ELSEVIER

Chemometrics and Intelligent Laboratory Systems 58 (2001) 59–72

Chemometrics and intelligent laboratory systems

www.elsevier.com/locate/chemometrics

Multi-way modelling of high-dimensionality electroencephalographic data

F. Estienne [a], N. Matthijs [a], D.L. Massart [a,*], P. Ricoux [b], D. Leibovici [c]

[a] *ChemoAC, Pharmaceutical Institute, Pharmaceutical and Biomedical Analysis, Vrije Universiteit Brussel, Laarbeeklaan 103, 1090 Brussels, Belgium*

[b] *ELF, Lyon, France*

[c] *Image Analysis Group, FMRIB Centre, Oxford University, John Radcliffe Hospital, Oxford OX3 9DU, UK*

Received 12 March 2001; accepted 22 June 2001

Abstract

The aim of this study is to investigate whether useful information can be extracted from an electroencephalographic (EEG) data set with a very high number of modes, and to determine which model is the most appropriate for this purpose. The data was acquired during the testing phase of a new drug expected to have effect on the brain activity. The implemented test program (several patients followed in time, different doses, conditions, etc....) led to a six-way data set. After it was confirmed that the exploratory analysis of this data set could not be handled with classical principal component analysis (PCA), and it was verified that multidimensional structure was present, multi-way methods were used to model the data. It appeared that Tucker 3 was the most suited model. It was possible to extract useful information from this high-dimensionality data. Non-relevant sources of variance (outlying patients for instance) were identified so that they can be removed before the in-depth physiological study is performed.

Keywords: Multi-way methods; Tucker 3; PARAFAC; Exploratory analysis; Electroencephalography; EEG

1. Introduction

The general aim of this study was to investigate the effect of a new antidepressant drug on the brain activity using electroencephalographic (EEG) data. The scope of the present paper is not to present advances in the field of neuroscience, but to show how multi-dimensional models can efficiently be applied to extract useful information from multi-way data even with high dimensionality (up to six modes in this study).

The principle of electroencephalography is to give a representation of the electrical activity of the brain [1]. Electroencephalography is mainly used for the detection and management of epilepsy. It is a non-invasive way of detecting structural abnormalities such as brain tumours. It is also used for the investigation of patients with other neurological disorders that sometimes lead to characteristic EEG abnormalities, or like in the present study, to determine the effect of a drug on the brain activity. This activity is measured

* Corresponding author. Tel.: +32-2-477-4734; fax: +32-2-477-4735.

E-mail address: fabi@vub.vub.ac.be (D.L. Massart).

0169-7439/01/$ - see front matter
PII: S0169-7439(01)00140-X

using metal electrodes placed on the scalp. Even if no general agreement was reached concerning the placement of the electrodes, most of the laboratories use the so-called International 10–20 system [2]. These measurements lead to electroencephalograms that can be used directly, as in case of abnormality they can present characteristic patterns, or can be treated with Fourier Transform to keep only the numerical values corresponding to the average energy of specific frequency bands.

2. Experimental

The data were acquired during the testing phase of a new antidepressant drug. This test program was a phase II (a small group of healthy volunteers is studied), mono-centric (all the experiments are performed in the same place), placebo-controlled, double blind (neither the patient, nor the doctor know whether the drug or the placebo is being administered) trial. The study was performed on 12 healthy male subjects, and the effect of four doses (placebo, 10, 30 and 90 mg) was investigated. This effect was followed in time over a 2-day period (8:00, 8:30, 9:30, 10:00, 10:30, 11:00, 11:30, 12:00 AM, 1:00 and 3:00 PM on the first day, 9 AM and 9 PM on the second day: 12 measurements). The EEGs were measured on 28 leads (augmented 10–20 system) located on the patient scalp (Fig. 1), and were repeated twice. The first measurement was performed in the so-called "resting" condition, where the patient is lying with eyes closed in a silent room. The second measurement was performed in the "vigilance-controlled" condition, where the subject is asked to perform simple tasks while the EEGs are acquired. Overall, 32256 EEG measurements were performed. Each of the EEG (at a given time, for one of the leads, on one patient, who was administered a certain dose of the substance, in one measurement condition) was decomposed using the Fast Fourier Transform into seven energy bands (α_1, α_2, β_1, β_2, β_3, δ, θ) commonly used in neuroscience [1]. Therefore, only the numerical value corresponding to the average energy of specific frequency bands is taken into account. The data was provided in the form of a table with dimensions (32256×7) with no possibility of coming back

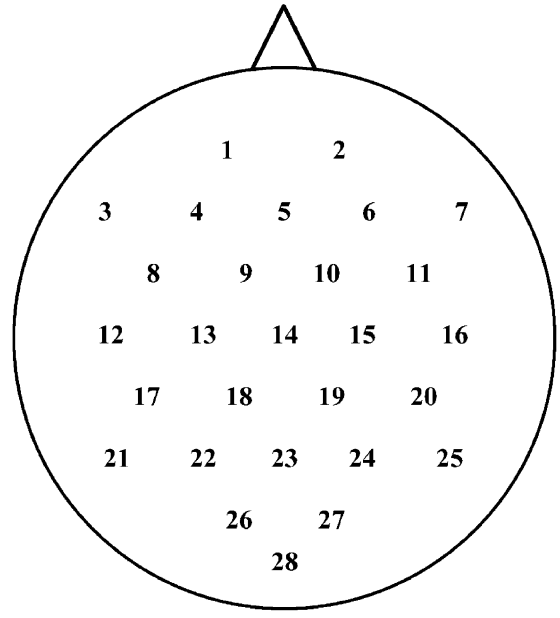

Fig. 1. Augmented 10–20 system, location of the 28 leads on the scalp.

to the original electroencephalograms. The (32256×7) table was reorganised into a multidimensional array. The resulting matrix is a six-way array with dimension ($7 \times 12 \times 28 \times 4 \times 12 \times 2$). The dimensions (or modes) are described as follows:

EEG dimension	Seven EEG bands (α_1, α_2, β_1, β_2, β_3, δ, θ)
Subject dimension	12 patients
Spatial dimension	28 leads
Dose dimension	Four doses (placebo, 10, 30 and 90 mg)
Time dimension	12 EEG measurements over 2 days
Condition dimension	Two measurement conditions (resting and vigilance controlled)

The calculations were performed on a Personal Computer with an AMD Athlon 600 MHz CPU and 256 Mega Bytes of RAM. The software used was house made or parts of *The N-way Toolbox* from Andersson and Bro [3]. The whole study was performed in the Matlab® environment.

3. Models

3.1. Unfolding PCA—Tucker 1

Unfolding Principal Component Analysis (PCA) consists in applying classical two-way PCA on the data matrix after it has been unfolded. The principle of unfolding is to consider the multidimensional matrix as a collection of regular two-way matrices and to put them next to another, leading to a new two-way matrix containing all the data. It is possible to unfold a three-way matrix along the three dimensions (Fig. 2). This results in three different matrices $\mathbf{X}_{(1)}$, $\mathbf{X}_{(2)}$ and $\mathbf{X}_{(3)}$ in which modes 1, 2 and 3 are, respectively, preserved. The score matrices obtained building a PCA model on each of those three matrices, respectively, called **A**, **B** and **C**, are the output of a Tucker 1 model. Tucker 1 is considered a weak multidimensional model, as it does not take into account the multi-way structure of the data. The **A**, **B** and **C** matrices are independently built. The Tucker 1 model is a collection of independent bilinear models, and not a multilinear model.

