



**HAL**  
open science

## Reconstruction problem by maximum entropy method applied on a mixture of experts

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

### ► To cite this version:

Vincent Vigneron, Jean-Marc Martinez, Marie-Christine Lepy, Jean Morel. Reconstruction problem by maximum entropy method applied on a mixture of experts. International Conference on Engineering Applications of Neural Networks (EANN '95), Aug 1995, Helsinki, Finland. hal-00221538

**HAL Id: hal-00221538**

**<https://hal.science/hal-00221538v1>**

Submitted on 30 Mar 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 Reconstruction problem by maximum entropy  
2 method

3 **V. Vigner**<sup>\*,1,2</sup>, **J. Morel**<sup>3</sup>, **M.C. Lépy**<sup>3</sup>, **J.M. Martinez**<sup>1</sup>

<sup>1</sup>CEA Saclay

DRN/DMT/SERMA

91191 Gif-sur-Yvette

FRANCE

<sup>3</sup>CREL

161, rue de versailles

78180 Le Chesnay

FRANCE

<sup>2</sup>CEA Saclay

DAMRI/LPRI, BP 52

91193 Gif-sur-Yvette cedex

FRANCE

4 **Abstract**

5 Layered Neural Networks, which are a class of models based on neu-  
6 ral computation , are applied to the measurement of uranium enrich-  
7 ment. The usual methods consider a limited number of  $\gamma$ -ray and  $X$ -  
8 ray peaks, and require previously calibrated instrumentation for each  
9 sample. But since, in practice, the source-detector ensemble geometry  
10 conditions are critically different, a mean of improving the above con-  
11 ventional methods is to reduce the region of interest ; this is possible  
12 by focusing on the  $K_{\alpha}X$  region where the three elementary components  
13 are present. The measurement of these components in mixtures leads to  
14 the desired ratio. Real data are used to study the performance of neu-  
15 ral networks and training is done with a Maximum Likelihood Method.  
16 We show that the encoding of data by Neural Networks is a promising  
17 method to measure uranium  $^{235}U$  and  $^{238}U$  quantities in infinitely thick  
18 samples.

19 **1 Introduction**

20 In the past few years, the topic of neural computing has generated widespread  
21 interest and popularity. The popularity of this technique is due in part to  
22 the analogy between Artificial Neural Networks (ANNs) and biological neu-  
23 ral networks. Many applications have been investigated using ANNs. We

---

<sup>0</sup> *Corresponding author:* e-mail [vvigne@soleil.serma.cea.fr](mailto:vvigne@soleil.serma.cea.fr).

24 demonstrate here how they can be used, with photons spectra, in the particu-  
25 lar case of uranium enrichment measurements, to determine the  $\frac{^{235}\text{U}}{\text{U}_{total}}$  isotope  
26 ratio. Indeed, with modern detectors and high technology, spectral data are  
27 collected with even finer sampling and with even large precision, what imposes  
28 a perpetual need for efficient interpretation methods.

29 Traditional non-destructive methods for uranium enrichment use several X-  
30 and  $\gamma$ -ray peaks, mainly in the 60 to 200 keV region. Most of these methods,  
31 which were developed more than 20 years ago, are based on measurements  
32 of the full energy peak at 185,7 keV ([10],[5],[9],[8]). They require a prior  
33 calibration of the system and the measurement conditions to be constant.  
34 Other methods have been developed using several  $\gamma$ -ray peaks [3],[2]. In fact,  
35 these latter methods require a self-calibration with a limited number of peaks,  
36 making them difficult to implement.

37 Calibration procedures and matrix effects can be avoided by first, focusing  
38 the spectra analysis on a limited region, called  $K_{\alpha}X$  region, containing the  
39 main uranium components, second, by using so-called infinitely thick samples.  
40 These samples are such that any further thickness increase does not affect the  
41  $\gamma$  emission.

42 The processing of the  $K_{\alpha}X$  region requires taking into account 3 elemental  
43 images corresponding to  $^{235}\text{U}$ ,  $^{238}\text{U}$  and X-ray fluorescence. This approach  
44 requires that all the parameters for constituting each elemental image are  
45 well-known and is based on the use of external data characterising the photon  
46 spectral emission together with the detector characteristics and geometry.

