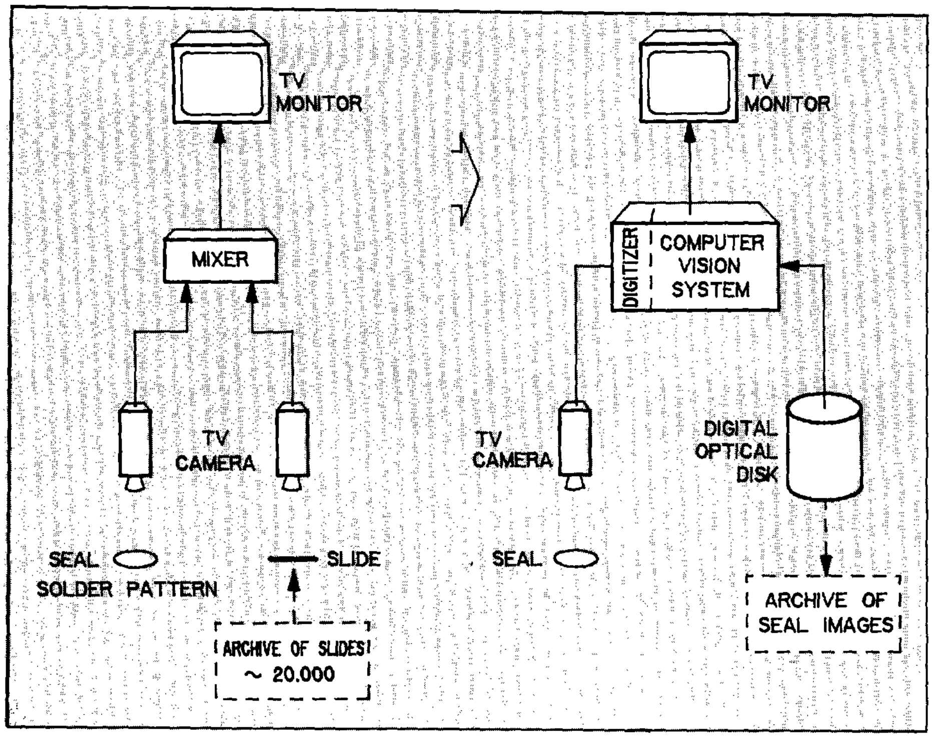


FIG. 6. *Automation of seals identification*

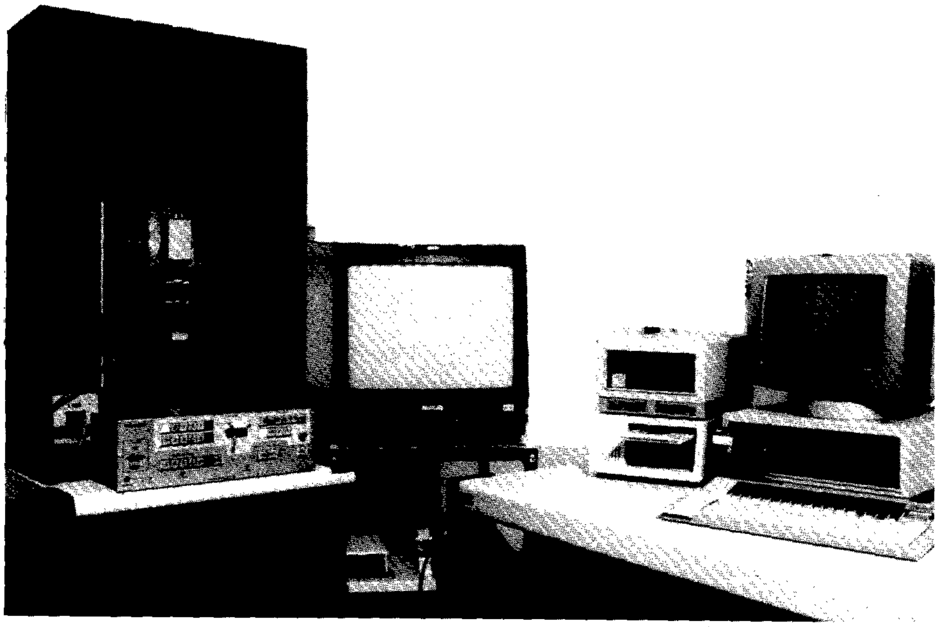


FIG. 7. *Computer vision system for seal verification*

small zones allows detecting a smaller discrepancy between images but is also more sensible to instability in illumination and in electronic equipment, e.g. TV camera. An image filtering is performed in each zone for increasing stability and then the correlation factors are calculated. If a factors drops below the selected threshold the system evidences this zone on the TV monitor.

Two important aspects in the system realization are the seal illumination and the seal positioning under the camera. The adopted solution consists in a horizontal light beam which is reflected by a glass plane at 45 degrees. The design of the seal positioning mechanism assures automatically the correct x and y position whereas the seal is rotated by hand in the correct position indicated by the computer on the TV monitor. Later on the rotation could be achieved by a computer controlled stepping motor.

A prototype system is operating in the laboratory and extensive measurements have shown good results for detection capability and stability /12,13/. The realization of a more engineered system is in progress.

Laser Surveillance

In this application the optical sensor is a laser device and the computer performs the analysis of the reflected light beam. JRC-Ispra developed, in collaboration with IAEA, the engineered prototype of the laser system LASSY which was primarily designed for surveillance at spent fuel storage pools. A beam of laser light scans a horizontal plane; when an assembly penetrates this place of light the changes of reflectivity and distance at the corresponding angle position of the laser eye are detected. The computer evaluates these changes and in case of alarm stores all concerned data on disk for a later interpretation.

The determination of the distance between the laser eye and the penetrating object can be achieved with one laser beam only since in the last version the phase shift measurement between emitted and received signal has been included. The addition of a second eye assures a higher precision in the determination of the disturbance position and a useful redundancy for a future underwater operation. The laser eye scans horizontally over 90 degrees and the computer takes up to 2048 measurement points in one scan. One measurement point contains the angle position of the eye given by a precise shaft encoder, the amplitude and phase value of the reflected signal.

The modular design of the present prototype allows for easy exchange of various optical and software components; this enables the adaptation to the different working media - air or water. The emitter and detector optics are coaxially aligned to assure a good tamper resistance. Using as light source a laser diode of 2-4 mW output power the system operates over 60 meters in air and with a light emitting diode (LED) the range is about 30 meters. The use of a very weak light source for safety reason required the development of dedicated electronic circuits performing the preamplification and synchronous demodulation of the received signal.

