Statistical modelling of neural networks in gamma-spectrometry applications

Vincent Vigneron, Jean Morel, Marie-Christine Lépy, Jean-Marc Martinez

Ref. : DMT/SERMA/LETR 95/175

Direction des Réacteurs Nucléaires
Département de Mécanique & de Technologie

950c

DEMANDE D'AUTORISATION DE COMMUNICATION OU DE PUBLICATION

PUBLICATION - TITRE :

Auteur (s) : V
Nature de la Publication :
Editeurs (CEA, AIEA, OCDE,...) : NT 6081 N. FR 9601559

COMMUNICATION - TITRE : STATISTICAL MODELLING OF NEURAL NETWORKS IN SPECTROMETRY APPLICATIO
Auteur (s) : VIGNERON V., J.M. MARTINEZ, S.MOREL, M.C.LÉTY

Titre de la Conférence : ICRM'95

Lieu et Date : CONFERENCE AND INTERNATIONAL SYMPOSIUM ON RADIONUCLEIDE METROLOGY AND ITS APPLICATIONS.
15-19 Mai 1995 - Paris

ORGANISATEURS : CEA - BNM

12326

Date de remise des textes : 20 mars 1995.
Commentaires : Travail en cooperation avec DAMRI

Ce texte a-t-il été déjà publié ? : Non.
Si OUI : Référence de la publication antérieure :
Cette publication contient-elle, à votre avis, des informations brevetables OUI (NON)

DATE DE LA DEMANDE : 28.03.95.

Chef de Laboratoire : J.M. EVRARD

Chef de Service : M. Proust.

)EMANDE D'AVIS (éventuellement) :

) - C.P.I./Saclay	**OUI NON**
) - Chargé de Mission pour les affaires industrielles	**OUI NON**
) - SYFRA	**OUI NON**
) - Partenaires concernés (EDF, FRA, Dpts de la DRN) : (joindre photocopie)	**Date de la Demande :**
- Autres unités opérationnelles : (joindre photocopie)	**Date de la Demande :**
- Autres avis demandés par le Chef du D.M.T. :	

vis du : **(le cas échéant)**

)ate :
:s avis sont à envoyer au DMT/DIR

ate d' Arrivée au D.M.T. 25-4-95 | **Décision D.M.T. n°** 261 Accord | **Date** 25-4-95

Le Chef du Département

ARRIVÉE - CIRST
02 MAI 98 | 003007

: Un texte complet -
ie : ~~DMT/SDC~~
DMT/~~SA~~
: Copie autorisation + Résumé + Texte à I.N.S.T.N./MIST/~~CIRST~~.

We regret that some of the pages in this report may not be up to the proper legibility standards, even though the best possible copy was used for scanning

STATISTICAL MODELLING OF NEURAL NETWORKS IN γ-SPECTROMETRY APPLICATIONS

V. Vigneron[1,3], J. Morel[2], M.C. Lépy[2], J.M. Martinez[1]

[1]CEA Saclay	[3]CREL	[2]CEA Saclay
DRN/DMT/SERMA	161, rue de versailles	DAMRI/LPRI, BP 52
91191 Gif-sur-Yvette	78180 Le Chesnay	91193 Gif-sur-Yvette cedex
FRANCE	FRANCE	FRANCE
vvigne@soleil.serma.cea.fr	fuche@crel.fdn.org	LEPY@BABAORUM.cea.fr
jmm@soleil.serma.cea.fr		

Abstract

Layered Neural Networks, which are a class of models based on neural computation , are applied to the measurement of uranium enrichment, i.e. the isotope ratio $\frac{^{235}U}{^{235}U+^{236}U+^{238}U}$. *The usual methods considere a limited number of γ-ray and X-ray peaks, and require previously calibrated instrumentation for each sample. But, in practice, the source-detector ensemble geometry conditions are critically different, thus a means of improving the above conventional methods is to reduce the region of interest: this is possible by focusing on the $K_\alpha X$ region where the three elementary components are present. The measurement of these components in mixtures leads to the desired ratio. Real data are used to study the performance of neural networks. Training is done with a Maximum Likelihood method. We show the encoding of data by Neural Networks is a promising method to measure uranium ^{235}U and ^{238}U quantities in infinitely thick samples.*

Key words NEURAL NETWORKS, MACHINE LEARNING, URANIUM ENRICHMENT, GENERALIZED LINEAR MODELS.

1 Introduction

In the past few years, the topic of neural computing has generated widespread interest and popularity. The popularity of this technique is due in part to the analogy between Artificial

Neural Networks (ANNs) and biological neural networks. Numerous applications have been investigated using ANNs, we demonstrate how they can be used to model spectra of mixtures to produce quantitative estimates of the concentrations of the components in the mixtures. Indeed, with modern sensor and high technology, spectral data are collected with ever finer sampling and with ever large precision. This leads to a perpetual need for more efficient methods of interpreting spectral data, namely for uranium enrichment measurement. The objective of uranium enrichment measurement methods is to determine the $\frac{^{235}U}{^{235}U+^{236}U+^{238}U}$ isotopic ratio.

If these are non-destructive methods, the determination uses several X-ray and γ-ray peaks, mainly in the 60 to 200 keV region. These methods, which were developed more than 20 years ago, are based on measurements of the full energy peak at 185,7 keV ([1], [2], [3], [4]).

They require a prior calibration of the system and imply that the measurement conditions are constant. Frequently, the matrix effects are avoided by measuring only samples said to be infinitely thick : these are samples whose thickness is such that any further increase has no effect on the fraction of the 185,7 keV γ-ray emitted. Other methods have been developed using several γ-ray peaks [5], [6], [7]. In fact, these latter methods requires a self-calibration with a limited number of peaks, making them difficult to implement.