47 It is precisely in this context that a Neural Network appears to be a useful  
48 tool. In fact, the training by ANNs can be considered as a search procedure for  
49 an "optimum" regression function among a set of acceptable functions using a  
50 set of training examples. From the statistical point of view, ANNs belong to  
51 the evaluation techniques for non-parametric models, still called *tabula rasa*.  
52 ANNs, like most statistical methods, are able to process vast amounts of data  
53 and to make predictions that are sometimes surprisingly accurate. This does  
54 not make them intelligent in the usual sense of the word. ANNs learn in  
55 much the same way that many statistical algorithms do estimation. But in  
56 contrast to usual automatic spectra analysis methods, ANNs use full-parallel  
57 computing, are simple to implement, not very sensitive to outliers and contain  
58 nonlinearities.

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

## 63 2 Experimental Aspect

### 64 2.1 Preliminaries

65 In the case of uranium spectra, the efficiency response is difficult to establish  
 66 due to insufficient number of peaks that can be used. This can be overcome  
 67 by reducing the region of interest of the spectrum so that the variation in the  
 68 detector efficiency is limited. This is possible by considering only the relatively  
 69 complex  $K_{\alpha}X$  region, which extends from 83 to 103 keV, where many peaks  
 70 are superimposed (Fig. 1).

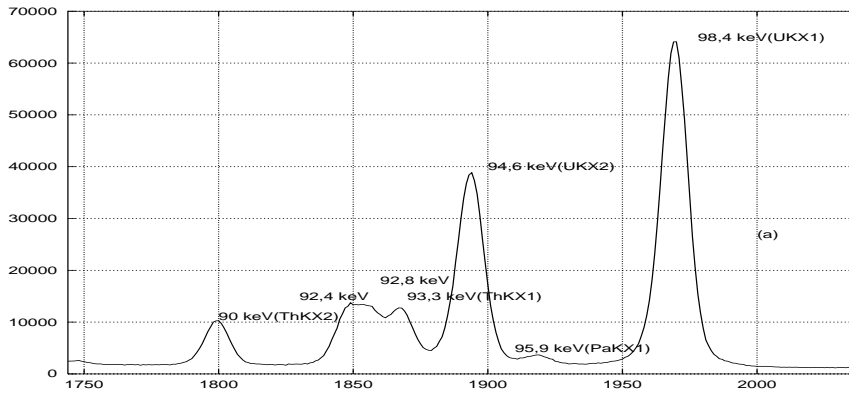


Figure 1: Principal useful  $X$ – and  $\gamma$ –rays in the spectral analysis of the  $K_{\alpha}X$  region.

71 This region contains enough information to allow the determination of  $^{235}\text{U}$   
 72 and  $^{238}\text{U}$  and is sufficiently small for considering the efficiency as constant. It  
 73 is however very complex to analyze, due to several interfering  $X$ - and  $\gamma$ -rays.  
 74 These can be grouped as follows :

- 75 •  $^{235}\text{U}$  and daughters : 84.21 keV ( $\gamma^{231}\text{Th}$ ), 89.95 keV ( $\gamma^{231}\text{Th}$ ,  $\text{Th}K_{\alpha_2}X$ ),  
 76 92.28 keV ( $\text{Pa}K_{\alpha_2}X$ ), 93.35 keV ( $\text{Th}K_{\alpha_1}X$ ), 95.86 keV ( $\text{Pa}K_{\alpha_1}X$ )
- 77 •  $^{238}\text{U}$  and daughters : 83.30 keV ( $\gamma^{234}\text{Th}$ ), 92.28 keV ( $\text{Pa}K_{\alpha_2}X$ ), 92.38  
 78 keV ( $\gamma^{234}\text{Th}$ ), 92.79 keV ( $\gamma^{234}\text{Th}$ ), 94.65 keV ( $\text{UK}_{\alpha_2}X$ ), 95.86 keV ( $\text{Pa}K_{\alpha_1}X$ ),  
 79 98.43 keV ( $\text{UK}_{\alpha_1}X$ ), 99.85 keV ( $\gamma^{234}\text{Pa}$ )
- 80 • Uranium X-ray fluorescence : 94.65 keV ( $K_{\alpha_2}X$ ), 98.43 keV ( $K_{\alpha_1}X$ ).

81 In the standard approach, the processing of the considered region takes into  
 82 account the 3 elemental images, the first corresponding to  $^{235}\text{U}$  and his daugh-  
 83 ters, the second to  $^{238}\text{U}$  and its daughters and the third to the uranium  $X$ -ray  
 84 fluorescence spectrum. These images are then represented by mathematical ex-  
 85 pressions taking into account the shapes of the  $X$ -ray (Voigt profile) and  $\gamma$ -ray

86 (Gaussian) peaks, their energies, and intensities. The determination is then  
 87 carried out conventionally with a least squares method, like the MGA-U code  
 88 [1]. The final enrichment is obtained by correcting for the presence of  $^{234}\text{U}$   
 89 with the 120.9 keV peak.