In case of alarm the computer prints out a reduced information set for quick look but stores also a detailed data set on disk. The printed data include date, time, the class of alarm and some indication on the position and size of the penetrating object. Using one eye, scanning horizontally, the computer can only indicate the projection of the object perpendicularly to the beam. The detailed data set include all measurement points of the whole scan.

The computer compares the signal amplitude and distance values to a continuously updated background which corresponds to the undisturbed reflected light from the containment walls. Detected deviations are analysed according to a hierarchical classification in order to reduce the false alarm rate. The evaluation parameters are continuously updated in an iterative process and adapted to the environment and scan conditions.

In alarm situations the computer provides an external signal which can be used to trigger a TV camera; this allows the LASSY integration with other C/S devices and long term tests monitoring. Combining a laser system with TV surveillance increases considerably the overall assurance since the two devices are based on different methods and technologies. The laser system compared to the TV camera gives more quantitative results on scene changes concerning the object position and size. A complete one-eye laser system has been realized in laboratory and field tests are in progress.

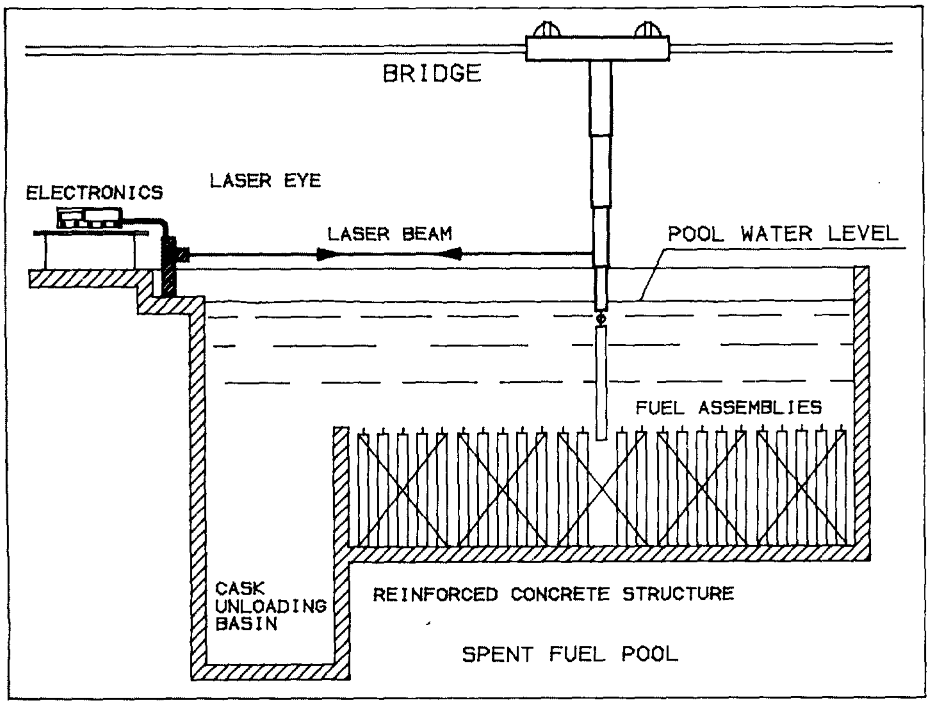


FIG. 8. *Application of laser surveillance*

FIG. 9. *Picture of laser eye*

Conclusions

This paper has presented some computer applications developed by JRC-Ispra in the field of optical surveillance concerning video data reduction or reviewed aid, integration in video surveillance, automatic verification of seal pictures and laser surveillance. The new technologies offer recently interesting solutions to overcome some present limitations in surveillance performance, e.g. digital optical disk for image storage, computer vision system for video picture processing, laser device for distance measurement. Nowadays it would be also feasible to apply computer assisted teleoperation to safeguards operations in areas of difficult access to the nuclear material.

The future role of optical surveillance in safeguards will depend on: 1) the possibility to characterize in a more quantitative way its performance; 2) the possibility to introduce a higher automation in the evalaution phase, and 3) its integration in systems for combining information from different safeguards devices. These developments demand for a more intensive use of dedicated computer systems integrating the new technologies which are presently available on the market.

References

/1/ BETTENDROFFER, E. — Surveillance du stockage de matières fissiles par système de vision — Technical Note No. I.05.03.86.02 (1986)

/2/ BETTENDROFFER, E. — Image Processing System for Videotape Review — Report EUR 11624.EN (1988)

/3/ MOL, M.J. — A Real-Time Scene Change Detector for the Review of Video Images. In: Proc. 29th INMM Annual Meeting (1988)

/4/ MOL, M.J. — Program Review for Real-Time Scene Change Detection in Video Images on a PC Based Vision System — Report EUR 11677.EN (1988)

/5/ MOL, M.J. — Extensions and Changes to Program Review — Report EUR 12055.EN (1989)

/6/ BETTENDROFFER, E., MORANDI, C., SOREL, F., TERMANINI, A. — Computer Aided Video Surveillance, CAVIS. In: Proc. 11th ESARDA Annual Symposium, May 1989, Luxembourg

/7/ BETTENDROFFER, E., MORANDI, C., TERMANINI, A. — CAVIS Technical Reference Manual — Technical Note No. I.89.80 (1989)

/8/ BETTENDROFFER, E., MORANDI, C., TERMANINI, A. — CAVIS User's Manual — Technical Note No. I.89.81 (1989)

9/ AMIC, S., DETOURBET, P., HAAS, R., SOREL, F. — Automation of the Verification Technique for Metal Seals. In: Proc. 9th ESARDA Annual Symposium, May 1987, London (UK)

/10/ AMIC, S. — Système d'archivage d'images sur disque optique — Note Technique No. I.88.16 (1988)

/11/ AMIC, S. — Logiciel pour archiver et vérifier les images de scellés à métal — Note Technique No. I.88.141 (1988)

/12/ COMBET, M., SOREL, F. — Reconnaissance automatique ds sceaux du type E — Report EUR 12739.FR (1990)

/13/ COMBET, M., CAMPOS, G. — Etude expérimentale de la reconnaissance des sceaux E — Technical Note I.90.29 (1990)

/14/ THOMSEN, K., MORANDI, C., HAMMER, J., SOREL, F. — Laser Surveillance System, the Engineered Prototype LASSY. In: Proc. 11th ESARDA Annual Symposium, May 1989, Luxembourg

Spent Nuclear Fuel Reprocessing : Chemometrical Treatment of Input Analytical Data

S. Schoof
Kernforschungszentrum Karlsruhe (FRG)
Institute für Radiochemie

Introduction

Reprocessing input solutions of spent nuclear fuel are analysed by independent laboratories and the results of the analysis are subsequently treated by a variety of proven mathematical and statistical methods /1/ in order to check for sampling errors, measurement errors, and diversion of material.