A means of overcoming calibration-related difficulties is to reduce the region of interest of the spectrum so that the variation in the detector efficiency is limited. This is possible by considering only the relatively complex $K_\alpha X$ region, which extends from 83 to 103 keV where numerous γ-ray (Gaussian distribution) and X-ray (Voigt profile) peaks are superimposed.

The processing of this region requires taking account of 3 elemental images corresponding to the presence of ^{235}U and its daughters, to ^{238}U and to X-ray fluorescence. These images are now represented by mathematical expressions taking account of the shapes of the X-ray and γ-ray peaks, their energies, intensities and a slight variation in detector efficiency. The determination is then carried out conventionally with the least squares method, this is the case of the MGA-U code [8]. This approach requires that all the parameters for constituting each elemental image are well-known and is based on the use of external data contained in the spectrum for the region considered.

It is precisely in this context that a Neural Network appears to be a useful tool for the characterization of so-called infinitely thick samples. In fact the training by ANNs can be considered as a search procedure for an "optimum" regression function among a set of acceptable functions using a set of training examples. From the statisticians point of view (Ripley, 1992), ANNs belong to evaluation techniques for non-parametric models, still called *tabula rasa* [1].

ANNs, like most statistical methods, are capable of processing vast amounts of data and making predictions that are sometimes surprisingly accurate : this does not make them

[1] In other words, the set of methods developed without providing a "true " model, as opposed to parametric models where the response to predictive values is not known. This point of view is in contradiction with that of numerous connectionists.

intelligent in the usual sense of the word. ANNs learn in much the same way that many statistical algorithms do estimation. But in contrast to usual automatic spectra analysis methods, ANNs use full parallel computing, are simple to implement, not very sensitive to outliers and contain nonlinearities.

In the following, we describe the identification method based on neural networks to quantify uranium quantities. Section II covers the experimental procedure and neural networks technique is explained in section III. Finally, Section IV gives the outlook and conclusion.

2 Experimental Aspect

2.1 Preliminaries

From the similarity to plutonium isotope measurement procedures, it should be possible to use all the spectral region to determine the enrichment. To do this, it would be necessary to establish, relatively, the detector efficiency for the considered sample. This overall efficiency is the product of 4 factors:

- detector response for a point source,
- transfer of this response to the sample,
- attenuation by the material between the detector and the nuclear material,
- sample self-absorption.

This overall efficiency is critically dependent on the energy. It is a function of several parameters, some of which are not known, as, for example, the exact dimensions of the sample and its container and the precise composition of the material. Also for uranium, this response is to establish due to the insufficient number of peaks that can be used to define the efficiency. This difficulty is overcome by using only the $K_\alpha X$ region which extends from 83 to 100 keV (figure 2). This region contains enough information to allow the determination of ^{235}U and ^{238}U and is sufficiently limited that the efficiency can be, initially, considered as constant. It is however very complex to analyze due to several interfering X- and γ-rays. These can be grouped as follows :

- ^{235}U and daughters : 84,21 keV (γ^{231}Th), 89,95 keV (γ^{231}Th), ($ThK_{\alpha_2}X$), 92,28 keV ($PaK_{\alpha_2}X$), 93,35 keV ($ThK_{\alpha_1}X$), 95,86 keV ($PaK_{\alpha_1}X$)
- ^{238}U and daughters : 83,30 keV (γ^{234}Th), 92,28 keV ($PaK_{\alpha_2}X$), 92,38 keV (γ^{234}Th), 92,79 keV (γ^{234}Th), 94,65 keV ($UK_{\alpha_2}X$), 95,86 keV ($PaK_{\alpha_1}X$), 98,43 keV ($UK_{\alpha_1}X$), 99,85 keV (γ^{234}Pa)
- Uranium X-ray fluorescence : 94.65 keV ($UK_{\alpha_2}X$), 98,43 ($UK_{\alpha_1}X$).

In the standard approach, the processing of the considered region taking account of the 3 elemental images, the first corresponding to ^{235}U and his daughters, the second to that

of ^{238}U and its daughters and the third to the uranium X-ray fluoresence spectrum. The determination of these components in the complex spectrum leads to the $\frac{^{235}U}{^{235}U+^{238}U}$ (see table 1). The final enrichment is obtained by making allowance for the prescience of ^{236}U. a correction can thus be made by using the 120,9 keV.

The litterature concerned in calculation of uranium enrichment is vast and is expending rapidely. We found the texts by Kull & Gunaven [2], and Neuilly [9] of particular interest. They propose a method based on an ordinary linear regression with the enrichment E, as the dependent variable, but they do not give any evaluation of their results. Liggett [10] provides an informal and readable account of the possible future direction of the field : namely Maximum Likelihood Estimators. In all cases, analysis are carried out in respecting two limitations : a constant variance over the random error and the normality of these errors.

2.2 Samples measurement

Six uranium oxide standards with different enrichments and infinite thickness were counted several times by γ-ray spectrometry to test the neural procedure. These were bare cylindrical pellets, with certified enrichments and the following characteristics [3] :

The Ge(HP) planar detector used in the measurement system had the following characteristics 1: surface, 2.00 cm^2 ; thickness, 1.00 cm ; resolution, 190 eV at 6 keV and 480 eV at 122 keV. All the measurements were made under the same conditions, i.e. with 0,05 keV per channel and a source-detector window distance of 11 cm. Ten 20.000-seconds counts were made for each standard pellet and each of the obtained spectra was processed by our analysis procedure. The ^{234}U concentration is relatively low, however , a $\frac{^{234}U}{^{235}U}$ relative mass ratio varying from 0,5 to 1,1%, depending on the pellet was determined by γ-ray spectrometry by using both the 53,2 and 120,9 keV peaks for ^{234}U and the 185,7 keV peak for ^{235}U.