## 90 2.2 Experimental protocol

91 Six uranium oxide standards with different enrichments, from 0.7 to 9.6%,  
 92 and infinite thickness were counted several times by  $\gamma$ -ray spectrometry to  
 93 test the neural procedure. These were bare cylindrical pellets, with certified  
 94 enrichments and their main characteristics are presented in Table 1.

Table 1: Characteristics of  $\text{UO}_2$  standards

<b>Diameter(cm)× Height(cm)</b>	$\frac{\text{U}}{\text{O}}$ <b>ratio (<math>g \cdot g^{-1}\%</math>)</b>	<b>Stated enrichment (<math>g \cdot g^{-1}\%</math>)</b>	$\frac{^{235}\text{U}}{^{235}\text{U}+^{238}\text{U}}$ <b>ratio (<math>g \cdot g^{-1}\%</math>)</b>
1,30× 2,00	88,00	0,7112 ±0,004	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

95 The Ge(HP) planar detector used in the measurement system had the  
 96 following specification : surface,  $2.00 \text{ cm}^2$  ; thickness,  $1.00 \text{ cm}$  ; FWHM, 190  
 97 eV at 6 keV and 480 eV at 122 keV. All the measurements were made under  
 98 the same conditions, i.e. with 0.05 keV per channel and a distance between  
 99 source and detector-window of 1.1 cm. Ten 20000-s. spectra for each standard  
 100 pellet were analysed by our procedure. The  $^{234}\text{U}$  concentration is relatively  
 101 low, although a  $\frac{^{234}\text{U}}{^{235}\text{U}}$  mass ratio varying from 0,5 to 1,1%, depending on the  
 102 pellet, was determined by  $\gamma$ -ray spectrometry by using both the 53.2 and the  
 103 120.9 keV peaks for  $^{234}\text{U}$  and the 185.7 keV peak for  $^{235}\text{U}$ .

104 In short, 65 sets of experimental data from real-life experiments were pre-  
 105 pared using the concentrations given in Table 1, and are illustrated in Fig.  
 106 2.

## 107 3 Layered Neural Network and Training method

### 108 3.1 Using Neural Networks

109 The purpose of this section, rather than the presentation of the neural network  
 110 theory, is to present the place of the connectionnist approach in  $\gamma$ -spectrometry

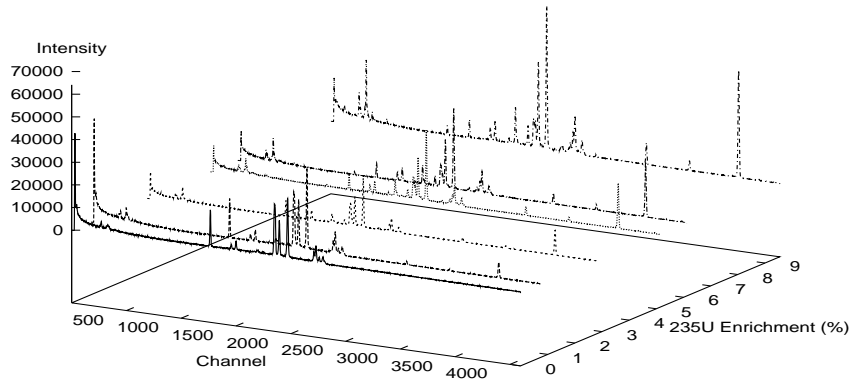


Figure 2: 3D-Representation of the  $UO_2$  spectra set.

111 problems. Neural Networks are non-linear black-box model structures, to be  
 112 used with conventional parameter estimation methods. Most details and basic  
 113 concepts are clearly described in a paper to be published [12]. ANN consists  
 114 of a large number of neurons, i.e. simple linear or nonlinear computing el-  
 115 ements, interconnected in complex ways and often organized into layers [6].  
 116 The collective or parallel behaviour of the network is determined by the way  
 117 in which the nodes are connected and the relative type and strength (excitory  
 118 or inhibitory) of the interactions among them [7].

119 The objective of ANNs is to construct a suitable model which, when applied  
 120 to a  $^{235}U$  enrichment spectrum, produces an output,  $y$ , which approximates  
 121 the exact uranium enrichment ratio. The principal idea of the connectionist  
 122 approach is to substitute a neural model and the learning procedure of the net-  
 123 work for classical fitting algorithms, which make use of complex mathematical  
 124 algorithms, generally based on the separation of a given curve, associated to  
 125 each individual peak, plus a background.