The purpose of this paper is to describe the application of two well-known chemometrical methods /2, 3/, i.e. principal component analysis and target factor analysis, to the evaluation of analysis data from spent nuclear fuel reprocessing input. This evaluation applies to measurements performed by one laboratory as well as to interlaboratory comparison.

Data Evaluation Method and Computer Program

General Description

Typical reprocessing input solution analysis data of spent fuel contained in a data matrix [D] (Fig. 1) consist of variables such as concentrations, concentration ratios of uranium and plutonium isotopes, isotope ratios, burn-up, total Pu and total U contents measured on r different input samples. In Figure 1 these variables are denoted isotopes for short.

Since these variables are not independent of each other, the data are redundant, i.e. they depend on a smaller number of latent or abstract factors, n, where $n < c$ = number of variables or isotopes. These abstract factors are independent of each other, they are sufficient to determine the entire set of (measured) variables. As will be shown later, the variance in the data investigated here can be described by $n = 3$ abstract factors only although up to 10 variables were measured. It should be stated that the degree and form of the dependence between the original variables is not relevant for the proposed data evaluation which is of purely statistical nature. In a principal component analysis (PCA) the abstract factors together with the number of abstract factors, n, are determined. The nature of these abstract factors is usually unknown, they might be related but not identical to burn-up, initial uranium enrichment, irradiation density during burn-up, etc.

In a second step of the PCA the original data [D] are compressed or reduced to a new data set [R*], which depends only on n abstract factors, thereby preserving the entire information of the original data set. Mathematically, the matrix [D] is split into a product of two matrices [R*]*[C*]. If there was no sampling or measurement error in the original data [D] and the number of abstract factors had been determined correctly, the product [D*] = [R*]*[C*] would be identical to the original data. Thus, the deviation between the original data [D] and the reproduced data [D*] can be used in a first attempt of outlier screening.

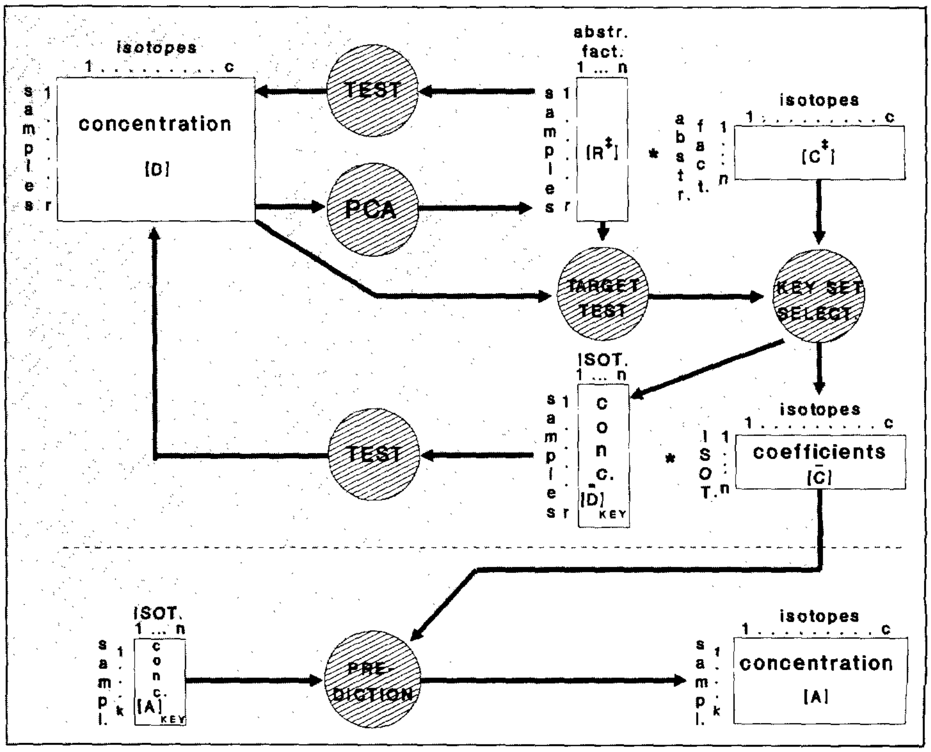


FIG. 1. *Chemometrical evaluation of reprocessing input analysis data*

For a successful PCA the number of samples, r, and the number of variables, c, should each be at least twice the number of abstract factors, n.

Since the physical nature of the abstract factors is not known, these factors are replaced by variables, i.e. isotope concentrations, etc., selected from the original data set in the target factor analysis (TFA). All variables are tested individually for their usefulness; those n variables which perform the test best and which are measurable easily and with high precision are selected as a key set. As will be shown later, U-235/238, Pu-241/240, and Pu-242/241 are chosen as key set variables in this context. These key set variables can be used to evaluate all other variables of the original data set (see equation 1).

The result of the target factor analysis is the key set $[\bar{D}]_{KEY}$ and a coefficient set $[\bar{C}]$ (see Fig. 1). The matrix product $[\bar{D}]_{KEY} * [\bar{C}]$ is the data matrix $[D]_{TFA}$ which as in PCA would be identical to the original data set [D] if there was no error in the original data. The impact of these errors in TFA is twofold : the errors of the data set [C*] and those of the isotopes of the key set will influence the deviation between original data set and reproduced data set $[D]_{TFA}$.