In short, 65 sets of experimental data from real-life calibration experiment were prepared using the concentrations illustrated in table (1), i.e. five ^{235}U-pure idealized spectra, and ten of each standard from 0,711 to 9,548%.

3 Layered Neural Network and Training method

3.1 Using Neural Networks

In this section, our aim is not the presentation of the neural network theory. Our purpose is to present the place of the connectionnist approach in γ-spectrometry problems. Most details and basic concepts are clearly described in a paper to be published [11]. ANN consists of a large number of neurons, i.e. simple linear or nonlinear computing elements, interconnected in complex ways and often organized into layers [12]. The collective or parallel behaviour of the network is determined by the way in which the nodes are connected and the relative type and strengh (excitory or inhibitory) of the interaction amongst them.

ANNs are used by engineers, physicists, neurophysiologists, or computer scientists in three main ways: as models of biological nervous systems and intelligence, as real-time adaptative signal processors or controllers for applications such as robots or nuclear plants [13] and as data analysis methods.

The objective of ANNs is to construct a suitable model which, when subjected to a ^{235}U enrichment spectrum, produces an output **y** which approximates the exact uranium enrichment ratio. The principal idea of the connextionist approach is to substitute a neural model and the learning procedure of the network for classical fitting algorithms. "Classical" solutions to the automatic spectrum analysis problem make use of complex mathematical algorithms, generally based on the separation of a given curve, associated to each individual peak plus a background.

An exemple of multi-layer network is given in figure 3. The notational convention is that the square represents a computational unit into which the x_j's are fed and multiplied by the respective ω_j's. The fundamental processing element of an ANN is a node (figure 3). Nodes are analogous to neurons in biological systems. Each node has a series of weighted inputs, ω_i which may be either external signal or output from other nodes. The inputs of the nodes are analogous to synapses, and the weights corresponds to the strengh of the synaptic connection. The sum of the weighted inputs is transformed with a linear or a non-linear transformation function. The most popular one is the sigmoïd function $f(x) = \frac{1}{1+e^{-x}}$. This function has an output in the range 0 to 1, where x is the weighted sum of the inputs. Other transfer functions have been investigated including the hyperbolic tangent, and simple linear functions. We create a representation of a standard Neural Network called *M*ulti-*L*ayered *P*erceptron (*MLP*), which is a very familiar statistical construct : the Multivariate Multiple Nonlinear Regression.

Transmission of information between units of two neighboring layers is performed through oriented links. These links are level-headed by connection weights. The essence of the construction is as follows :

- input layer : this layer contains input units. Each unit receives input-variables, selected through a free parameters reduction procedure.
- hidden layer : this layer acts as an array of feature detectors picking up features without regard to position. the information coming to the input units is recoded on the hidden layer into an internal representation Thus, the input-layer units contribute to the input of each second-layer unit. It is fully-connected to the output.
- output layer : it applies a sigmoïd activation fonction to the weighted sum of the hidden outputs.

The role of the hidden layer is fundamental. A network without hidden units will be unable to perform the necessary multi-input multi-output mappings, in particular with non-linear problems.

Input pattern can always be encoded, if there are enough hidden units, in a form so that the appropriate output pattern can be generated from the corresponding input pattern.

The training data are denoted by $\chi = (\mathbf{x}, \mathbf{y^d})_{t=1}^{N}$ where $\mathbf{x}$ is the feature vector corresponding to the t^{th} observation. The expected response $\mathbf{y^d} = (y_1, y_2, \ldots y_M)$ is related to the inputs $\mathbf{x} = (x_1, x_2, \ldots x_N)$ according to :

$$\mathbf{y} = \phi(\mathbf{x}, \omega), \tag{1}$$

where ω are called *connection weights.*

From the χ set of examples, the learning of the network consists in the modification of the synaptic weights in order to minimize an objective function in relation to the set of examples that will be presented to the network. Cost function represents the discrepancy between desired output, say $\mathbf{y^d}$, and predicted output $\mathbf{y}$ of the model, e.g. the square form:

$$J(\omega) = \frac{1}{2}\sum_{p=1}^{P}(\mathbf{y^d} - \mathbf{y})^t(\mathbf{y^d} - \mathbf{y}), \tag{2}$$

where p is the number of the units of the output layer, The Backpropagation (BP) training algorithm [14] compares the actual and desired output of the network and adjusts the weights of the network in order to reduce the bias according to the iterative step :

$$\omega_j \leftarrow \omega_j + \Delta\omega_j, \tag{3}$$

where

$$\Delta\omega_j = \eta.\frac{\partial J}{\partial \omega_j} = \eta.\delta.x_j. \tag{4}$$

In equation (4), $\eta(> 0)$ is called the *learning rate*, and the learning rule the *Generalized Delta rule*. Clearly, δ denotes the vector of sensitivities at layer k of the network.

The practical difference between this device and the statistical version lies in the way the training data are used to dictate the values for ω. It turns out that there are 2 mains aspects to the processing : (1) specifying the architecture of a suitable network, (2) training the network to perform well with reference to a training set. Clearly, the connectionist and the statistician approaches differ in the way to handle with (1) and (2). The connectionist will resolve (1) by constructing a network of nodes and links from which a regression function can be written down, whereas the statistician will usually use general techniques as Maximum Likelihood Estimation (MLE), Mallow's C_p, Bayesian inference, ...

3.2 Training of the ANN

To check that this method was general and reliable, we have applied it to 65 sets of experimental data from real-life calibration experiment : five ^{235}U-pure idealized spectra, and ten of each standard precited (see table 1). Each spectrum contains 4096 points. The enrichment values were discretly distributed from 0 to 9.548 %. The computation of the spectra

are compared on two regressions models : **MLP MODELE** (depicted in figure 3), where the input are spectral data and **MIXTURES OF EXPERTS MODELE** (figure 4) where the input are the enrichment values. The specifications for the networks created for the calibration of the simulated data are listed in table 3. They were found to be optimal according the rigourous methodology describe in [11], for low prediction bias.