126 An exemple of multi-layer network is given in Fig. 3.a. The notation  
 127 convention is such that the square represents a computational unit into which  
 128 the input variables  $x_j$ 's are fed and multiplied by the respective weights  $\omega_j$ 's.  
 129 The fundamental processing element of an ANN is a node (Fig. 3.b). Nodes are  
 130 analogous to neurons in biological systems. Each node has a series of weighted  
 131 inputs,  $\omega_i$ , which may be either an external signal or the output from other  
 132 nodes. The sum of the weighted inputs is transformed with a linear or a  
 133 non-linear transformation function (often the logistic function  $f(x) = \frac{1}{1+e^{-x}}$ ).  
 134 In the statistical context, this standard Neural Network called *Multi-Layered*  
 135 *Perceptron (MLP)*, is analogous to the Multivariate Nonlinear Regression.

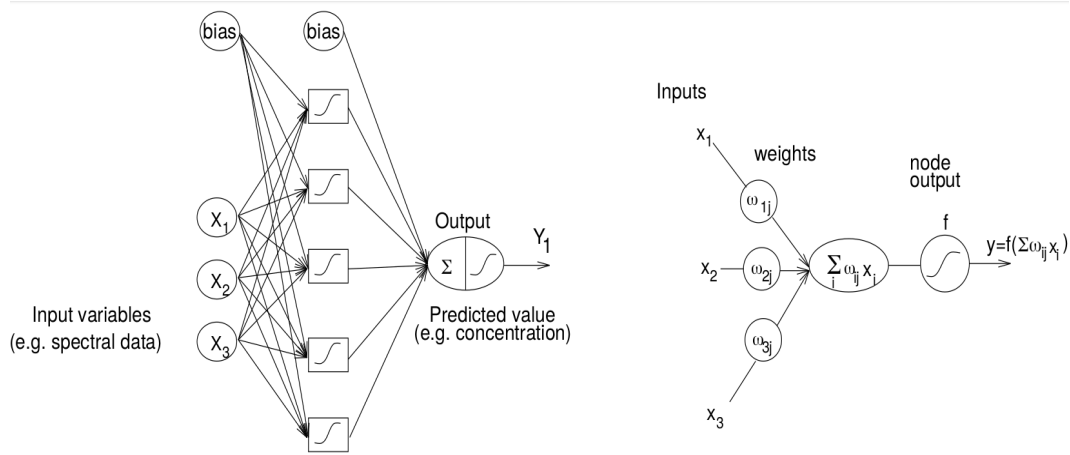


Figure 3: (a) MLP 3-5-1 with nonlinear threshold and (b) schematic representation of a node in an ANN.

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

- 139 • input layer : this layer contains input units. Each unit receives input-  
 140 variables, selected through a free parameters reduction procedure.
- 141 • hidden layer : this layer acts as an array of feature detectors picking  
 142 up features without regard to position. The information coming to the  
 143 input units is coded on the hidden layer into an internal representation  
 144 Thus, the input-layer units contribute to the input of each second-layer  
 145 unit. It is fully-connected to the output.
- 146 • output layer : it applies a sigmoid activation function to the weighted  
 147 sum of the hidden outputs.

148 The role of the hidden layer is fundamental. A network without hidden units  
 149 will be unable to perform the necessary multi-input multi-output mappings,  
 150 in particular with non-linear problems. Input pattern can always be encoded,  
 151 if there are enough hidden units, in a form so that the appropriate output  
 152 pattern can be generated from the corresponding input pattern.

153 The training data are denoted by  $\chi = (\mathbf{x}, \mathbf{y}^d)_{t=1}^N$  where  $N$  is the number  
 154 of observations and  $\mathbf{x}$  is the feature vector corresponding to the  $t^{th}$  obser-  
 155 vation. The expected response  $\mathbf{y}^d = (y_1, y_2, \dots, y_M)$  is related to the inputs  
 156  $\mathbf{x} = (x_1, x_2, \dots, x_N)$  according to

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

157 where  $\omega$  are the *connection weights*.

158 The approximation results are non-constructive, and in practice the weights  
159 have to be chosen to minimize some fitting criterion, e.g. least squares

$$J(\omega) = \frac{1}{2} \sum_p^N (\mathbf{y}_p^d - \phi(\mathbf{x}_p, \omega))^2, \quad (2)$$

160 with respect to all the parameters, where  $\mathbf{y}_p^d$  is the target for the  $p$ th example  
161 pattern. The minimization has to be done by some numerical search procedure.  
162 This is called *nonlinear optimization*. The parameter estimate is defined as  
163 the minimizing argument :

$$\hat{\omega} = \operatorname{argmin}_{\omega} J(\omega) \quad (3)$$

164 Most efficient search routines are based on local iteration along a "downhill"  
165 direction from the current point. We then have an iterative scheme of the  
166 following kind :

$$\hat{\omega}^{(i+1)} \leftarrow \hat{\omega}^{(i)} - \eta \times \frac{\partial J}{\partial \omega^{(i)}} \quad (4)$$

