

## PREDICTION OF THE DIOXIN, PCB, AND TOTAL WHO-TEQ VALUES ON THE BASIS OF FIVE CONGENER CONCENTRATION: TOWARD A NEW SCREENING STRATEGY FOR POPs IN FISH?

Marchand P<sup>1</sup>, Antignac JP<sup>1</sup>, Gade C<sup>1</sup>, Venisseau A<sup>1</sup>, Brosseaud A<sup>1</sup>, Bertrand D<sup>2</sup>, Qannari El M<sup>2</sup>, Sirot V<sup>3</sup>, Tard A<sup>3</sup>, Volatier JL<sup>3</sup>, Leblanc JC<sup>3</sup>, André F<sup>1</sup>, and Le Bizec B<sup>1</sup>

<sup>1</sup>Laboratoire d'Etude des Résidus et Contaminants dans les Aliments (LABERCA), Ecole Nationale Vétérinaire de Nantes (ENVN), BP 50707, 44307 Nantes Cedex 3, France ; <sup>2</sup> Unité de Sensométrie et de Chimiométrie, Ecole Nationale d'Ingénieurs des Techniques des Industries Agricoles et Alimentaires (ENITIAA), Nantes, France ; <sup>3</sup> AFSSA-DERNS, Equipe appréciation du risque (AQR), Maisons-Alfort, France.

### Introduction

Current regulation regarding levels of PCDD/Fs and PCB in food is based on WHO-TEQ values, which are calculated from concentrations measured for 17 dioxin and 12 PCB congeners. Time and cost related to these analysis are very significant, due to the sample preparation procedure (purification of different fractions containing each analyte family) and data analysis (quantification of the different congeners on the basis of GC-HRMS chromatograms traces and isotope dilution method). From this point of view, the possibility to predict the final WHO-TEQ values on the basis of a limited number of congeners should be very beneficial. In the present study, multivariate statistical techniques (PCA, hierarchical clustering, multiple linear regression) were used in order to investigate the correlations in-between the different analytes measured in fish samples (i.e. relations between congeners considered as statistical variables). The final purpose was to propose a diagnostic model permitting to predict the different TEQ values (Dioxin TEQ, PCB TEQ and Total TEQ) on the basis of concentration results obtained for a minimal number of PCDD/PCDF/PCB congeners.

### Material and Methods

#### *Samples*

The first part of the present study (i.e. elaboration of the predictive model) was based on the analysis results obtained for 94 fish samples, according a total diet study food composite sampling approach, which were collected within the framework of a national research project coordinated by the Institut National de la Recherche Agronomique (INRA) and the Agence Française de Sécurité Sanitaire des Aliments (AFSSA). These samples covered 28 fish species as well as 4 catching locations. The second part of the study (i.e. validation of the elaborated model) was based on the experience of LABERCA as National Reference Laboratory (NRL) in charge of these substances, toward three additional and independent data set. The first and second validation set corresponded to 60 and 87 fish samples, respectively, which were analysed in the frame of the 2004 and 2005 French monitoring plans, respectively. The last data set consisted in 3 fish samples assayed for interlaboratory studies conducted in 2004 and 2005.

#### *Reagents and chemicals*

Organic solvents (pentane, hexane, cyclohexane, isooctane, toluene, acetone, dichloromethane, diethylether, ethanol and methanol) were of picrograde® quality and provided by Promochem (Molsheim, France). Acetic and sulphuric acids were purchased from SDS (Peypin, France). Sodium sulphate and potassium oxalate were from Merck (Darmstadt, Germany). Silica gel was from Fluka. Native <sup>12</sup>C and <sup>13</sup>C-labelled PCDD/PCDF/PCB congeners were provided by Promochem.

#### *Sample preparation*

For each analysed fish sample, 10-20 g aliquots of fresh material (corresponding to 0.5-1.5 g equivalent fat) were freeze-dried, powdered, and transferred into Accelerated Solvent Extraction (ASE) cells. Pressure and temperature were set to 100 bar and 120 °C respectively. Four successive extraction cycles (5 min each) were performed using a mixture toluene/acetone 70:30 (v/v) as extraction solvent. The extract was evaporated to dryness, permitting the gravimetric determination of the fat content. Extracts were dissolved in hexane and a classical 3 steps purification process was then performed, using successively activated silica, florisil and celite/carbon stationary phases. <sup>13</sup>C-labelled internal standards were introduced in all samples before extraction and used for quantification according to the classical isotope dilution method.

#### *GC-HRMS analysis*

GC-HRMS detection was performed on a Hewlett-Packard 6890 gas chromatograph (Palo-Alto, USA), equipped with a DB-5MS column (30 m x 0.25 mm i.d., 0.25 µm film thickness), and coupled to a Jeol JMS-700D high

## Levels in feed and food

resolution mass spectrometer. GC program was as follows: 120°C (3 min), 20°C/min until 170°C (0 min), and then 3°C/min until 275°C (7 min). Injector as well as transfer line temperature were set to 280 °C. Acquisition was performed in SIM mode with a resolution higher than 10 000 (10 % valley). Electron impact ionization energy was at 38-40 eV and ion source temperature was maintained at 280 °C. Each of the monitored PCDD/PCDF/PCB congeners was identified on the basis of their molecular ion  $[M]^+$  and the corresponding  $^{37}\text{Cl}$  isotopic contribution. Indicator PBDE congeners were also monitored.

### Statistical analysis

Each PCDD/PCDF/PCB congener was considered as a statistical variable and each analyzed sample as an observation. The values assigned to each variable were the concentrations measured for the corresponding congener, expressed in  $\text{pg}\cdot\text{g}^{-1}$  fw. For each observation, additional informative variables were introduced including the extracted fat amount, the WHO-TEQ calculated for dioxins (Diox TEQ), PCB (PCB TEQ) and the sum of dioxin+PCB (Total TEQ), the nature/species of the sample, as well as its location of collection. Statistical analysis included principal component analysis (PCA), hierarchical clustering of the variables, and step-by-step incremental multiple linear regression, and were realized using Statistica© software (Statsoft, Inc., Tulsa, USA).