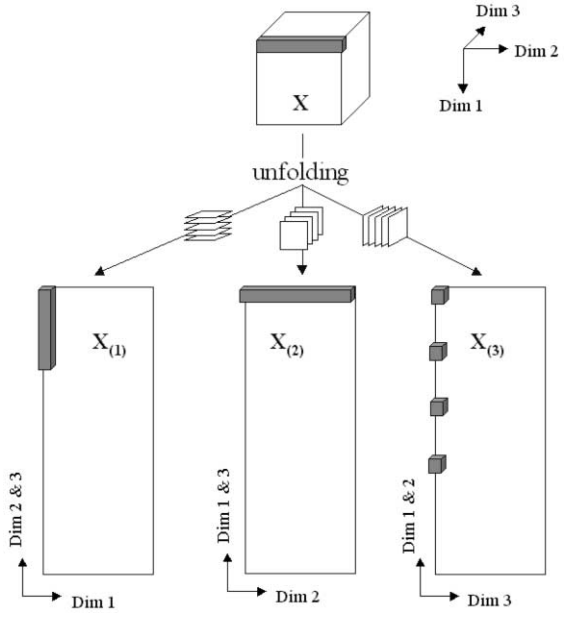

Fig. 2. Three possible ways of unfolding a three-way array **X**. $\mathbf{X}_{(1)}$, $\mathbf{X}_{(2)}$ and $\mathbf{X}_{(3)}$ are the two-way matrices obtained after unfolding with preserving the 1st, 2nd and 3rd mode, respectively.

3.2. Tucker 3

The Tucker 3 [4,5] model is a generalisation of bilinear PCA to data with more modes. The Tucker 3 model (limited here to a three-way case for sake of simplicity) can be formulated as in Eq. (1).

$$x_{ijk} = \sum_{l=1}^{w1} \sum_{m=1}^{w2} \sum_{n=1}^{w3} a_{il} b_{jm} c_{kn} g_{lmn} \qquad (1)$$

where x_{ijk} is an $(l \times m \times n)$ multidimensional array, w_1, w_2 and w_3 are the number of components extracted on the 1st, 2nd and 3rd mode, respectively, a, b, and c are the elements of the **A**, **B** and **C** loadings matrices for the 1st, 2nd and 3rd mode, respectively, and g are the elements of the core matrix **G**.

The information carried by these matrices is therefore of the same nature as the information contained in the equivalent matrices of the Tucker 1 model. The difference comes from the fact that these matrices are built simultaneously during the Alternating Least Squares (ALS) fitting process of the model in order to account for the multidimensional structure. Tucker 3 is a multilinear model. Moreover, the **G** matrix defines how individual loading vectors in the different modes interact. This information is not available in the Tucker 1 model. The Tucker 3 model can also be seen in a more graphical way as shown in Fig. 3, it appears as a weighted sum of outer products between the factors stored as columns in the **A**, **B** and **C** matrices.

One of the interesting properties of the Tucker model is that the number of components for the different modes does not have to be the same (as is the case in the PARAFAC model). In Tucker 3, the components in each mode are usually constrained to orthogonality, which implies a fast convergence. A limitation of this model is that the solution obtained is not unique, an infinity of other equivalent solutions can be obtained by rotating the result without changing the fit of the model.

3.3. PARAFAC

The PARAFAC model [6,7] is another generalisation of bilinear PCA to higher order data. It can be mathematically described as in Eq. (2):

$$x_{ijk} = \sum_{l=1}^{w} a_{il} b_{jl} c_{kl}. \qquad (2)$$

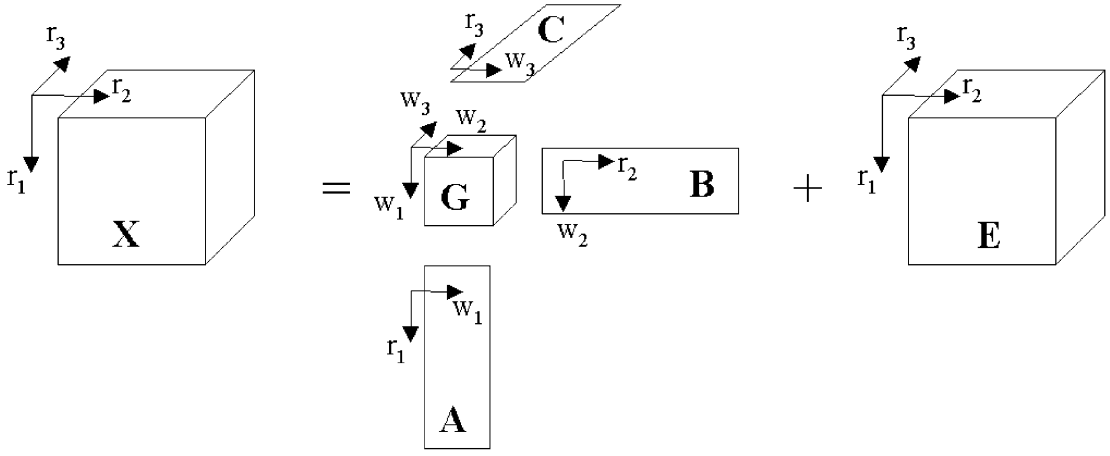

Fig. 3. Representation of the Tucker 3 model applied to a three-way array **X**. **A**, **B** and **C** are the loadings corresponding, respectively, to the 1st, 2nd and 3rd dimension. **G** is the core matrix. **E** is the matrix of residuals.

Like Tucker 3, PARAFAC is a real multilinear model. It can be considered as a special case of the Tucker 3 model, in which the number of components extracted along each mode would be the same, and the core matrix would contain only non-zero elements on its diagonal. This specific structure of the core makes PARAFAC models much easier to interpret than Tucker 3 models. The PARAFAC model can also be seen in a more graphical way as shown in Fig. 4. The most interesting feature of the PARAFAC model is uniqueness. The model provides unique factor estimates, the solution obtained cannot be rotated without modification of its fit. As components on each mode are not constrained to or-