167 where  $\hat{\omega}^{(i)}$  is the parameter estimate after iteration number  $i$ ,  $\eta (> 0)$  is the  
168 step size and  $\frac{\partial J}{\partial \omega^{(i)}}$  an estimate of the gradient of  $J(\omega^i)$ . The practical difference  
169 between this device and the statistical version lies in the way the training data  
170 are used to dictate the values for  $\omega$ . It turns out that there are 2 main aspects  
171 to the processing : (1) specifying the architecture of a suitable network, (2)  
172 training the network to perform well with reference to a training set.

### 173 3.2 Application of the ANN

174 To check that this method was general and reliable, we have applied it to 65  
175 sets of experimental data from real-life experiment : five  $^{235}\text{U}$ -pure idealized  
176 spectra, and ten of each precited standard (see Table 1). Each spectrum  
177 contains 4096 points. The computations of the spectra are compared on two  
178 regression models: the **MLP MODEL** ( Fig. 3), where the inputs are spectral  
179 data, and the **MIXTURES OF EXPERTS MODEL** (Fig. 4) [4] where the inputs are  
180 the enrichment values.

181 The specifications for the networks created for the calibration of the sim-  
182 ulated data are listed in Table 2. They were found to be optimal according  
183 to the rigourous methodology described in [12], for low prediction bias. The  
184 choice of the right architecture is mainly intuitive and implies arbitrary deci-  
185 sions. But an attempt to apply ANN directly fails due to dimensionability. In  
186 accordance with this, the dimension of the input vector has been reduced dra-  
187 matically by Principle Components Analysis (PCA), leading to the adequate



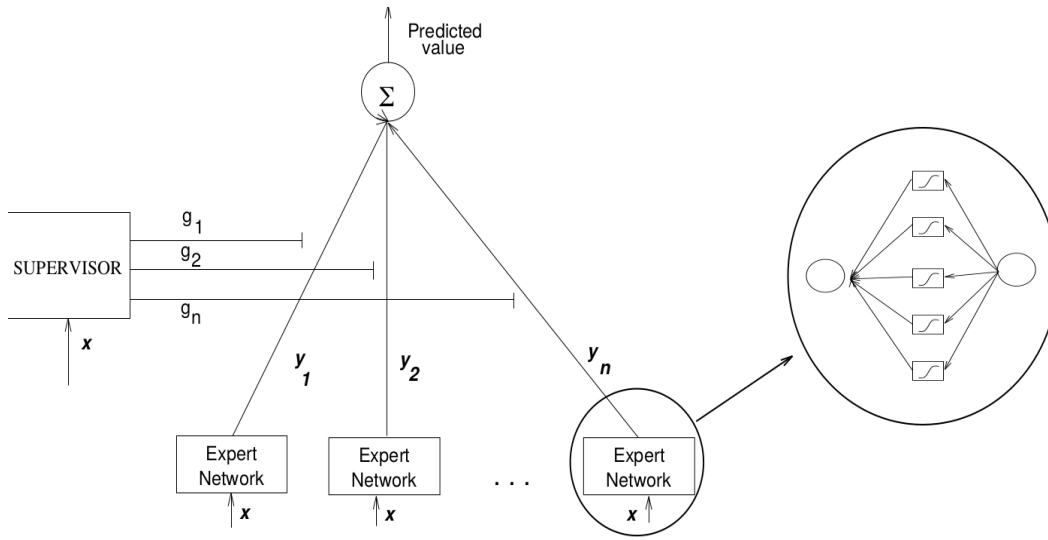


Figure 4: Mixtures of Experts model.

188 reduction of weights emerging from the first layer of the ANN.

Table 2: ANNs specifications and parameters

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	sigmoidal	sigmoidal	sigmoidal
output layer transfer function	linear	linear	exponential

189 The **MLP MODEL**, depicted in Fig. 3, consists of an input layer of 6 or 3  
 190 units leading up through one layer of hidden units to an output layer of a single  
 191 unit that corresponds to the desired enrichment. This network represents a  
 192 poor parametrized model, but the training dataset  $(\mathbf{x}; \mathbf{y}^{(d)})_{t=1}^{65}$  was small. The  
 193 network is initialized with random weights and trained. For each pattern, the  
 194 bias, Eq. 2, is evaluated. This quantity decreases rapidly (Fig. 5) in the  
 195 beginning, and the training is stopped when the network reaches a minimum  
 196 error on the training set, because this is an efficient way to avoid overfitting.  
 197 After 32 000 successful training passes, the bias rate range from -0.05 to 0.04%  
 198 for the 6-3-1 net (from -0.031 to 0.061% for the 3-5-1 net).