The choice of the right architecture is mainly intuitive and implies arbitrary decisions. But an attempt to apply ANN directly falls victim to the curse of dimensionability. Ockham's razor applies to supervised learning and ANN. This (parcimony) principle states that "*what can be done with less is done in vain with more*". In acccordance with this, the dimension of the input vector has been reduced dramatically by Principle Components Analysis, leading to the adequat reduction of weights emerging from the first layer of the ANN. The aim is to embrace the maximum amount of useful information in a small number of principal component scores (taken in the $K_\alpha X$ region) without overly reducing the computational power of the network[2].

Each input variable should be sifted so that its mean (averaged over the training set) is close to 0 because *correlations* between input variables introduce "preferred directions for weight changes" [16]. Decorrelation was performed by a PCA (Karhunen-Loeve expansion). This transformation preserves the form of the data base.

The **MLP MODELE** depicted in figure 3), consists of an input layer of 6 or 3 units leading up through one layer of hidden units to an output layer of a single unit that corresponds to the desired enrichment. This network represents a poor parametrized model, but the training dataset $(\mathbf{x}; \mathbf{y}^{(\mathbf{d})}, t = 1 \ldots 65)$ was small. The network is initialized with random weights and trained. For each pattern, the bias (2) is evaluated. This quantity decreases rapidly (figure 5.a) in the beginning. The training is stopped when the network reaches a minimum error on the training set, because this is an efficient way to avoid overfitting. After 32.000 successful training passes, the bias rate range from -0.05 to 0.04% for the 6-3-1 net (from -0.031 to 0.061% for the 3-5-1 net).

The aim of the second proposition is to attenuate these bias fluctuations.

In the case of ***MIXTURES OF EXPERTS MODELE (MEX)*[17]**, each item is associated with a vector of measurable features, and a target $\mathbf{y}^{\mathbf{d}}$ which represents the enrichment. The network receives the input $\mathbf{x}$ and creates the output vector $\mathbf{y}$ as a predicted value of the unknown $\mathbf{y}^{\mathbf{d}}$. This model consists of 210 independant fully-connected networks (figure 4) : One expert is put for one channel of the $K_\alpha X$ region.

Each expert is an observator, trying to find a "signal" due to radioactive decay in a large amount of noise, the variance of each count being proportional to the level and thus depending on the enrichment of a particular sample and on the background level of the

[2] The connectionist must deal with overfitting effects : indeed, when the net is overparametrized, convergence makes no sense : the net has learned the function and the interferences, he will tend to generalize poorly. In this discrimination, the supressed variables reduce the $\frac{signal}{noise}$ ratio [15].

particular observation. A cooperation-competition procedure driven by a supervisor between the expert's outputs leads to the choice of the most appropriate concentration. The basic idea is that competition leads to specialization.

Let $\mathbf{y_1}, \mathbf{y_2}, \ldots$ denote the output vectors of the experts, and $g_1, g_2, \ldots$ the supervisor output units, then the output of the entire architecture, $\mathbf{y}$, is $\mathbf{y} = \sum_{i=1}^{210} g_i \mathbf{y_i}$. The supervisor decides whether expert i is currently applicable or nor. The winning expert is the network with the smallest bias $(\mathbf{y^d} - \mathbf{y_i})$. The learning algorithm describe in [17] is inspired from an other paradigm called the *Maximum Likelihood Method.*

3.3 Discussion of the Results using ANN

As the initial base included only 65 examples, we wanted to keep a maximum of examples for the training base.

Redundances in the data-set enrichments present one main advantage : as we measure more than one response for each case, information from all the measured responses can be combined to provide more precise parameter estimation and to determine a more realistic models.

In all simulations, the measure of the system's performance is the Mean Square Error. The bias rates are compared in table 3 and on figure 5.a and 6. The figure 5.a shows the learning curves (i.e. the learning performances) for the two MLP networks using BP random training procedure. The horizontal axis gives the number of epochs ; the vertical axis gives the Mean Square Errors value (MSE). Clearly, the 6-3-1 network learned significantly faster than the 3-5-1. This difference can be explained by the information gain of the six-inputs network vs the 3-inputs network.

Figure 6 compares the results of the three models. The bias between the predicted and the desired enrichments is plotted for each of the 65 samples. The darkest line is put for the MEX. The results suggest that the strong dispersion of the bias with MLP is significantly attenuated when MEX is applied. This judgement must be moderated for the 6,122-enrichment-ratio samples.

The figure 5.b is concerned with the Multi-Expert model. The plotted points are predicted enrichment value (one for each of the 210 experts) when a $5,785\% - {}^{235}U$ spectrum is presented to the MEX model. The credit assignement procedure on these 210 contributions is supervised to produce a final estimation. In the most right column of table (3) can be seen the final predicted values of the simulations with MEX. Compared with the MLPs, this shows that MEX method is really reliable : for example the bias between the predicted and the calculated 2,785% enrichments range from 2,784 to 2,790%. As noted above, after 32.000 successful training passes, the larger bias happen for 5,111 and 6,122% enrichments. The relative lack of precision can be ascribed to the small size of the training dataset.

A comparison of the absolute bias curves suggest that of the three system studied, the Mixtures of Experts is capable of showing the most robust performance.

In fact, the modular approach presents three main advantages on the MLP models: it is able to model behavior, it learns faster than a global model and representation is easier to interpret : the modular architecture takes advantage of task decomposition, but the learner must decide which variables to allocate to the networks.