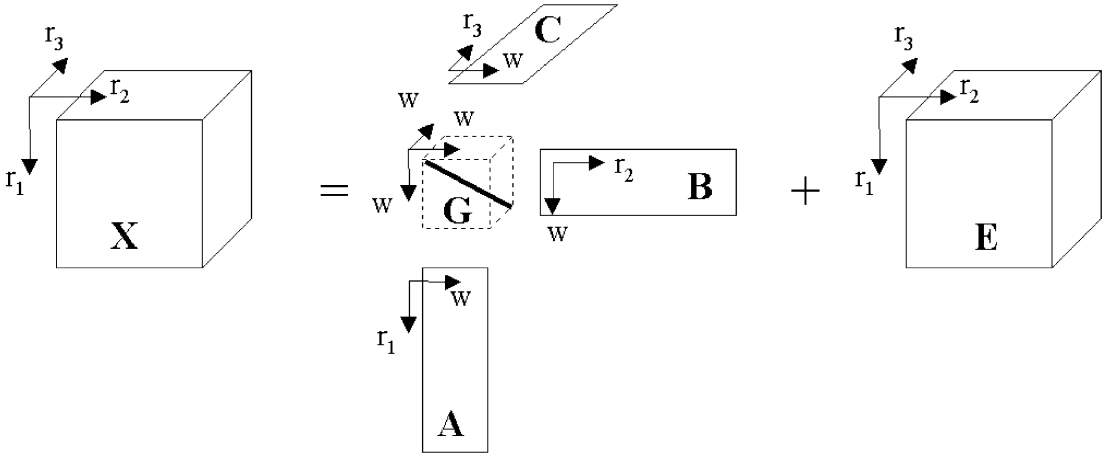

Fig. 4. Representation of the PARAFAC model applied to a three-way array **X**. **A**, **B** and **C** are the loadings corresponding, respectively, to the 1st, 2nd and 3rd dimension. **G** is the super-diagonal core matrix. **E** is the matrix of residuals.

thogonality, the convergence is usually quite slower than observed with the Tucker 3 model.

4. Results and discussion

4.1. Linear and bilinear models

Because of the nature of the data set, it was very difficult to explore it visually the way it is usually done, for instance, with spectral data. In order to get a better insight of the data, some averages were computed directly from the original variables. This corresponds to building simple linear models. The global average (on patients, doses and conditions) for the energy bands can then be displayed on a map of the brain for each of the measurement times. It is then possible to see in a rough way the evolution of the activity of the brain as a function of time and of the location in the brain (Fig. 5). It can be seen that the activity of the brain seems to globally increase to reach a maximum at time 6 (11:00 AM, first day).

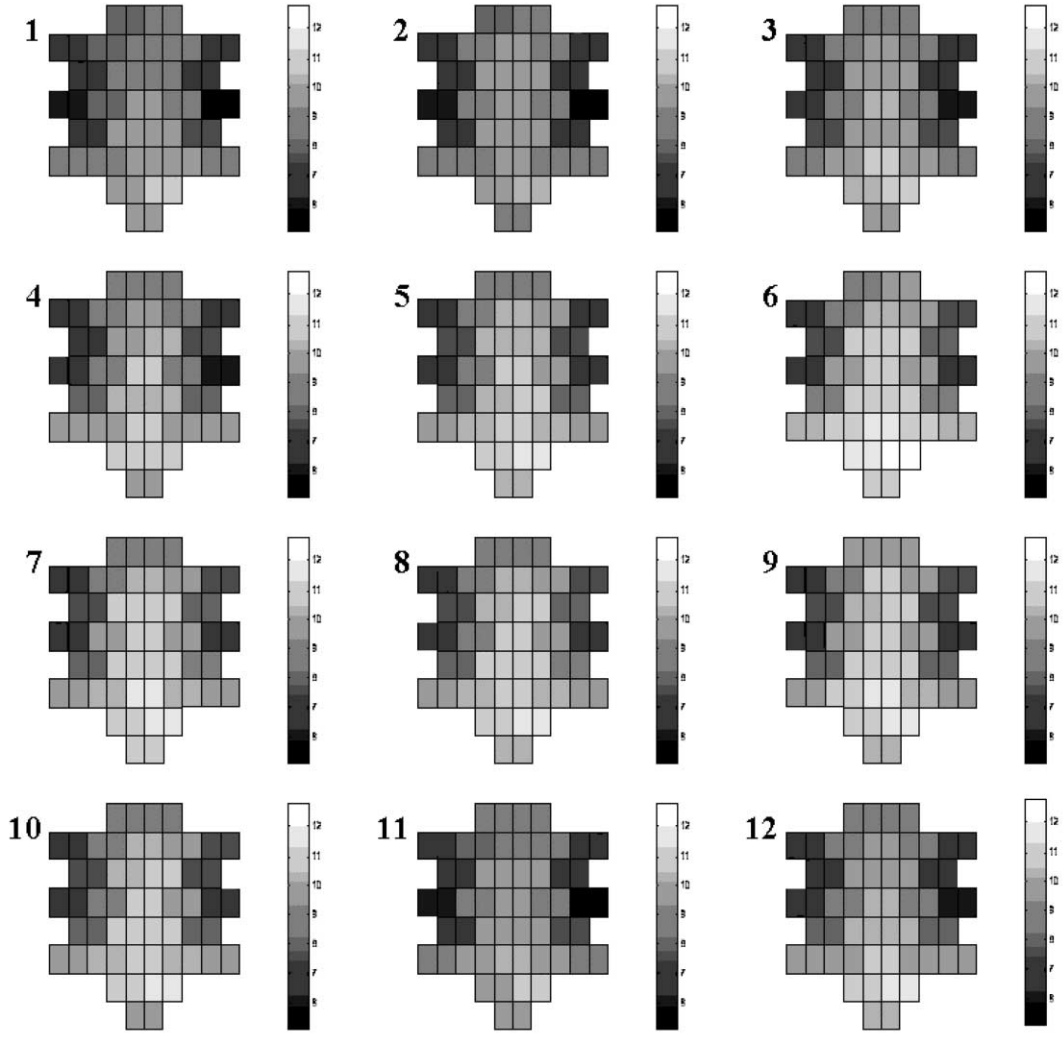

Fig. 5. Original data (averaged on patients, doses, conditions, and the seven energy bands) displayed, for each of the measurement times, on a grid representing the electrodes locations. Dark zones indicate low activity.

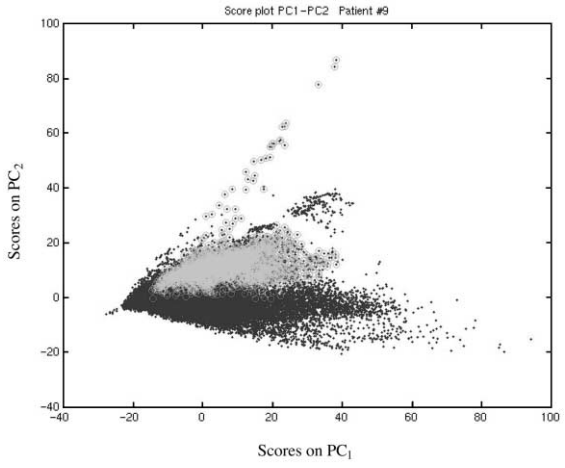

Fig. 6. Results of PCA on the (7×32256) matrix: scores on PC_1 versus scores on PC_2. Points corresponding to patient #9 are highlighted.

The activity seems to increase mainly in the back part of the brain. The plot corresponding to time 11 (9:00 AM, second day) shows that the state of the brain seems to be similar on the first and second day at equivalent times. Studying such plots for individual energy bands shows that the different bands are not all present and varying in the same parts of the brain (i.e. some are more present and active in the front or back part of the brain).