Writing explicitly the results of those matrix calculations for sample p and the above mentioned key set gives :

$$
\begin{aligned}
(U\text{-}234)_p &= c_{11}*(U\text{-}235/238)_p + c_{21}*(Pu\text{-}241/240)_p + c_{31}*(Pu\text{-}242/241)_p \\
(U\text{-}235)_p &= c_{12}*(U\text{-}235/238)_p + c_{22}*(Pu\text{-}241/240)_p + c_{32}*(Pu\text{-}242/241)_p \\
&\;\;\vdots \\
(Pu/U)_p &= c_{1c}*(U\text{-}235/238)_p + c_{2c}*(Pu\text{-}241/240)_p + c_{3c}*(Pu\text{-}242/241)_p
\end{aligned}
\qquad (1)
$$

Here $(U\text{-}234)_p$ means the calculated concentration of U-234 in sample p and the coefficients $\bar{c}_{hj}$ ($h = 1,...,n$; $j = 1,...,c$) are the data in the matrix $[\bar{C}]$. $(U\text{-}235/238)_p$, $(Pu\text{-}241/240)_p$, and $(Pu\text{-}242/241)_p$ are the measured isotope concentration ratios of sample p. The calculated concentrations can be used for searching outliers in the original data.

The geometrical meaning of the principal component analysis and the target factor analysis is the following. The original data of the r samples represent r points in a c-dimensional space. It is found by PCA where those r points are located in an n ($< c$)-dimensional space spanned by the orthogonal abstract factors $c_1 \ldots c_n$, whereby the rows in [R*] represent the coordinates of the data points in the abstract factor space and the rows in [C*] are the coordinates of the abstract factors in the original c-dimensional space, i.e. PCA performs a coordinate transformation. In a general target factor analysis it is tried to transform these abstract factor axes into axes that represent fundamental structural or physical parameters. Such physical parameters are difficult to determine for the data investigated here; instead, appropriate variables, i.e. data columns of the *original data* [D], are used to span the n-dimensional data space, i.e. they are used as new coordinate axes.

Many years ago the well-known isotope correlation technique (ICT) /4, 1/ had been developed for the data analysis of reprocessing input analysis data. In ICT several empirical and/or theoretical, based on reactor physics, linear regressions between isotope concentrations, isotope concentration ratios, etc., are used for outlier screening and prediction of new data. It now turns out that these one-dimensional linear regressions are special cases of the three-dimensional regressions found with PCA and TFA : two coefficents of the above regressions are very small in comparison to the third.

The most important feature of TFA is that of data prediction for new samples. As shown in the lower part of Figure 1, for new samples only the concentrations of the key set isotopes, $[A]_{key}$, (e.g. U-235/238, Pu-241/240, Pu-242/241) have to be measured in order to calculate the concentrations of all c variables in [A]. Since the coefficient matrix $[\bar{C}]$ is based on the original data [D], in addition to pointing out sampling and measurement errors, it can be checked whether the fuel of the new samples belongs to the same category as that of the original data.

Mathematical description

The details of the mathematical procedure are well described in the book published by Malinowski and Howery /2/. In order to facilitate understanding of the following, the nomenclature of that book will be used as far as possible.

It can be shown /2/ that the first n most important eigenvectors of the covariance matrix [Z]

$$[Z] = [D]^T * [D] \ , \qquad (2)$$

where $[D]^T$ is the transpose of [D], are identical to the abstract factors of the data matrix [D], whereas the last c - n eigenvectors represent the influence of the errors in the data set. The relative importance of the eigenvectors, i.e. the abstract factors, C_h, is given by the accompanying eigenvalues λ_h. The first eigenvector, C_1, describes most of the variance in the sample data, the second eigenvector, C_2, which is orthogonal to the first, describes most of the variance which is not explained by the first and so on.

For the determination of the number of abstract factors several methods are given in Ref. /2/. Since for the data under investigation no errors prior to the data evaluation were known, Malinowski's empirical indicator function IND /2/ was used, which usually is minimum for the proper number of abstract factors.

The data matrix [D] is then split into two matrices [C*] and [R*]. The matrix [R*] comprises, arranged in rows, the compressed data of the r samples with respect to the abstract factors, whereas the matrix [C*] describes the dependence of the n abstract factors upon the c variables, arranged in rows. The product [D*]

$$[D^*]_{r*c} = [R^*]_{r*n} * [C^*]_{n*c} \qquad (3)$$

should match the original data set [D] as closely as possible. The first and second subscripts of a matrix indicate the number of rows and columns of the matrix, respectively. The deviation between the two matrices represented by the residual mean square deviation, RMS_{PCA},

$$RMS_{PCA} = \sum_{j=1}^{c} RMS_j/c \qquad (4)$$

$$RMS_j = SQRT\ \{\sum_{i=1}^{r} (d_{ij}^* - d_{ij})^2 \ / \ (r-1)\ \} \ / \ d_j \qquad (5)$$

is a relative measure of successful data compression. RMS_j is the relative root mean square error of variable j, d_{ij}^* and d_{ij} are the elements of [D*] and [D], respectively, while d_j is the mean of variable j. Single gross deviations with respect to RMS_j can be considered as outliers.

In the target factor analysis applied here suited data columns $\bar{D}_h$ of the orignal data [D] are used as key set variables. For each of the $\bar{D}_h$ the correspondent data transformation vector T_h is calculated according to

$$T_h = [\lambda^*]^{-1} * [R^*] * \bar{D}_h \qquad (6)$$

where $[\lambda^*]^{-1}$ is the transform of the eigenvalue matrix, [R*] is the compressed data matrix after PCA in Eq. (2).

The applicability of $\bar{D}_h$ being used as a new coordinate called target vector is tested individually by checking the difference between $\bar{D}_h$ and the vector D_h obtained by transformation of the compressed data set [R*] :

$$\bar{\bar{D}}_h = [R^*] * T_h \qquad (7)$$

Out of those vectors $\bar{\bar{D}}_h$ for which the difference $\bar{D}_h - D_h$ is minimum several sets of n data (target) vectors each are built. Malinowski and Howery propose to retain as a key set the set $[\bar{\bar{D}}]_{KEY}$ which best satisfies

$$[\bar{\bar{D}}]_{KEY} * [\bar{C}] - [D] = min, \qquad (8)$$

where

$$[D]_{TFA} = [\bar{\bar{D}}]_{KEY}\ [\bar{C}] \qquad (9)$$

is the data set reproduced by target factor analysis and

$$[\bar{\bar{D}}]_{KEY} = [\bar{\bar{D}}_a,...,\bar{\bar{D}}_h,...,\bar{\bar{D}}_n] \qquad (10)$$

is the target factor matrix composed columnwise of original data vectors $\bar{\bar{D}}_a,...,\bar{\bar{D}}_h,...,\bar{\bar{D}}_n$.