This method is at the same time very general and very specific. It is very general in the sense that no hypothesis is made on the aspect of the spectra : it does not depend if the spectra are well resolved or not, if they are very likely or not, if you select most significative areas of spectrum only (MLP models) or a global part of the spectrum (MEX model). No physical model is required whereas "classical" procedures may use three Lorentzians for example. But, at the same time, the method is very specific because the ANN must learn representative spectra of the family spectra to identify. Furthermore, other tests proved to us that ANNs are resistant to noise Presently, we must put the blame on the excessively short size of the training dataset.

4 Conclusions

We have studied the feasability of applying ANNs to uranium enrichment measurement. On data with a non-linear relationship between spectral response and analyte concentration, a neural network is shown to be able to predict ^{235}U concentrations. The results of the simulations which are presented in table (3) show the interest of the ANN method : it remains reliable in the general case. The ANN calibration results are especially interesting because of the presence of non-linear instrumental artefacts.

Our results appear to be at the state of art in automated quantifying methods for isotopes in a mixture of components. The basic principle is to use input and output data to provide information on how to set the parameters where no definite mathematical model can be assigned *a priori*. Thus we have adaptative prediction. This requires that the network parameters be set correctly for the work to be carried out as desired. This method has been demonstrated as a reliable tool for dealing with data from low resolution detectors even in under adverse conditions and has been already successfully used by [18] in a X-ray fluorescence application. Final network with connections and weighting functions could be easily implemented using commercial digital processing hardware.

But, there is no single learning procedure which is appropriate for all tasks. It is of fundamental importance that special requirements of each task are analyzed and that appropriate training algorithms are developed for families of tasks. However an efficient use of the networks requires as careful as possible analysis of the problem, an analysis that is often ignored by impatient users.

Acknowlegments

The authors are extremely grateful to the staff of SAR and LPRI, and in particular J.L. Boutaine, A.C. Simon and R. Junca for their support and useful discussions during this work. V. Vigneron wishes to thank C. Fuche from CREL for her support in this research.

References

[1] T.D. Reilly, Walton R.B., and Parker J.L. The enrichment metter. Report LA-4065-MS, Los Alamos National Laboratory, USA, 1970.

[2] L.A. Kull, R.O. Gonaven, and G.E. Glancy. A simple γ spectrometric technique for measuring isotopic abundances in nuclear materials. *Atomic Energy Review*, 144:681+, 1976.

[3] J. Morel, H. Goenvec, J. Dalmazzone, and G. Malet. Références pour la détermination de l'uranium 235 dans les combustibles nucléaires. *IAEA Nuclear Safeguard Technology 1978*, 1979.

[4] P. Matussek. Accurate determination of the ^{235}U isotope abundance by γ spectrometry, a user's manual for the certified reference material EC-NMR-171/NBS-SRM-969. Report KfK3752, Kernforschungszentrum Karlsruhe, Germany, 1985.

[5] R.J.S Harry, J.K. Aaldijk, and J.P. Brook. Gamma spectrometric determination of isotopic composition without use of standards. SM201/6, IAEA, France, 1980.

[6] R. Hagenauer. Non destructive uranium enrichment determination in process holdup deposits. *Nucl. Mater. Manag. Proc*, 1991.

[7] J.L. Pinault. La méthode d'analyse spectrale : une nouvelle approche pour la déconvolution des spectres X et γ. *Journées de Spectrométrie γ et X*, CEA-N-2756:175+, 1993.

[8] R. Gunning, W.D. Ruther, P. Miller, J. Goerten, M. Swinhoe, H. Wagner, J. Verplancke, M. Bickel, and S. Abousahl. MGAU: A new analysis code for measuring ^{235}U enrichments in arbitrary samples. UCRL-JC-114713, Lawrence Livermore National Laboratory, 1994.

[9] M.G. Neuilly. *Safeguards technical manual*, volume 3. IAEA, Vienna, 1980.

[10] W. Liggett. Calibration for measurements with background correction applied to uranium-235 enrichment. *Nuclear Instruments and Methods*, 1983.

[11] V. Vigneron and J.M. Martinez. A looking at Neural Network methodology. *to be published*, 1995.

[12] Y. LeCun. Efficient learning and second-order methods. Technical report, ATT & Bell Laboratories, 1993. Preprint.

[13] J.M. Martinez, Parey C., and Houkari M. Learning optimal control using neural networks. In *Proc. Fifth International Neuro-Nimes*, pages 431–442, Paris 12 University, 1992.

[14] D.E. Rumelhart and J.L. McClelland. *Parallel Distributed Processing*. MIT Press, Cambridge, MA, 1988. Chap.8.

[15] V. Vigneron, A.C. Simon, R. Junca, and J.M. Martinez. Neural techniques applied on measurement of uranium quantities by fluorescence X. *Preprint*, 1995.

[16] G. Saporta. *Probabilités, analyse de données et statistiques.* Technip, 1992.

[17] M.I. Jordan. *Lectures on neural networks.* Ecole CEA-INRIA-EDF au Bréau-Sans-Nappe, MIT, 1994.

[18] K. Abbas and R. Junca. Analyse par fluorescence X de solutions de nitrate dúranyle et de plutonium en présence de produits de fission. Note technique DTA/DAMRI/SAR/S/93-412, CEA Saclay, France, 1993.