Classical two-way PCA can also be used to explore this data set. Bilinear models are then constructed. The intensities of the seven energy bands are considered as variables, and the 32256 measurement conditions as objects. The PCA results (Fig. 6) show that there is some structure in the data. Points of the score plot corresponding to an individual patient are located in relatively well-defined areas. The same thing can be observed for points corresponding to a certain electrode or dose. However, the results are too complex to be readily interpretable, and justify the use of multi-way methods to explore this data set.

4.2. Assessing multilinear structure

Many data sets can be arranged in a multi-way form. This does not mean that multi-way methods should be applied on such data sets as using such methods makes sense only if multilinear structure is present in the data. For instance, if slices of a three-way array are completely independent, no structure (or correlation) is present along this mode, and multi-way methods should not be used. Two-way PCA can be used to ensure that some multidimensional structure is actually present in the data. The data can be reduced to a smaller dimensionality (smaller number of modes) array by extracting parts of the array corresponding to one element of a given mode. For instance, considering only patient #11, the 30-mg dose, and the resting condition, the resulting matrix is a three-way array with dimension ($28 \times 12 \times 7$). Only the spatial, time, and variable dimensions are then taken into account. This matrix has to be unfolded before ordinary PCA can be performed. If the data is unfolded preserving the first dimension, the resulting matrix will have dimension ($28 \times (12 \times 7)$). The scores of a PCA model performed on this data give information about the 28 electrodes, and the loadings give simultaneously information about the time and the variables, 12 repetitions (one per time) of the information about the seven variables are expected. It is verified that there is a structure remaining in the loadings of the PCA model (Fig. 7). The loadings for each variable globally vary following a common time profile. This is an indication of a dimensional structure between the time and variable dimensions in the data used. A ($7 \times (28 \times 12)$) array can also be obtained by rearranging the previous matrix. This time, the loadings show combined informa-

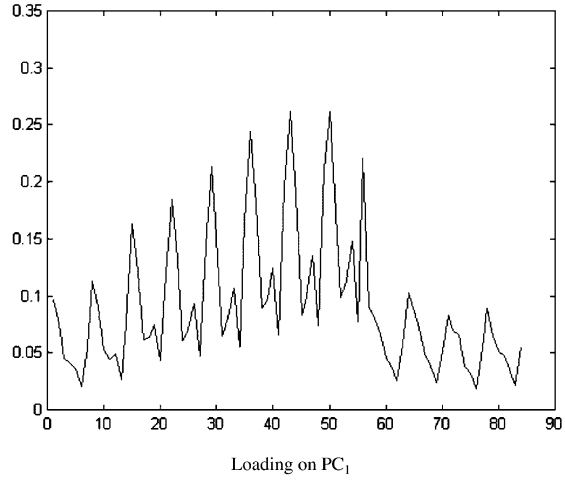

Fig. 7. Loadings on PC_1 for a (28×12×7) model (patient #11, 30-mg dose, resting condition).

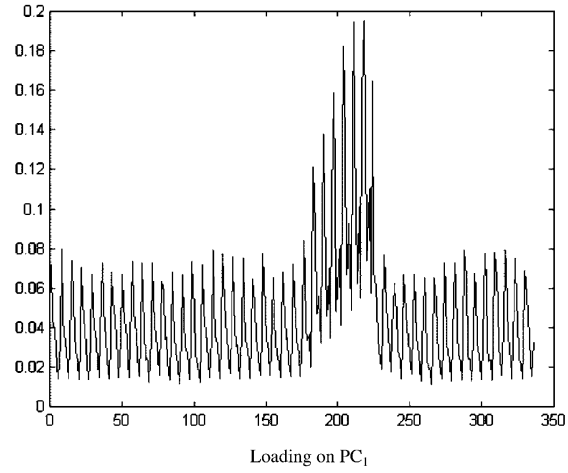

Fig. 8. Loadings on PC_1 for a $(28 \times 7 \times 12 \times 4)$ model (patient #11, resting condition).

tion about the electrodes and time dimension. The plot shows 12 repetitions (one per time) of the 28 electrodes. It can be observed that the loading values of the electrodes once again globally follow a time profile, indicating that there is some multi-way structure relating these two modes. Considering only the part of the data set corresponding to patient #11 and the resting condition leads to a four-way array with dimension $(28 \times 7 \times 12 \times 4)$. The loadings of the PCA model built on this array unfolded preserving its first mode should give information about variables, time, and doses simultaneously. A structure due to the dose dimension is visible (Fig. 8). Dose 3 (30 mg) seems to be standing out.

4.3. Multilinear models optimization

The PARAFAC model should preferably be used as its simplicity makes the interpretation of the results easier and also because of its uniqueness property. However, it has first to be investigated whether the data can be modelled with PARAFAC. This verification can be performed using the Core Consistency Diagnostic [8]. This approach is used to estimate the optimal complexity of a PARAFAC model (or any other model that can be considered as a restricted Tucker 3 model). It can be seen as building a Tucker 3 model with the same complexity as the PARAFAC model and with unconstrained components and analysing its core. In practice, the core consistency diagnostic is performed by calculating a Tucker 3 core matrix from the loading matrices of the PARAFAC model. If the PARAFAC model is valid and optimal in terms of complexity, the core matrix of this Tucker 3 model, after rotation to optimal diagonality, should contain no significant non-diagonal element. The data was first restricted to simpler three-way cases, and three-way PARAFAC models were built. For instance, in the case of models built for data restricted to one patient, one condition, and

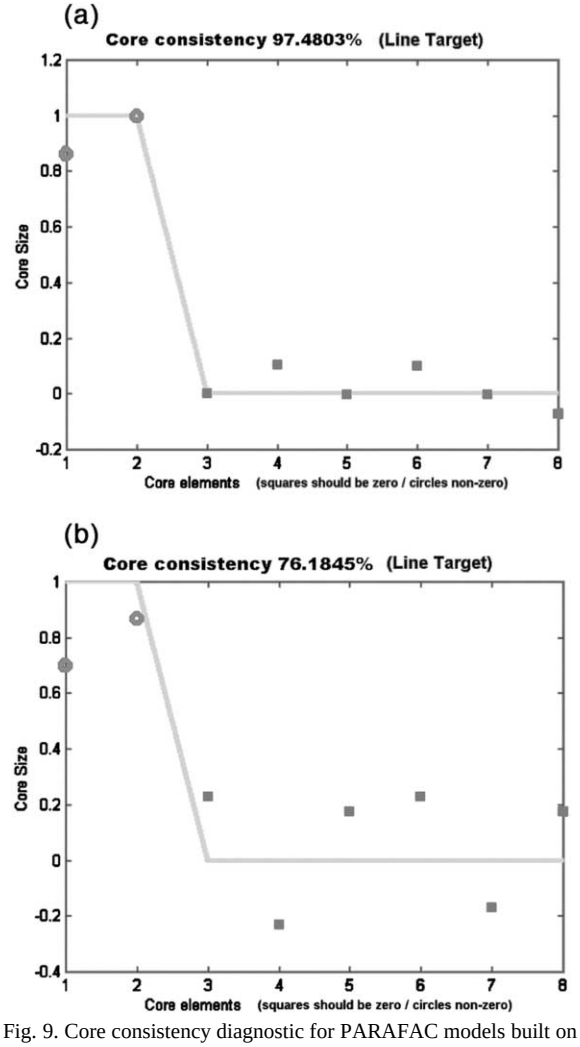

Fig. 9. Core consistency diagnostic for PARAFAC models built on three-way data. (a) Patient 6, resting condition, 30-mg dose. (b) Patient 11, resting condition, 30-mg dose.