199 In the case of mixtures of experts (MEX), each item is associated with a  
 200 vector of measurable features, and a target  $\mathbf{y}^d$  which represents the enrich-  
 201 ment. The network receives the input  $\mathbf{x}$  and creates the output vector  $\mathbf{y}$  as

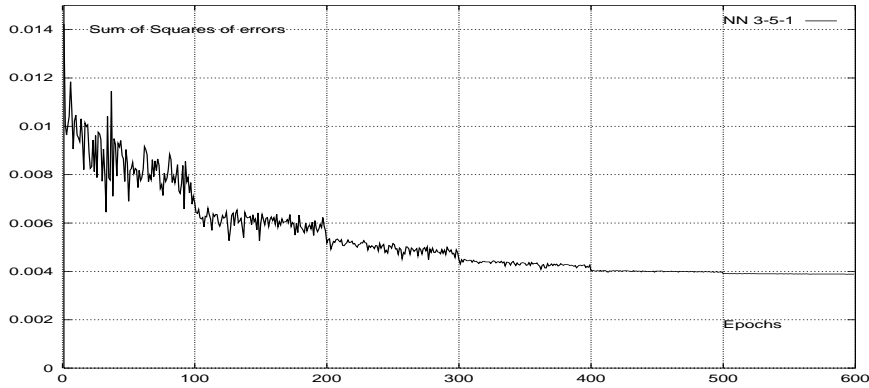


Figure 5: Sum of squares of bias on the training set for MLP architectures

202 a predicted value of the unknown  $\mathbf{y}^d$ . This model consists of 210 independant  
 203 fully-connected networks (Fig. 4) : One expert is put for one channel of  
 204 the  $K_\alpha X$  region, each expert being an observer, trying to find a "signal" due  
 205 to radioactive decay in a large amount of noise, the variance of each count  
 206 being proportional to the level and thus depending on the enrichment of a  
 207 particular sample and on the background level of the particular observation.  
 208 A cooperation-competition procedure driven by a supervisor between the ex-  
 209 pert's outputs leads to the choice of the most appropriate concentration.

210 Let  $\mathbf{y}_1, \mathbf{y}_2, \dots$  denote the output vectors of the experts, and  $g_1, g_2, \dots$   
 211 the supervisor output units, then the output of the entire architecture,  $\mathbf{y}$ ,  
 212 is  $\mathbf{y} = \sum_{i=1}^{210} g_i \mathbf{y}_i$ . The supervisor decides whether expert  $i$  is currently appli-  
 213 cable or not. The winning expert is the network with the smallest bias ( $\mathbf{y}^d - \mathbf{y}_i$ ).

214

## 215 4 Discussion of the Results using ANN

216 As the initial base included only 65 examples, we wanted to keep a maximum  
 217 of examples for the training base. Redundances in the data-set enrichments  
 218 present one main advantage : as we measure more than one response for each  
 219 case, information from all the measured responses can be combined to provide  
 220 more precise parameter estimation and to determine a more realistic model.

221 In all simulations, the measure of the system's performance is the Mean  
 222 Square Error. The bias rates obtained by using MEX are benchmarked against  
 223 the results obtained by using MLPs in Table 3 and on Fig. 5 and Fig. 7. The  
 224 Fig. 5 shows the learning curves (i.e. the learning performances) for the two  
 225 MLP networks using a random training procedure. The horizontal axis gives  
 226 the number of epochs ; the vertical axis gives the Mean Square Errors value  
 227 (MSE). Clearly, the 6-3-1 network learned significantly faster than the 3-5-

228 1. This difference can be explained by the information gain of the 6-inputs  
 229 network vs the 3-inputs network.

Table 3: Ranges of calculated Enrichments with MLP and MEX

Declared enrichment	MLP 3-5-1	MLP 6-3-1	MEXs
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

230 The Fig. 6 concerns the Multi-Expert model. The plotted points are pre-  
 231 dicted enrichment value (one for each of the 210 experts) when a 5.111%  $-^{235}U$   
 232 spectrum is presented to the MEX model. The credit assignment procedure  
 233 on these 210 contributions is supervised to produce a final estimation. In the  
 234 right most column of Table 3, the final predicted values of the simulations with  
 235 MEX can be seen. Compared with the MLPs, this shows that MEX method is  
 236 really reliable ; for example, the bias between the predicted and the calculated  
 237 2.785% enrichments range from 2.784 to 2.790%. As noted above, after 32  
 238 000 successful training passes, the larger bias happens for 5.111 and 6.122%  
 239 enrichments. This relative lack of precision can be ascribed to the small size  
 240 of the training dataset.