Since the data investigated here are very different in magnitude, minimizing the mean of the relative column root mean square error as for PCA, i.e. minimizing RMS_{TFA} in Eq. (11) instead of minimizing Eq. (8)

$$RMS_{TFA} = \sum_{j=1}^{c} RMS_j^* \ / \ c = min \qquad (11)$$

$$RMS_j^* = SQRT\ \{\sum_{i=1}^{r} (d_{ij}^* - d_{ij})^2 \ / \ (r-1)\ \} \ / \ d_j \qquad (12)$$

d_{ij}^* = elements of the reproduced matrix $[D]_{TFA}$
d_j = mean of variable j

gave better reproductions for the uranium and plutonium isotope concentrations as well as for Pu/U which are the most important variables in a reprocessing input analysis.

The complete transformation matrix

$$[T]_{KEY} = [T_a,...,T_h,...,T_n] \qquad (13)$$

is composed columnwise of the transformation vectors T_h defined in Eq. (5). The coefficient matrix

$$[C] = [T]_{KEY}^{-1} * [C^*] \qquad (14)$$

where $[T]_{KEY}^{-1}$ is the transform of $[T]_{KEY}$, represents the coordinates of the sample points in the new n-dimensional space spanned by the chosen n target vectors, i.e. by n variables of the original data set.

The final step in the data evaluation procedure is that of data prediction. If the data measured for n (< c) isotopes on k new samples occur in the matrix $[A]_{KEY} = [A]_{k*n}$, then the data $[A]_{k*c}$ for those samples calculated for all c isotopes are simply obtained by matrix multiplication using the coefficient matrix $[\bar{C}]$ calculated in eq. (14) :

$$[A]_{k*c} = [A]_{k*n} * [\bar{C}]_{n*c} \qquad (15)$$

Computer Program

Malinowski and Howery's software package FACTANAL /5/ was adapted to Ryan-McFarland FORTRAN77 and modified to meet special needs. This code is the only one known to us, which, besides principal component analysis (and the KNN method), performs target factor analysis, but it has no provision for any graphical display. It is capable of handling data matrices of sizes up to 40 variables * 40 samples.

Data Origin and Preprocessing

Ten input batch samples of a WAK reprocessing campaign were analysed by four laboratories, each determining up to c = 16 variables, i.e. concentrations (weight percent) of uranium and plutonium isotopes, ratios of isotope concentrations (percent) as well as total uranium and total plutonium contents (g/kg total mass); one laboratory determined also the burn-up, Ft. Details can be found in Ref. /1/.

These data were transformed into dimensional values of [atoms per initial metal atoms] = [IMA]. The data of the laboratory chosen for the detailed PCA and TFA are shown in Table I; the data of all four laboratories were used for an interlaboratory comparison.

Since the absolute values of the variables, i.e. isotope concentrations, etc., are very different, it was found that prior to the principal component analysis scaling of the data columnwise by factor F_j results in the optimum reproduction of data in PCA (minimum RMS_{PCA} in Eq. (4)) as wel as in TFA (minimum RMS_{TFA} in Eq. (11)). Factor F_j is defined as

$$\begin{aligned} F_j &= 1 \ / \ STD(D_j) \\ j &= 1,...,c \qquad (16) \\ c &= \text{number of original data columns} \end{aligned}$$

where the standard deviation, $STD(D_j)$, of the column vector D_j (variable in the sense described above) is defined as

$$STD(D_j) = SQRT\ \{\sum_{i=1}^{r} (d_{ij} - d_j)^2 \ / \ (r-1)\ \} \ / \ d_j \qquad (17)$$

d_{ij} = element of the original data matrix [D]
d_j = mean of d_{ij}
r = number of samples.

Batch samples 86 to 94 were chosen for principal component analysis and target factor analysis leaving sample 95 out for the demonstration of data prediction.

TABLE I — Original data measured by one laboratory and transformed into dimensions of (atoms/initial metal atoms)

	1 Ft	2 U-234	3 U-235	4 U-236	5 U-238
86	2.953E+00	1.956E-04	9.688E-03	3.713E-03	9.4801-01
87	3.151E+00	1.855E-04	8.980E-03	3.832E-03	9.467E-01
88	2.945E+00	1.957E-04	9.870E-03	3.716E-03	9.484E-01
89	3.024E+00	1.760E-04	9.568E-03	3.760E-03	9.477E-01
90	3.076E+00	1.954E-04	9.241E-03	3.865E-03	9.473E-01
91	2.909E+00	2.056E-04	9.991E-03	3.727E-03	9.486E-01
92	2.758E+00	1.864E-04	1.069E-02	3.686E-03	9.499E-01
93	3.043E+00	2.150E-04	9.762E-03	3.799E-03	9.474E-01
94	2.696E+00	1.864E-04	1.142E-02	3.639E-03	9.497E-01
95	2.920E+00	1.664E-04	9.980E-03	3.688E-03	9.487E-01

	6 Pu-238	7 Pu-239	8 Pu-240	9 Pu-241	10 Pu-242
86	1.181E-04	5.197E-03	2.062E-03	1.130E-03	4.014E-04
87	1.184E-04	4.955E-03	2.077E-03	1.152E-03	4.414E-04
88	9.912E-05	4.917E-03	1.938E-03	1.057E-03	3.751E-04
89	1.057E-04	5.008E-03	1.997E-03	1.097E-03	4.010E-04
90	1.147E-04	4.974E-03	2.040E-03	1.123E-03	4.266E-04
91	9.770E-05	4.938E-03	1.919E-03	1.049E-03	3.683E-04
92	8.351E-05	4.722E-03	1.801E-03	9.749E-04	3.266E-04
93	1.027E-04	4.890E-03	1.974E-03	1.075E-03	3.984E-04
94	8.150E-05	4.886E-03	1.815E-03	9.837E-04	3.135E-04
95	9.410E-05	4.886E-03	1.926E-03	1.042E-03	3.672E-04