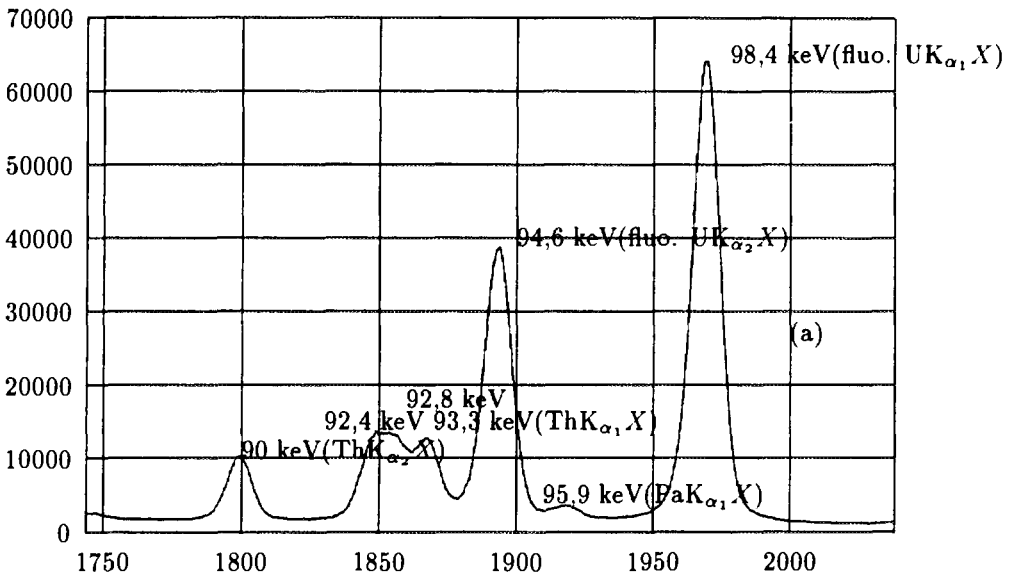


Figure 1: Principal useful X- and γ-rays in the spectral analysis of the $K_\alpha X$ region.

Diameter(cm)× Height(cm)	$\frac{U}{O}$ **ratio (g/g%)**	**Stated enrichment** ($10^{-2}g.g^{-1}$)	$\frac{^{235}U}{^{235}U+^{238}U}$ **ratio** ($10^{-2}g.g^{-1}$)
1,30× 2,00	88,00	0,7112 ±0,00004	0,7112
1,30× 1,90	88,00	1,416 ±0,001	1,416
0,80× 1,10	88,00	2,785 ±0,004	2,786
0,80× 1,02	87,96	5,111 ±0,015	5,112
0,80× 1,00	87,98	6,222 ±0,018	6,225
0.92× 1,35	87,90	9,548 ±0,04	9,558

Table 1: Percentage Enrichments of UO_2 standards

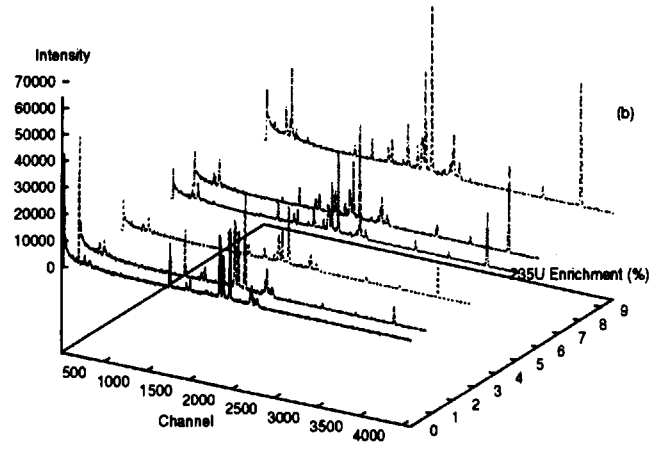


Figure 2: 3D-Representation of the UO_2 spectra set.

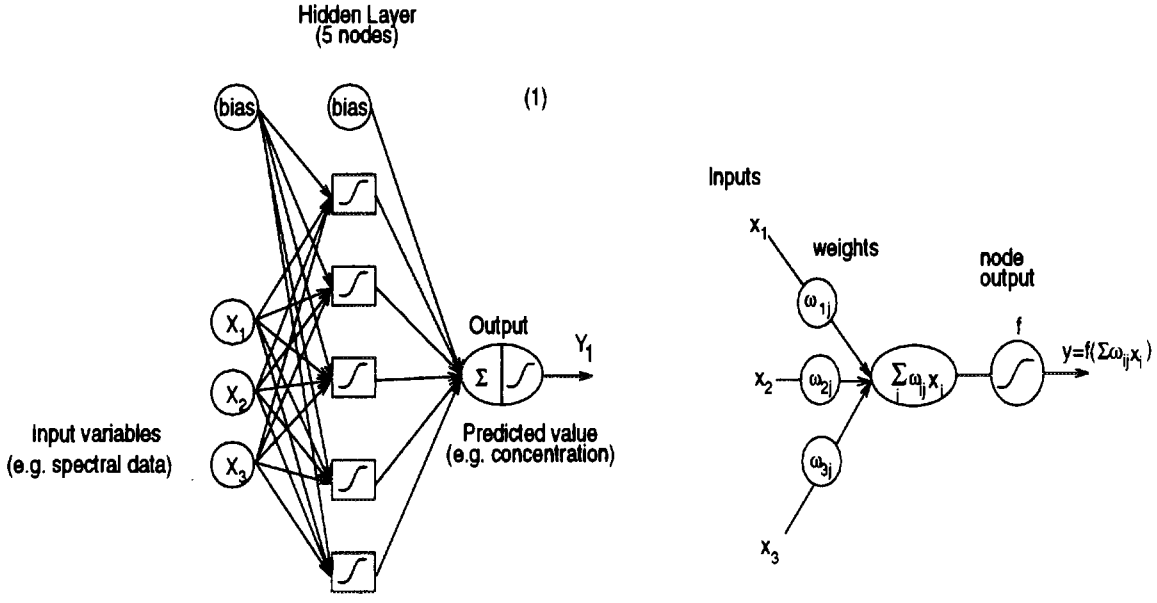


Figure 3: MLP 3-5-1 with nonlinear threshold and schematic representation of a node in an ANN.