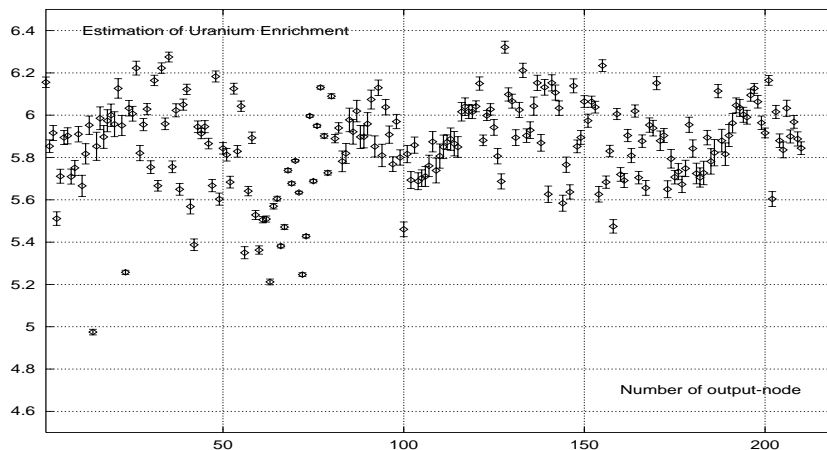


Figure 6: Example of enrichment value (at 5,785 %) predicted by the Mixtures of Experts

241 Fig. 7 compares the results of the three models. The bias between the  
 242 predicted and the desired enrichments is plotted for each of the 65 samples.

243 The darkest line is put for the MEX. The results suggest that the strong dis-  
 244 persion of the bias with MLP is significantly attenuated when MEX is applied.  
 245 This judgement must be moderated for the 6.122-enrichment-ratio samples. A  
 246 comparison of the absolute bias curves suggest that, among of the three sys-  
 247 tems studied, the Mixtures of Experts is capable of showing the most robust  
 248 performance.

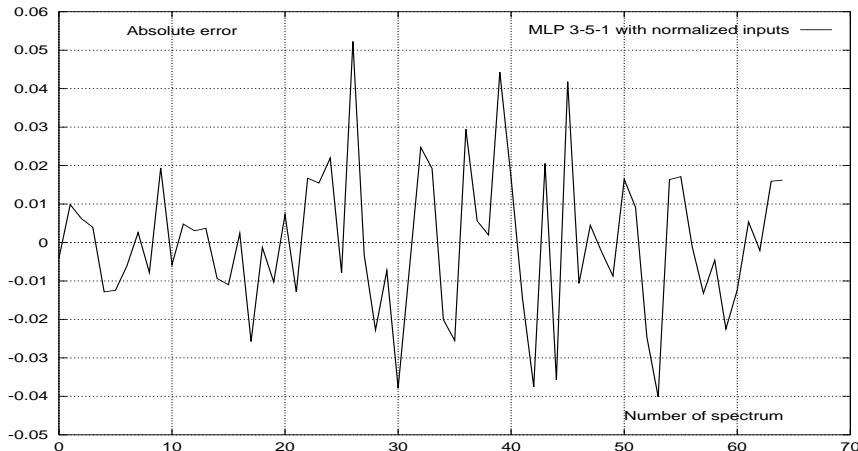


Figure 7: Absolute bias in the enrichment estimation

249 In fact, the modular approach presents three main advantages on the MLP  
 250 models: it is able to model behavior, it learns faster than a global model  
 251 and the representation is easier to interpret. The modular architecture takes  
 252 advantage of task decomposition, but the learner must decide which variables  
 253 to allocate to the networks. This method is, at the same time, very general  
 254 and very specific. It is very general in the sense that no hypothesis is made  
 255 on the aspect of the spectra : it does not depend on whether the spectra are  
 256 well resolved or not, whether they are very likely or not, whether you select  
 257 most significant areas of spectrum only (MLP models) or a global part of the  
 258 spectrum (MEX model). But, at the same time, the method is very specific  
 259 because the ANN must learn representative spectra of the family spectra to  
 260 identify. Furthermore, other tests proved to us that ANNs are resistant to  
 261 noise. Presently, we must put the blame on the excessively short size of the  
 262 training dataset.