one dose, the dimensions modelled are the spatial dimension (position of the electrodes), the time dimension, and the variables dimension. In all cases studied here, a two-component PARAFAC model was always optimal. However, the performances of the PARAFAC models depended greatly on the patient studied. For patient #6, for instance (Fig. 9a), the model is much better than for patient #11 (Fig. 9b). This indicates that the data do not seem to follow a PARAFAC model, or at least the modelling is not easy, the data can therefore not be fit adequately by this model. By increasing the number of dimensions modelled, it was verified that a PARAFAC model is probably not appropriate for this data set. In order to assess the validity of the PARAFAC model on a data set, it is also useful to estimate the fit of both Tucker 3 and PARAFAC models in order to evaluate if the larger flexibility of the Tucker model leads to a significant improvement in the fit. The fit of the two components PARAFAC model and the (222222) six-way Tucker 3 model (two components extracted on each of the six modes) is actually almost identical (around 93.5% of explained variance). However, this complexity does not seem to be optimal at all in the case of the six-way Tucker 3 model. In order to keep computation time reasonable, the optimal complexity of the six-way Tucker 3 model was evaluated (Fig. 10) taking into account only a number of components quite close to 2. The complexity was therefore investigated only from (111111) to (333333). It appeared that the optimal complexity is (333221), which can be detailed as follows :

EEG dimension	Three components
Subject dimension	Three components
Spatial dimension	Three components
Dose dimension	Two components
Time dimension	Two components
Condition dimension	One component

This complexity corresponds to the beginning of the last plateau on the curve (more exactly in this case on a part of the curve just after a significant reduction of the slope). The model is on purpose not chosen to be parsimonious as it would for instance have been possible to select the complexity corresponding to the beginning of the plateau containing the (222222) model. It is however always possible to discard some components from the model if it appears from the interpretation of the core that they are not useful in the reconstruction of the original matrix **X**.

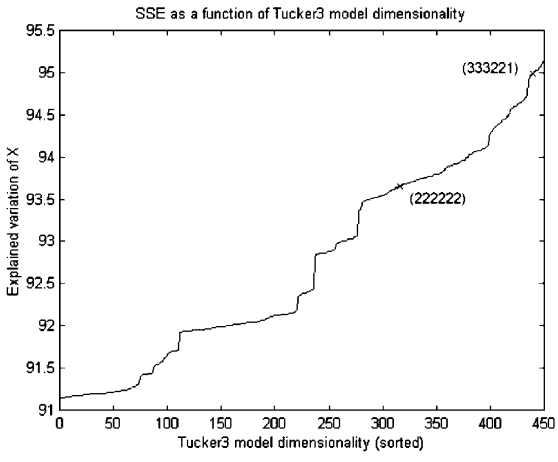

Fig. 10. Variance explained by the Tucker 3 models as a function of the model complexity.

4.4. Six-way tucker model

The six-way Tucker 3 model leads to a core array **G** with dimensions $(3 \times 3 \times 3 \times 2 \times 2 \times 1)$ and six component matrices **A**, **B**, **C**, **D**, **E**, and **F** related each to one of the modes.

4.4.1. Loadings on the variable dimension

The first matrix **A** holds the loadings for the EEG dimension (seven EEG bands). By calculating from the original data the average energy (over the five other modes) of each frequency band, it can be seen (Fig. 11a) that the first component is used to describe the average energy of the bands. The second component, as well as the third one (Fig. 11b), will at this stage be interpreted as showing the effect of some other parameters (time or effect of the substance) on the distribution of the bands.

4.4.2. Loadings on the patient dimension

The second matrix **B** holds the loadings for the patients dimension (12 patients). The main information in the loading plots is that some extreme values are present. Patient #6 appears as an extreme value

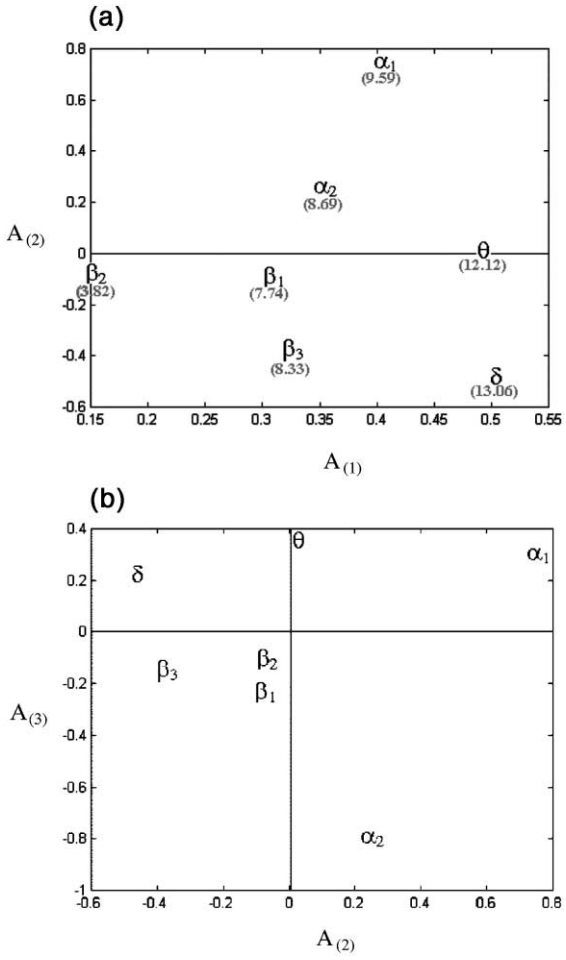

Fig. 11. Loadings on the variable dimension, six-way model with complexity (3 3 3 2 2 1). (a) $A_{(1)}$ versus $A_{(2)}$. The mean energies of the bands are also given. (b) $A_{(2)}$ versus $A_{(3)}$.

on component 1 (Fig. 12a). Patient #11 appears as an outlier on component 3 (Fig. 12b). At this stage, without looking at the core array **G** in order to remove the rotational indeterminacy of the Tucker 3 model, it is not possible to go further in the discussion about this matrix.