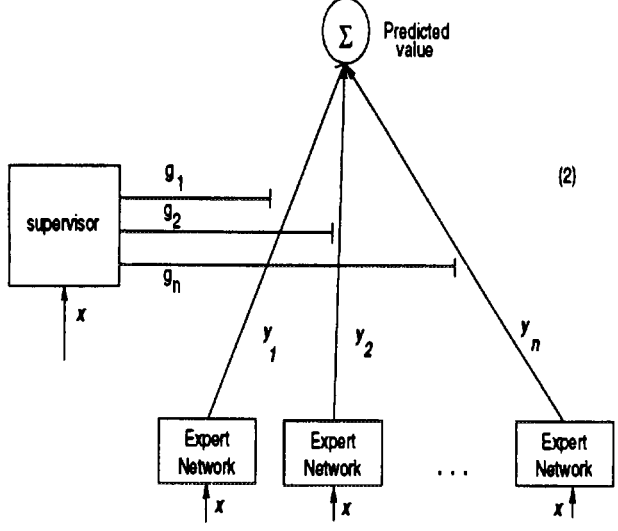


Figure 4: Mixtures of Experts model.

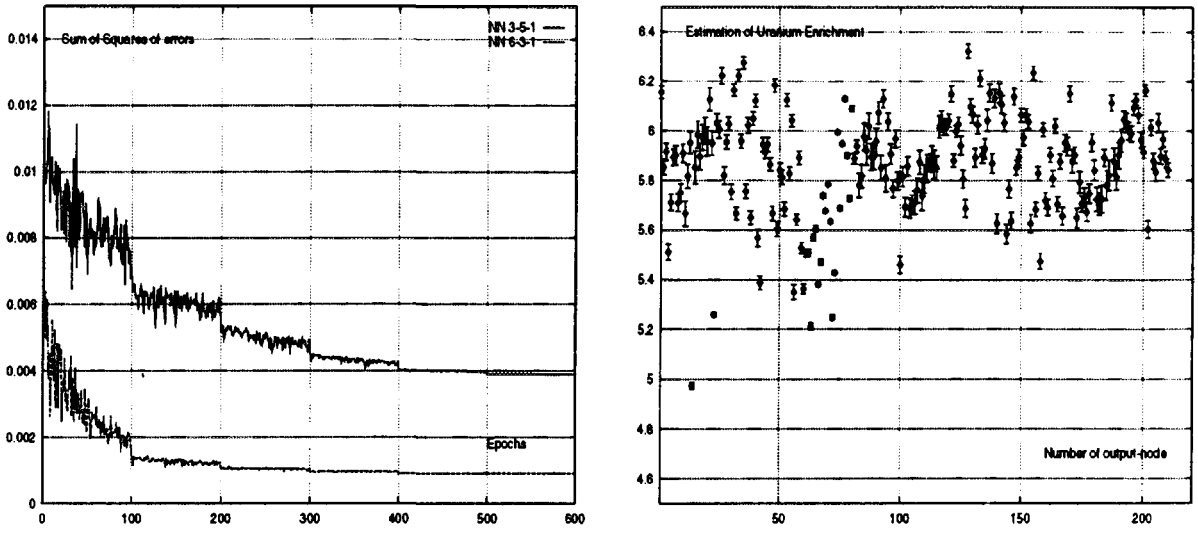


Figure 5: (a) Sum of squares of bias on the training set for MLP architectures (b) Example of enrichment value (at 5,785 %) predicted by the Mixtures of Experts

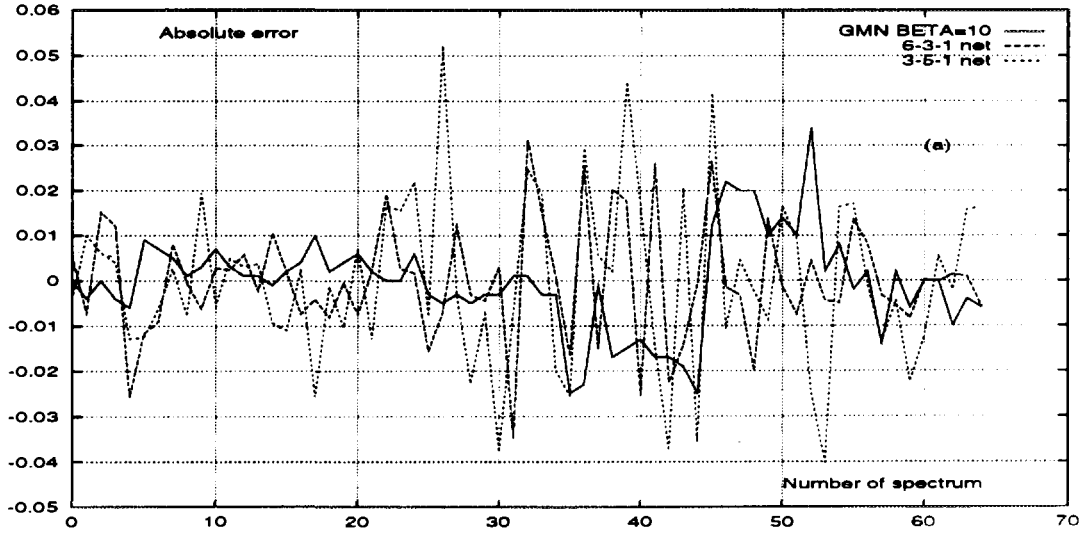


Figure 6: Absolute bias in the enrichment estimation

parameter	MLP 6-3-1	MLP 3-5-1	Mixtures of Experts
Type of input	spectral data	spectral data	enrichment value
input nodes	6	3	1
hidden node	3	5	1050
output node	1	1	210
learning rule	BP	BP	Maximum Likelihood
input layer transfer function	linear	linear	linear
hidden layer transfer function	sigmoïdal	sigmoïdal	sigmoïdal
output layer transfer function	linear	linear	exponential