	11 U235/238	12 Pu240/239	13 Pu241/240	14 Pu242/240	15 Pu242/241
86	1.022E-02	3.967E-01	5.483E-01	1.947E-01	3.550E-01
87	9.485E-03	4.191E-01	5.549E-01	2.125E-01	3.830E-01
88	1.041E-02	3.942E-01	5.451E-01	1.935E-01	3.550E-01
89	1.010E-02	3.988E-01	5.495E-01	2.008E-01	3.654E-01
90	9.755E-03	4.101E-01	5.504E-01	2.091E-01	3.800E-01
91	1.053E-02	3.887E-01	5.468E-01	1.919E-01	3.509E-01
92	1.126E-02	3.813E-01	5.414E-01	1.813E-01	3.350E-01
93	1.030E-02	4.036E-01	5.444E-01	2.019E-01	3.708E-01
94	1.203E-02	3.715E-01	5.418E-01	1.727E-01	3.187E-01
95	1.052E-01	3.942E-01	5.412E-01	1.907E-01	3.523E-01

	16 Pu/U
86	9.215E-03
87	9.101E-03
88	8.776E-03
89	8.946E-03
90	9.017E-03
91	8.683E-03
92	8.255E-03
93	8.772E-03
94	8.354E-03
95	8.625E-03

Results

Principal Component Analysis

The results of the principal component analysis are summarised in Table II. The importance of each absctract factor is given by its eigenvalue, showing the predominant influence of the first factor in comparison to the following ones. The mean deviation, RMS_{PCA}, between original and reproduced data, [D*], decreases with the addition of abstract factors whereas the total variance, VAR CUM, explained by the abstract factors is already 99.998% for one abstract factor and reaches 99.9998% for three abstract factors included in the principal component analysis. This means that one abstract factor should be sufficient to explain the variance in the data if the prediction of the isotope concentrations, which are much smaller in magnitude than burn-up and isotope concentration ratios, were precise enough with one factor only. Finally, the number of measurement values for which the deviation between original and recalculated values are within certain limits, are given in the last four columns of Table II, which shows that 112 out of 144 data points are recalculated with deviation less than 1% and only six measurements deviate by more than 5% for three factors included in the analysis.

It was tried to determine the exact number of abstract factors using Malinowski's indicator function on unscaled data for which it had been defined. The analysis of each of the data sets of the four laboratories indicated three to five abstract factors, thus showing that this function is not adequate for this kind of data evaluation. Wold's procedure /6/ might give better results. The minimum number of three was chosen for the following analysis assuming that the higher number found for some of the laboratories is due to measurement errors.

In order to give better insight into the dependence of variables on abstract factors the PCA results are represented graphically in Figs. 2 and 3 for **unscaled** data. Fig. 2 demonstrates that the first two abstract factors, C_1 and C_2, have the greatest influence on burn-up, Ft, U-238 and the isotope concentration ratios whereas all other isotope concentrations as well as Pu/U are located in one point at the origin if all 16 variables are included. Repeating the principal component analysis, including only isotope concentrations and Pu/U (Fig. 3), reveals that U-235 and U-238 are strongly correlated to C_2 and C_1, respectively, whereas all other isotopes and Pu/U depend only, but to a smaller extent, on C_2 (and of course on C_3). Fig. 4 shows, first, that the variance in the **scaled** sample data is maximum for the first abstract factor, C_1, according to the assumption of PCA and, second, that the samples in the two-dimensional factor space spanned by the two largest abstract factors, C_1 and C_2, scatter around the first abstract factor, C_1, which shows that one abstract factor is not sufficient to explain the variance in the data for all isotope concentrations and Pu/U. Apparently, there is no sample which does not fit into the overall pattern in Fig. 4; thus, all samples seem to stem from similar fuel assemblies. It also shows that there is no dependence of the individual fuel assembly history on abstract factors; otherwise the sample pairs 86+87, 88+89, 90+91, 92+93, which stem from the fuel assemblies # 168, 171, 176, and 172, respectively /1/, should be close to each other in the abstract factor space.

Table III shows the difference (%) between original and recalculated data relative to the original data for three abstract factors and the root mean square error relative to the variable mean, RMS_j, for each variable. These mean deviations are less than 1.5% except for U-234 (5.8%) and for Pu-238 (2.16%) which is due to low precision in measurement. The overall mean deviation, RMS_{PCA} (Eq.(4)), amounts to 1.04%. No single gross outliers with respect to the variable mean deviation, RMS_j, are found.

Target Factor Analysis

The best results in a target factor analysis according to Eq. (11) were obtained with U-235/238, Pu-241/240 and Pu-242/241 as target vectors or key set variables.

TABLE II — Results of the principal component analysis on scaled data of Table I

Factor	Eigenvalue	RMS_{PCA} Eq. (4)	VAR CUM (%)	Number of points with deviation (original-calculated data) < 1%	1-5%	5-10%	> 10%
1	6.87E+06	2.11E-02	99.9982	41	60	26	17
2	1.03E+02	4.64E-03	99.9997	88	51	4	1
3	1.12E+01	3.15E-03	99.9998	112	26	6	0
4	8.18E+00	2.98E-03	100.000	128	16	0	0
5	9.57E-01	2.96E-03	100.000	136	8	0	0
*	*	*	*	*	*	*	*
*	*	*	*	*	*	*	*
8	3.25E-01	8.62E-05	100.000	144	0	0	0

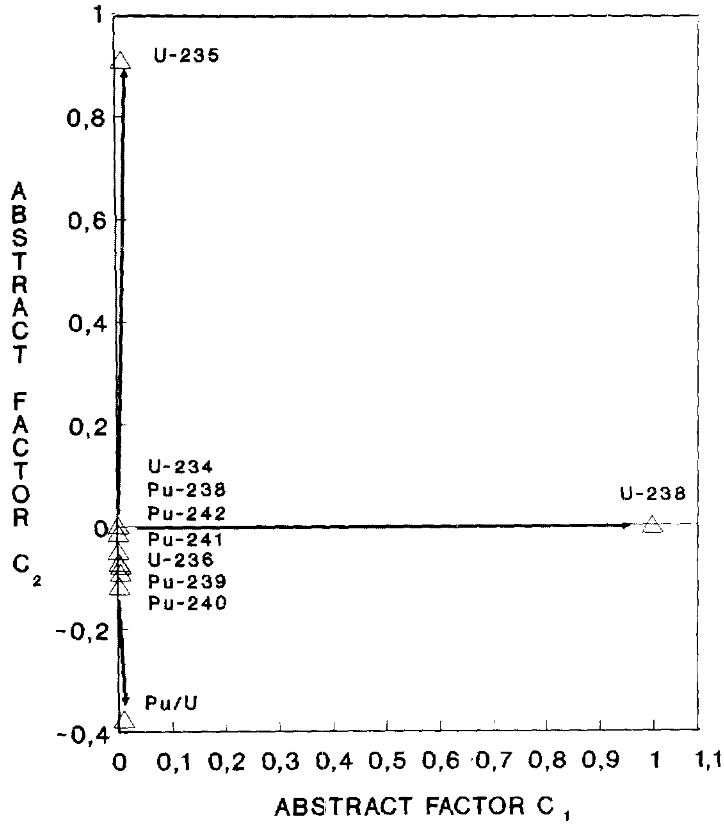


FIG. 3. *Geometrical relationship between abstract factor axes, C_1 and C_2, and Pu/U, uranium and plutonium isotope concentration axes, data unscaled*