4.4.3. Loadings on the spatial dimension

The third matrix **C** holds the loadings for the spatial dimension (28 electrodes). The first remarkable thing in the plot of $C_{(1)}$ versus $C_{(2)}$ is the symmetry of the loadings (Fig. 13a). All electrodes that are symmetrical on the brain (Fig. 1), for instance electrodes #17 and 20 appear very close to each other on the loading plot. Moreover, considering all these pairs of symmetrical electrodes, the one located on the right part of the brain appears to have systematically higher loading values. For instance, electrode #20 has higher loadings values than electrode #17. This rule holds for all of the pairs of electrodes, except for electrodes #12 and 16. It will be established when interpreting the core matrix that this is due to a specific problem with one of these leads for one of the patients. If the loading values on component 1 are reported on the map of the electrodes on the brain, a representation of the activity of the brain is obtained (Fig. 13b), it looks very similar to what was obtained

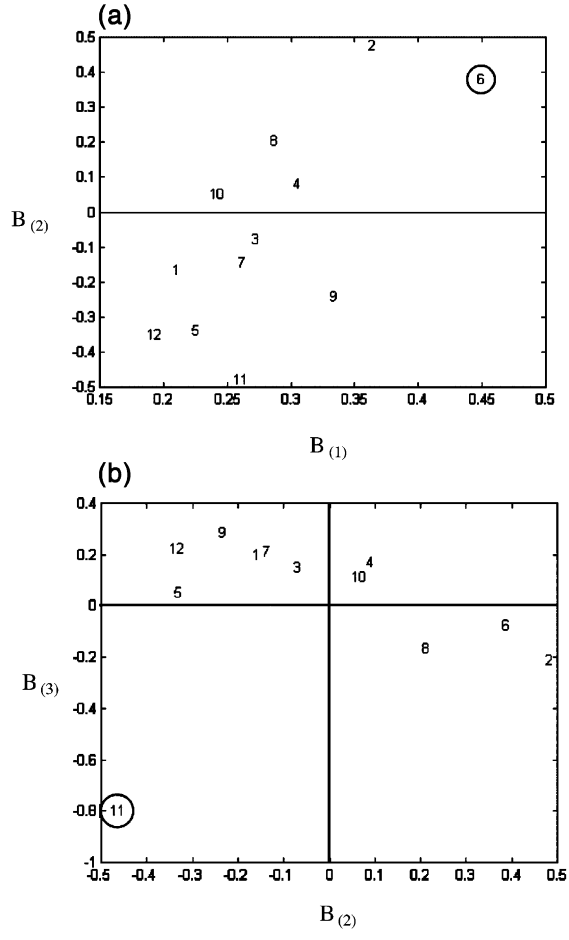

Fig. 12. Loadings on the patient dimension, six-way model with complexity (3 3 3 2 2 1). (a) $B_{(1)}$ versus $B_{(2)}$. (b) $B_{(2)}$ versus $B_{(3)}$.

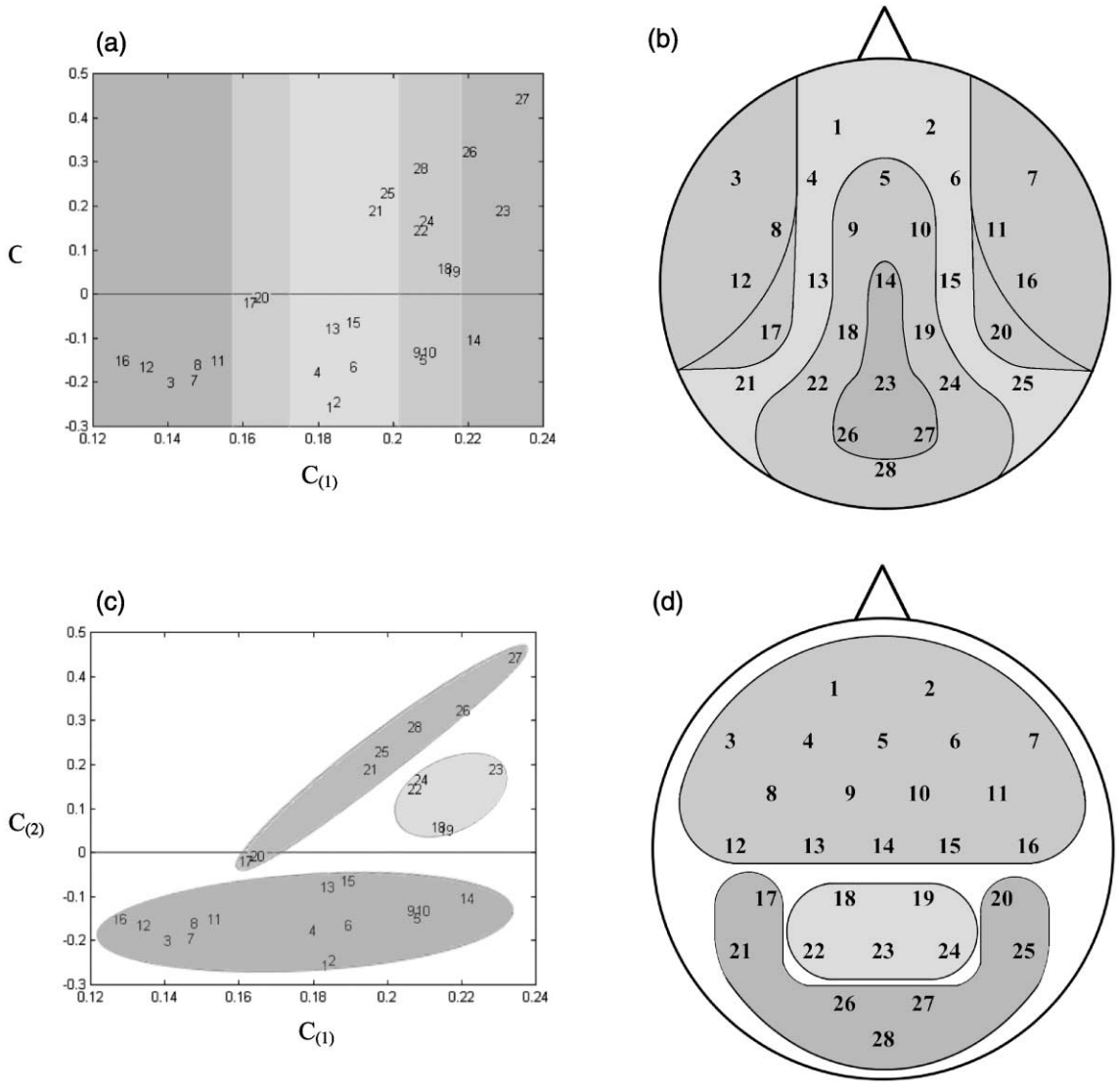

Fig. 13. Loadings on the spatial dimension, six-way model with complexity (3 3 3 2 2 1). (a) $C_{(1)}$ versus $C_{(2)}$. (b) Ranking of the electrodes on $C_{(1)}$ reported on the map of the brain. (c) $C_{(1)}$ versus $C_{(2)}$. (d) Patterns on the loading plots are reported on the map of the brain.

with linear models in the data exploration part (Fig. 5).

If the second component of the **C** matrix is now considered (Fig. 13c), and the loading values are reported on the map of the electrodes on the brain, a clear separation between the front and back part of the brain can be observed (Fig 13d). Considering directions in the plots, a central part of the brain can be identified. These patterns are interpreted as showing the activity of the substance on different parts of the brain.

It is important to note that, at this stage, only with the information present in the loading matrices, it is not possible to know whether the high loadings on $C_{(1)}$ for the central part of the brain mean high or low activity. A basic knowledge of brain physiology indicates that this indeed corresponds to high activity. It is however necessary to get rid of the rotational inde-

terminacy of the Tucker 3 model by interpreting the core matrix to extract this information from the model.