Table 2: ANNs specifications and parameters

Declared enrichment	**MLP 3-5-1**	**MLP 6-3-1**	**MEXs** $\beta = 10$
0,711%	0.691-0.723	0.700-0.720	0.702-0.710
1,416%	1.394-1.426	1.406-1.435	1.406-1.416
2,785%	2.732-2.822	2.762-2.799	2.784-2.790
5,111%	5.066-5.148	5.089-5.132	5.112-5.136
6,122%	6.105-6.162	6.117-6.133	6.088-6.112
9,548%	9.531-9.570	9.541-9.550	9.542-9.552

Table 3: Min-Max of calculated Enrichments with MLP and MEX

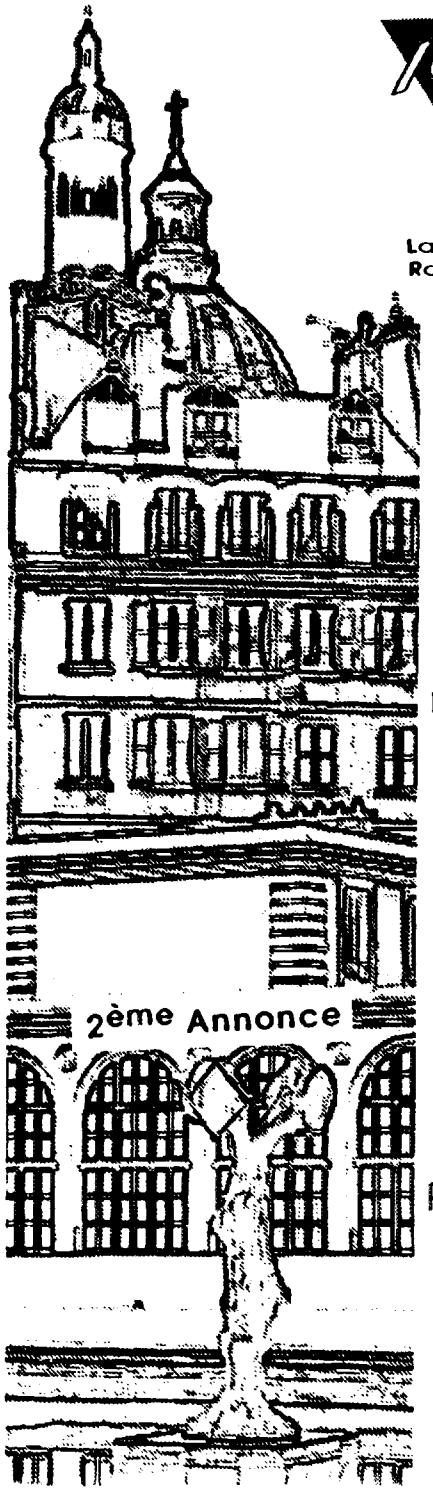
ICRM

International Commitee for Radionuclide Metrology

cea – BNM

Laboratoire Primaire des Rayonnements Ionisants

Symposium international sur la métrologie des radionucléides et ses applications



au Ministère de l'Enseignement Supérieur et de la Recherche

15-19 Mai 1995

PARIS (FRANCE)

RADMET'95

Exposition commerciale

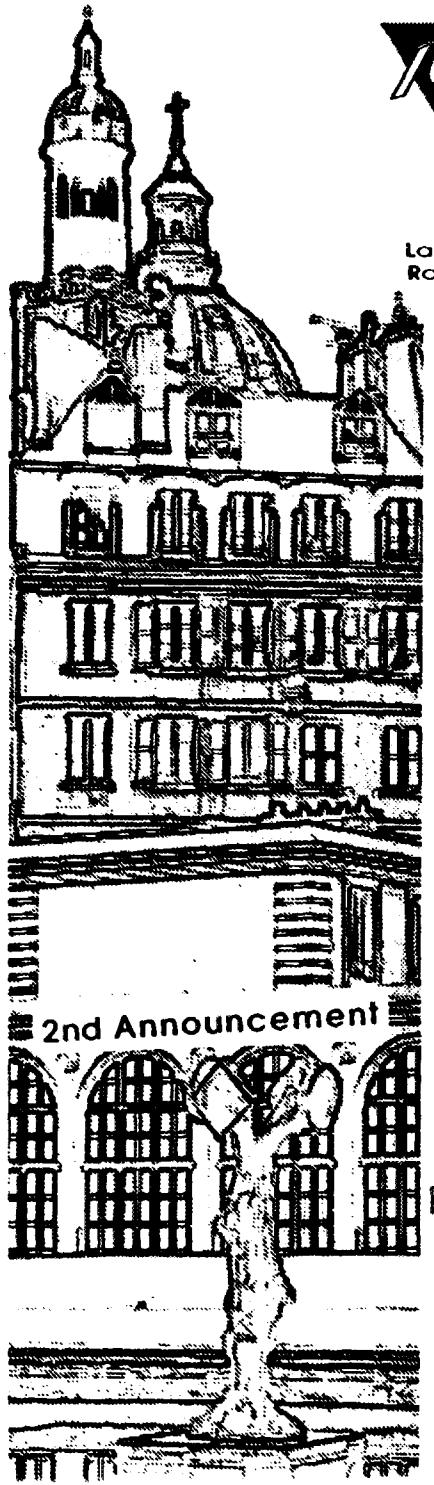
ICRM

International Commitee for Radionuclide Metrology

cea – BNM

Laboratoire Primaire des Rayonnements Ionisants

International Symposium on Radionuclide Metrology and its Applications



Held at the Ministère de l'Enseignement Supérieur et de la Recherche

15-19 May 1995

PARIS (FRANCE)

RADMET'95

Commerciale Exibition

EDITION SOREM CE SACLAY B.P. 2