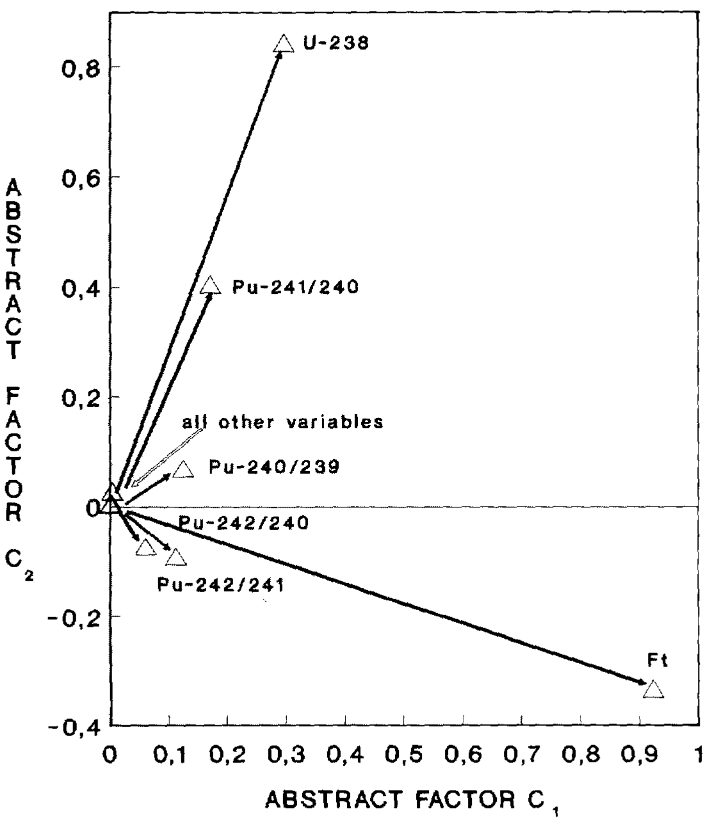


FIG. 2. *Geometrical relationships between abstract factor axes, C_1 and C_2, and variable axes 1-16, data unscaled.*

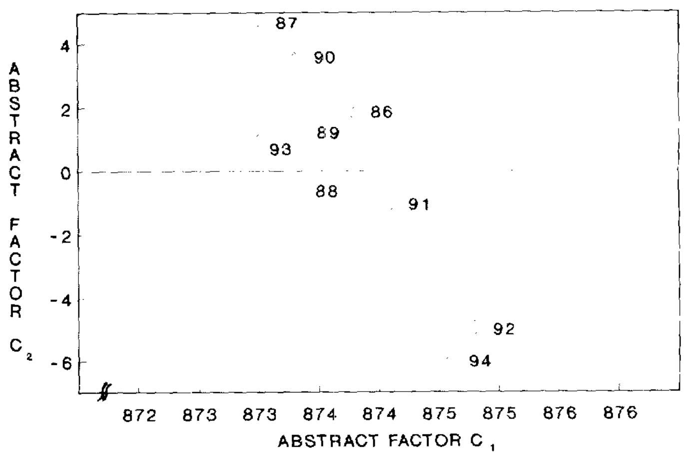


FIG. 4. *Samples 86-94 in the abstract factor space spanned by C_1 and C_2, data scaled*

For this key set the mean deviations, RMS_j^* (Eq. (12)), of each variable are shown in Table III. The overall mean deviation, RMS_{TFA} (Eq. (11)), equals 1.58% and is slightly higher than the mean difference of the PCA, RMS_{PCA} = 1.04%, this is due to the measurement errors introduced by the key set variables. The uranium isotope concentrations (except for U-234) are reproduced with mean deviations of less than 0.5%, whereas the plutonium isotope concentrations and Pu/U show deviations of less than 2.5%. The latter deviations will certainly become smaller with the number of samples increasing. Again, no gross outliers with respect to the variable mean deviation, RMS_j^*, were detected if one keeps in mind that any outlier test will be more significant the smaller the variable mean deviation is.

Data Prediction

For sample 95 data were predicted on the basis of its measured concentration ratios of U-235/238, Pu-241/240, and Pu-242/241 using Eq. (1) and the coefficients c_{hj} calculated by TFA (Eq. (14)) for samples 86-94. Table IV shows the relative deviations between predicted and measured data. Except for U-234, which must be an outlier all deviations are smaller than the variable mean deviation, RMS_j^*, of Table III, thus demonstrating that the target factor analysis can be used for prediction and/or outlier detection of new samples.

Interlaboratory Comparison

In the interlaboratory comparison for each variable a principal component analysis was performed on the matrix consisting of data measured by c = 4 laboratories on r = 7 samples. The reduction in the number of samples is due to the fact that one of the laboratories measured only 7 out of 10 samples. Scaling of data is not necessary since there is no difference in absolute magnitude; therefore, the exact number of abstract factors can be determined well by

TABLE III — Deviation (%) between data calculated by principal component analysis using 3 abstract factors and original data

	1 Ft	2 U-234	3 U-235	4 U-236	5 U-238	6 Pu-238	7 Pu-239	8 Pu-240
86	0.66	−6.23	0.01	0.31	0.00	−3.14	−0.88	−0.76
87	−0.17	5.58	−0.49	0.14	0.01	0.88	1.45	0.37
88	−0.54	−0.85	1.60	0.69	−0.01	0.64	0.03	0.02
89	−1.12	7.78	1.01	−0.22	0.00	4.12	0.66	0.84
90	1.00	1.18	−0.83	−0.95	0.00	−0.88	−0.16	0.12
91	0.31	−5.18	1.08	0.35	0.00	0.13	−0.85	0.28
92	0.64	5.42	1.48	0.00	0.00	−1.56	1.11	0.51
93	−0.52	−6.03	−1.60	0.23	0.00	−1.10	−1.31	−1.08
94	−021	0.53	−2.14	−0.53	0.01	1.24	0.02	−0.27
RMS_j(%)	0.70	5.28	1.44	0.51	0.01	2.16	0.93	0.62
RMS_j^*(%)	0.69	5.50	0.04	0.42	0.33	5.74	2.30	2.33