## 263 5 Conclusion

264 The simulation studies on  $UO_2$  real spectra have shown that Neural Networks  
 265 can be very effective to predict  $^{235}U$  enrichment. They appear to be useful  
 266 when a fast response is needed with a reasonable accuracy, when no hypoth-

267 esis is made on the aspect of the spectra or when no definite mathematical  
268 model can be assigned *a priori*. The resistance to noise is certainly one of the  
269 most powerful characteristics of this method. Final network with connections  
270 and weighting functions could be easily implemented using commercial digi-  
271 tal processing hardware. The good results obtained show that this method  
272 can be considered at the state of art to produce quantitative estimates of  
273 the concentrations of isotopic components in mixtures with fixed experimental  
274 conditions : they may be better than those obtained with standard methods  
275 in similar cases. This method has also been already successfully used in an  
276 X-ray fluorescence application [11].

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

## 282 Acknowledgments

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

## 287 References

- 288 [1] GUNNING, R., RUTHER, W. D., MILLER, P., GOERTEN, J., SWINHOE,  
289 M., WAGNER, H., VERPLANCKE, J., BICKEL, M., AND ABOUSAHL,  
290 S. MGAU: A new analysis code for measuring  $^{235}\text{U}$  enrichments in arbi-  
291 trary samples. Report UCRL-JC-114713, Lawrence Livermore National  
292 Laboratory, (Livermore), 1994.
- 293 [2] HAGENAUER, R. Non destructive uranium enrichment determination in  
294 process holdup deposits. In *Nucl. Mater. Manag. Proc* (1991).
- 295 [3] HARRY, R. J. S., AALDIJK, J. K., AND BROOK, J. P. Gamma spec-  
296 trometric determination of isotopic composition without use of standards.  
297 In *Report SM201/6 IAEA, Viena* (1980).
- 298 [4] JORDAN, M. I. *Lectures on neural networks*. Summer school, CEA-  
299 INRIA-EDF, MIT, 1994. Private Communication.

300 [5] KULL, L. A., GONAVEN, R. O., AND GLANCY, G. E. A simple  $\gamma$  spec-  
301 trometric technique for measuring isotopic abundances in nuclear materi-  
302 als. *Atomic Energy Review 144* (1976), 681–689.

303 [6] LECUN, Y. Efficient learning and second-order methods. Tech. rep., ATT  
304 & Bell Laboratories, 1993. Private Communication.

305 [7] MARTINEZ, J. M., PAREY, C., AND HOUKARI, M. Learning optimal  
306 control using neural networks. In *Proc. Fifth International Neuro-Nimes*  
307 (Paris, 1992), CEA (Saclay), EC2, pp. 431–442.

308 [8] MATUSSEK, P. Accurate determination of the  $^{235}\text{U}$  isotope abundance  
309 by  $\gamma$  spectrometry, a user’s manual for the certified reference material  
310 EC-NMR-171/NBS-SRM-969. In *Report KfK-3752* (Karlsruhe, 1985).

311 [9] MOREL, J., GOENVEC, H., DALMAZZONE, J., AND MALET, G.  
312 Références pour la détermination de l’uranium 235 dans les combustibles  
313 nucléaires. *IAEA Nuclear Safeguard Technology* (1978).

314 [10] REILLY, T. D., WALTON, R. B., AND PARKER, J. L. The enrichment  
315 meter. In *Report LA-4065-MS* (Los Alamos, 1970).

316 [11] SIMON, A. C., AND JUNCA, R. *Dosage de l’uranium et du pluto-*  
317 *onium par l’appareil de fluorescence X SAPRA  $\gamma X/\text{Ir}$ : simplification*  
318 *des codes de calcul dans le cas de solutions pures.* In *Note technique*  
319 *DTA/DAMRI/SAR/S/94-433* (Saclay, 1994).

320 [12] VIGNERON, V., AND MARTINEZ, J. M. A looking at Neural Network  
321 methodology. *Submitted to Nuclear Instruments and Methods in Physics*  
322 *Research* (1995).

## 323 List of Figures

324	1	Principal useful $X$ - and $\gamma$ -rays in the spectral analysis of the	
325		$K_{\alpha}X$ region. . . . .	3
326	2	3D-Representation of the $UO_2$ spectra set. . . . .	5
327	3	(a) MLP 3-5-1 with nonlinear threshold and (b) schematic rep-	
328		resentation of a node in an ANN. . . . .	6
329	4	Mixtures of Experts model. . . . .	8
330	5	Sum of squares of bias on the training set for MLP architectures	9
331	6	Example of enrichment value (at 5,785 %) predicted by the Mix-	
332		tures of Experts . . . . .	10
333	7	Absolute bias in the enrichment estimation . . . . .	11