4.4.4. Loadings on the dose dimension

The first component on the dose dimension $D_{(1)}$ can be interpreted quite easily (Fig. 14). It shows that 10 mg is quite close to placebo, indicating that this dose is not efficient. A 90-mg is more different compared to placebo indicating a better effect of this dose, and the most different is 30 mg. This can appear surprising, but the medical doctors in charge of the study expected this result. The higher dose does not systematically lead to the higher effect with this kind of substances. The second dimension, making a difference between 30 mg and the other doses, is much more difficult to interpret at this stage, but the phenomenon will be explained when interpreting the core matrix **G**.

4.4.5. Loadings on the time dimension

The first component on the time dimension $E_{(1)}$ shows the normal time profile of the evolution of the state of the brain during day time (Fig. 15). The activity globally increases from 8:00 AM (time 1) until 11:00 AM (time 6). This would of course still have to be confirmed by removing the rotational indeterminacy using **G**, but it already fits what was seen in the linear data exploration part (Fig. 5). Afterwards, the activity reduces. The loading value for the second day at 9:00 AM (point 11) is located between the ones corresponding to 8:30 and 9:30 AM in the first day, confirming this interpretation. The second dimension is interpreted as showing the time profile of the effect of the drug activity. It has to be specified that the drug was administered immediately after 8:30 AM (time 2). No effect of the substance can therefore be expected before 9:30 AM (time 3). The loadings on component 2 are indeed negative before 8:30 AM and become positive from 9:30 AM, regularly increasing until 11:30–12:00 AM. After 12:00 AM, the activity drops and becomes zero (no activity, same negative loading values as before the administration of the drug), and stays at this level during the second day.

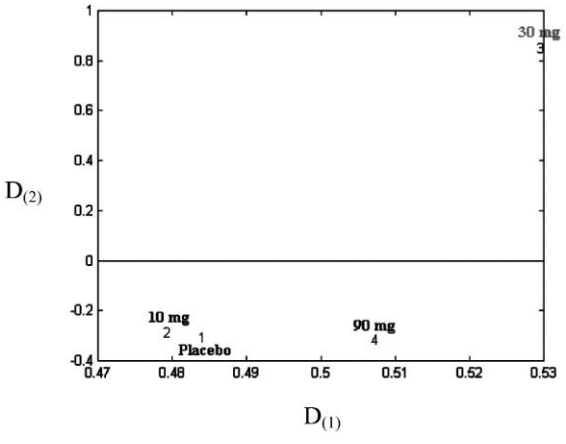

Fig. 14. Loadings on the dose dimension, six-way model with complexity (3 3 3 2 2 1). $D_{(1)}$ versus $D_{(2)}$.

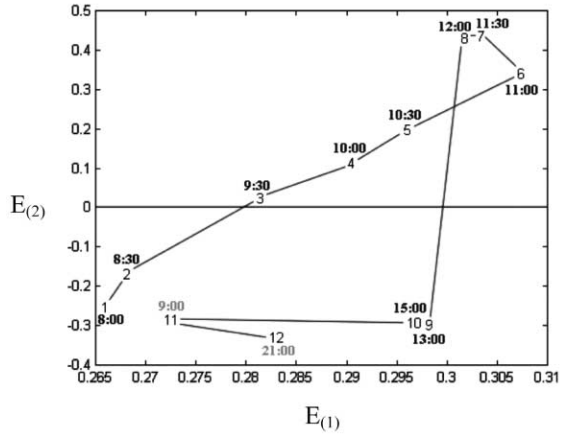

Fig. 15. Loadings on the time dimension, six-way model with complexity (3 3 3 2 2 1). $E_{(1)}$ versus $E_{(2)}$.

4.4.6. Loadings on the condition dimension

The last component matrix **F** gives information about the two different measurement conditions. It is in fact a vector as only one component was extracted along this mode. The loadings values are 0.701 for the resting condition, and 0.713 for the vigilance-controlled condition. The loadings are positive for both the conditions, this indicates that when interpreting the model, this mode can only have a scale effect. This means that the effect of the drug can only be larger or smaller depending on the condition, but one cannot expect to see opposed effects due to this parameter. The loading values for each condition are

also very similar. This indicates that the two conditions do not imply any effect on the brain activity that is significant for the model. This dimension was further investigated. Five-way models were built on data taking into account one of the conditions, the other condition, and the average of the data in the two conditions. All these models gave almost perfectly identical results, showing that the two conditions can in fact be considered as replicates of the same five-way data set. This mode is therefore not relevant in the data set.

4.4.7. The core matrix ***G***

The important elements of the core are shown in Table 1, together with their squared value (that represents the relative importance of the core element), and the variance explained by these elements. By building symbolic products as described by Henrion [9], it is possible to go over the rotational indeterminacy of the model and interpret the first elements of the core. The first element of the core explains most of the variance and reflects the normal evolution of the activity of a human brain during daytime, showing which bands are the most present, and how their intensity evolves in time. Even if the corresponding core values are very low (which is not surprising as phenomena with very small magnitude are investigated, compared to, for instance, the difference between two patients), the next elements also bring very relevant information. One of the most interesting elements in this core matrix is element #4. It shows that $B_{(3)}$, third component on the patient mode and $D_{(2)}$, second component on the dose mode interact. It can be reminded that $B_{(3)}$ differentiates between patient #11 and the other patients, spotting him as an outlier (Fig. 12b). It was also seen that $D_{(2)}$ differentiates between the 30-mg dose and the other doses (Fig. 14). This core element shows that patient #11 is an outlier due to an over-reaction to the most efficient dose. This interpretation was confirmed studying a five-way model restricted to patient #11. In this model, the 30-mg dose appeared to be even more extreme than on the six-way model. In the same way, mainly starting from the loading plots of the patient dimension, and looking for extreme points, it was possible to find core elements explaining very small amounts of the total variance of the system, but representative for special behaviours of specific patients. Core element #7, for instance, relates $B_{(1)}$, the first component on the patients mode (showing patient #6 as an outlier), to $A_{(3)}$, third component on the EEG mode (differentiating α_2 from the other energy bands). This core element accounts for a specific repartition of the energy bands for patient #6. This was confirmed by investigating a five-way model restricted to this patient. On this model, the distribution of energy bands showed in particular extremely high values of the α bands. The special behaviour of electrode #12 compared to its symmetrical on Fig. 13a can be explained by focusing on patient #9. All measurements on this patient have an extreme value for electrode #12. This was confirmed by studying a five-way model restricted to this patient, clearly differentiating electrode #12 from the others. It can be seen that the energy values for this electrode are wrong, the high-energy bands (especially β_2 and β_3) are strongly over-estimated. This happens for all the measurements performed on this patient for the 90-mg dose (which also corresponds to a certain period in time, as the doses are tested successively with a 'wash-out' period between each dose). This systematic and very localised problem seems to indicate that the corresponding electrode was either damaged or badly installed on the scalp during this part of the data acquisition.