	9 Pu-241	10 Pu-242	11 U235/238	12 Pu240/239	13 Pu241/240	14 Pu242/240	15 Pu242/241	16 Pu/U
86	−0.39	0.17	0.02	0.14	0.38	0.80	0.40	−0.68
87	−0.18	−0.22	−0.47	−0.97	−0.50	−0.16	0.40	0.79
88	0.25	0.38	1.59	−0.05	0.19	0.07	−0.16	−0.59
89	0.80	−0.23	0.99	0.17	−0.08	−1.22	−1.18	0.59
90	0.11	−0.11	−0.81	0.32	0.02	−0.05	−0.04	0.09
91	0.10	1.01	1.09	1.12	−0.19	0.45	0.64	−0.10
92	0.39	0.05	1.51	−0.49	−0.03	−0.31	−0.21	0.52
93	−0.59	−0.35	−1.65	0.12	0.42	0.45	−0.04	−0.80
94	−0.48	−0.71	−2.16	−0.66	−0.20	−0.02	0.23	0.21
RMS_j(%)	0.45	0.46	1.45	0.59	0.30	0.59	0.54	0.59
RMS_j^*(%)	2.40	2.53		0.62		0.34		2.00

Root mean square error, RMS_j, calculated with Eq. (5) and RMS_j^* calculated with Eq. (12)

TABLE IV — Relative deviation (%) between predicted and original data of sample 95

Ft −0.49	U-234 17.09	U-235 −0.01	U-236 0.84	U-238 −0.82
Pu-238 2.65	Pu-239 −1.06	Pu-240 −1.04	Pu-241 −0.22	Pu-242 0.89
Pu-240/239 −0.80	Pu-242/240 0.83	Pu/U −0.50		

Prediction from sample 95 measured data U-235/238, Pu-241/240, Pu-242/241, and coefficients calculated from data of samples 86-94

Table V — Results of principal component analyses for each variable

	RMS_j (%) in				n	outlier	max. deviation	
	lab 1	lab 2	lab 3	lab 4		#	(%)	in
U-234	4.86	15.1	8.81	3.81	1	93	−10.4	lab 3
U-235	0.53	0.79	1.36	0.59	1	94	− 2.6	lab 3
U-236	0.38	0.34	0.97	1.27	2	95	− 1.8	lab 4
U-238	0.04	0.01	0.03	0.02	1			
Pu-238	6.39	5.72	8.34	19.2	2	92	−29.0	lab 4
Pu-239	2.15	0.88	0.91	1.82	1	95	− 3.8	lab 1
Pu-240	2.12	1.00	0.98	1.61	1	95	− 3.8	lab 1
Pu-241	2.44	1.19	1.21	1.45	1	95	− 4.4	lab 1
Pu-242	2.85	6.00	1.84	4.62	1	93	−10.3	lab 2
U-235/238	0.54	0.80	1.38	0.60	1	94	− 2.6	lab 3
Pu-240/239	0.17	0.22	0.13	0.40	1			
Pu-241/240	0.34	0.37	0.31	0.37	1			
Pu-242/240	1.85	5.41	1.49	3.73	1	93	− 9.9	lab 2
Pu-242/241	1.85	5.22	1.62	3.71	1	93	− 9.6	lab 2
Pu/U	1.74	1.95	1.60	2.06	1	95	3.7	lab 4

Data matrices comprise measurement data of 4 laboratories on samples 88, 90-95. Root mean square error of columns, RMS_j, according to Eq. (5), n = number of abstract factors. For each variable only the maximum deviation >1.0% between calculated and original data is given.

Malinowski's indicator function. Table V shows that, except for U-236 and Pu-238, the data can be explained by one abstract factor only, i.e. there is no bias between laboratory measurements. For each laboratory the RMS_j error according to Eq. (5) is given. This reveals that the Pu-238 bias mentioned above is due to laboratory 4, whereas for U-236 a distinct bias between lab. 1 and lab. 2, on the one hand, and lab. 3 and lab. 4, on the other hand, seems to exist. Finally, in the last three columns of Table V the largest deviation $> 1\%$ between calculated and original data is shown for each variable. These deviations partially agree with the outliers detected in ref. /1/. Further insight into interlaboratory differences might be obtained by applying Malinowski and Howery's uniqueness and unity tests /2/.

Conclusions

Despite the restricted number of samples involved, the investigation demonstrates that principal component analysis and a subsequent target factor analysis of reprocessing input analysis data is a useful tool for screening outlying data, for the prediction of new data on the basis of known data, and, to a certain extent, for the comparison of laboratory measurements.

References

/1/ KOCH, L., SCHOOF, S. (Eds.) — The Isotope Correlation Experiment ICE. ESARDA Report 2/81, EUR 7766.EN (1981), KfK 3337

/2/ MALINOWSKI, E.R., HOWERY, D.G. — Factor Analysis in Chemistry. Wiley, New York (1980)

/3/ MASSART, D.L., VANDEGINSTE, B.G.M., DEMING, S.N., MICHOTTE, Y., KAUFMAN, L. — Chemometrix. A Textbook. Series: Data Handling in Science and Technology, Vol. 2, Elsevier, Amsterdam (1988)

/4/ CHRISTENSEN, D.E., SCHNEIDER, R.A., STEWART, K.B. — Summary of Experience with the Use of Isotopic Correlation Safeguards Techniques. BNWL-SA-4274 (March 20, 1972)

/5/ MALINOWSKI, E.R., HOWERY, D.G., WEINER, P.H., SOROKA, J.H., FUNKE, P.T., SELZER, R.S., LEVINSTONE, A. — Program "FACTANAL" (Program No. QCMP040, Quantum Chemistry Program Exchange, Indiana Univ., Chemistry Building 204, Bloomington, Ind. 47405) (1976)

/6/ WOLD, S. — Cross-Validatory Estimation of the Number of Components in Factor and Principal Components Models. *Technoimetrics*, **20** (1978) 397-405

LB-AB-90-018-EN-N