Table 1
Important core elements of the six-way model with complexity (3 3 3 2 2 1)

	Core element	Explained variance (%)	Core value	Squared core value
1	(1, 1, 1, 1, 1, 1)	95.95	4702.23	22111057.12
2	(2, 2, 1, 1, 1, 1)	1.63	613.57	376475.69
3	(2, 1, 2, 1, 1, 1)	0.60	374.71	140414.91
4	(1, 3, 1, 2, 1, 1)	0.39	−301.13	90682.65
5	(1, 3, 1, 2, 2, 1)	0.23	−234.06	54788.31
6	(3, 3, 1, 1, 1, 1)	0.20	−214.79	46137.64
7	(3, 1, 2, 1, 1, 1)	0.18	−208.41	43438.22
8	(1, 2, 2, 1, 1, 1)	0.09	150.35	22605.75
9	(1, 2, 1, 2, 2, 1)	0.09	−146.34	21417.87
10	(1, 3, 1, 1, 2, 1)	0.08	−137.65	18947.96

4.5. Analyzing subject variability

Since many of the core elements seemed to be used only to account for specific behaviours of individual

patients, it was decided to study more thoroughly the patients mode. The idea was to simplify the problem by removing the non-typical patients. This way, the number of relevant core elements should be reduced, as well as the optimal complexity of the model. For this purpose, it was decided to center the patients mode in order to highlight the differences between patients, and hopefully identify easily the suspected outliers. Moreover, as it was shown to be not relevant, the 6th dimension (related to the two measurement conditions) was collapsed. The average of the two conditions was used leading to a simpler five-way array.

The plots of the loading matrix **B** obtained for the patient dimension show that outliers already spotted with the six-way model appear now much more clearly (Fig. 16a,b). Patient #11 appears as an outlier already on component 2, while patient #6 (and also perhaps #2) is extreme on component 1, and patient #9 seems to be atypical on component 3. This shows that the centering of this mode succeeded in enhancing the differences between patients.

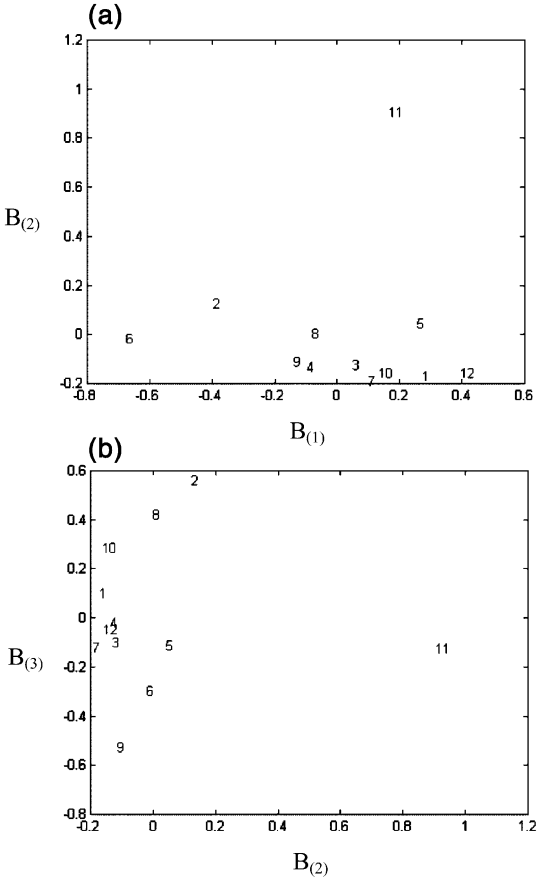

Fig. 16. Loadings on the patient dimension, five-way model with complexity (3 3 3 2 2). (a) $B_{(1)}$ versus $B_{(2)}$. (b) $B_{(2)}$ versus $B_{(3)}$.

Table 2
Important core elements of the five-way model with complexity (3 3 3 2 2 1)

	Core element	Explained variance (%)	Core value	Squared core value
1	(1, 1, 1, 1, 1)	74.66	−879.87	774171.86
2	(2, 3, 1, 1, 1)	8.69	300.26	90160.43
3	(1, 2, 1, 2, 1)	3.34	186.20	34671.54
4	(2, 2, 1, 2, 1)	1.99	−143.95	20723.73
5	(1, 2, 1, 2, 2)	1.75	134.90	18200.44
6	(3, 2, 1, 1, 1)	1.70	132.82	17641.84
7	(2, 2, 1, 2, 2)	1.37	−119.46	14272.36
8	(1, 2, 2, 2, 1)	0.82	−92.30	8520.95
9	(1, 2, 1, 1, 2)	0.78	90.08	8114.57
10	(2, 2, 1, 1, 2)	0.74	−87.65	7683.06

The core matrix also gave interesting information (Table 2). First, as an important source of variance was previously reduced by centering the data, the total variance explained by the five-way Tucker 3 model was, as could be expected, much smaller (from 94.9% for the six-way model to 68.8% for the five-way model centered on the patients dimension). The explained variance is also much more distributed between the core elements, which is logical as the variance of the system is less dominated by the differences between patients. It is also obvious that the complexity of the model could be very much reduced. This is especially true for the spatial (3rd) and time (5th) modes where two components might suffice.

5. Conclusion

Multi-way models, in particular Tucker 3, were used on data with a high number of modes. It was shown that this multi-way model was able to extract meaningful information from this very complex data set, when classical PCA brought no usable information. Each mode could be interpreted and the core matrix enabled to understand relations between

modes. Since it was established that some atypical patients made the modelling and the interpretation of the results much more complicated, the second part of this study, aiming at interpreting the anatomical results of the models in details, will be performed having these patients removed from the data set. Since some major sources of variance will be removed from the data, the optimal complexity of the models will have to be investigated in details again. Another interesting point is that the performances of the PARAFAC model seemed to depend very much on the behaviour of the patients, it will therefore be interesting to evaluate the modelling abilities of this model on the simplified data set. The results of this second part of the study will be presented in a forthcoming publication. It is anyhow already possible to say that the optimal complexities of the models established on the simplified data set are indeed much lower. The simplified data set also happens to conform much better to a PARAFAC model. This model can therefore be used, which will hopefully enable an easier interpretation of the results.

References

[1] M.J. Aminoff, Electrodiagnosis in Clinical Neurology, third edition, Churchill Livingstone, Edinburgh, 1987.
[2] H.H. Jasper, Report of the committee on methods of clinical examination in electroencephalography, Electroencephalogr. Clin. Neurophysiol. 10 (1958) 370.
[3] C.A. Andersson, R. Bro, Chemom. Intell. Lab. Syst. 52 (2000) 1–4.
[4] L.R. Tucker, Psychometrika 31 (1966) 279–311.
[5] P.M. Kroonenberg, Three-mode Principal Component Analysis, Theory and Applications, DSWO Press, Leiden, 1983.
[6] R. Harshman, UCLA working papers in phonetics 16 (1970) 1–84.
[7] J.D. Carrol, J. Chang, Psychometrika 45 (1970) 283–319.
[8] R. Bro, H.A.L. Kiers, J. Chemom. (2000) in press.
[9] H. Henrion, Chemom. Intell. Lab. Syst. 25 (1994) 1–23.