## Results and Discussion

The principal component analysis (PCA) performed on the first data set (Fig. 1) demonstrated a clear correlation in-between the 3 global TEQ values (Diox TEQ, PCB TEQ and Total TEQ) and several PCDD/PCDF/PCB congeners. The possibility of a predictive power associated to these congeners should be then further considered.

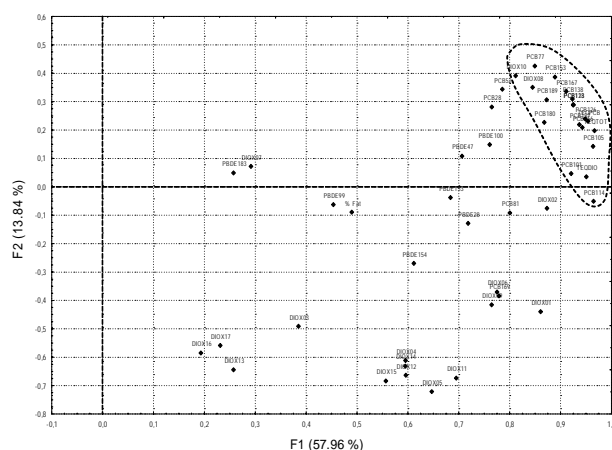


Figure 1 : Representation of the different variables on the two first axis extracted by the PCA performed on the first data set (n=94 samples).

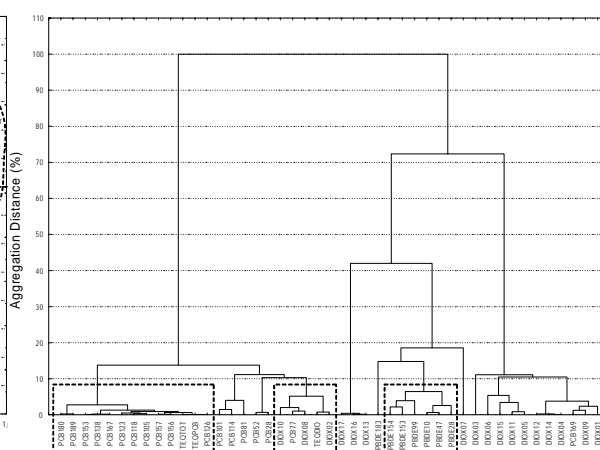


Figure 2 : Result of the hierarchical clustering of the variables from the first data set (n=94 samples).

(Diox1: 2.3.7.8-TCDD; Diox2: 1.2.3.7.8-PeCDD; Diox3: 1.2.3.4.7.8-HxCDD; Diox4: 1.2.3.6.7.8-HxCDD; Diox5: 1.2.3.7.8.9-HxCDD; Diox6: 1.2.3.4.6.7.8-HpCDD; Diox7: OCDD; Diox8: 2.3.7.8-TCDF; Diox9: 1.2.3.7.8-PeCDF; Diox10: 2.3.4.7.8-PeCDF; Diox11: 1.2.3.4.7.8-HxCDF; Diox12: 1.2.3.6.7.8-HxCDF; Diox13: 1.2.3.7.8.9-HxCDF; Diox14: 2.3.4.6.7.8-HxCDF; Diox15: 1.2.3.4.6.7.8-HpCDF; Diox16: 1.2.3.4.7.8.9-HpCDF; Diox17: OCDF).

The hierarchical clustering analysis (using the Ward aggregation method and the 1- $\rho$  metric) performed on the same data set (Fig. 2) gave more insight into the relationships between the different congeners. Three groups of congeners appeared on the dendrogram. The first group included the PCB TEQ and Total TEQ (these two variables appearing highly correlated together) and some PCB congeners. The second group included the Dioxin TEQ and a very limited number of dioxin congeners. The last one corresponded to the PBDE congeners, which appeared non significantly correlated to the dioxin or PCB congeners. These results tend to indicate the possibility to predict the Dioxin TEQ and PCB TEQ on the basis of 2 or 3 congeners chosen among the above mentioned groups, the Total TEQ being very probably predicted by the same congeners as the PCB TEQ.

After these descriptive analysis, a step-by-step incremental multiple linear regression was performed on the same data set. Thus, the Total TEQ was attempted to be predicted by 8 congeners. The results (Figure 3) demonstrated the very good efficiency of this linear model ( $R^2 > 0.9999$ ). A noticeable observation concerned the first congener included in the model (PCB126), which permitted alone to reach a  $R^2 > 0.99$ . In order to ensure a more precise confidence level and to propose a model based both on some dioxin and PCB congeners, the step 5 of the analysis was retained. Indeed, the resulting linear model permitted to reach a  $R^2$  value  $> 0.999$  using two PCB (PCB126, PCB105) and three PCDD/PCDF (Diox01=2.3.7.8-TCDD, Diox02=1.2.3.7.8-PeCDD,

## Levels in feed and food

Diox10=2.3.4.7.8-PeCDF) congeners. The final predictive equation was the following one: Total TEQ = 1.280391\*[Diox01] + 1.278663\*[Diox02] + 0.611997\*[Diox10] + 0.102590\*[PCB126] + 0.000916\*[PCB105]. In a second phase, the PCB TEQ and Dioxin TEQ values were predicted on the basis of the same congeners, leading to two other equations (coefficient set) given in figure 3. As shown on the graphical representations of the predicted versus observed values, the proposed predictive models were very satisfactory.

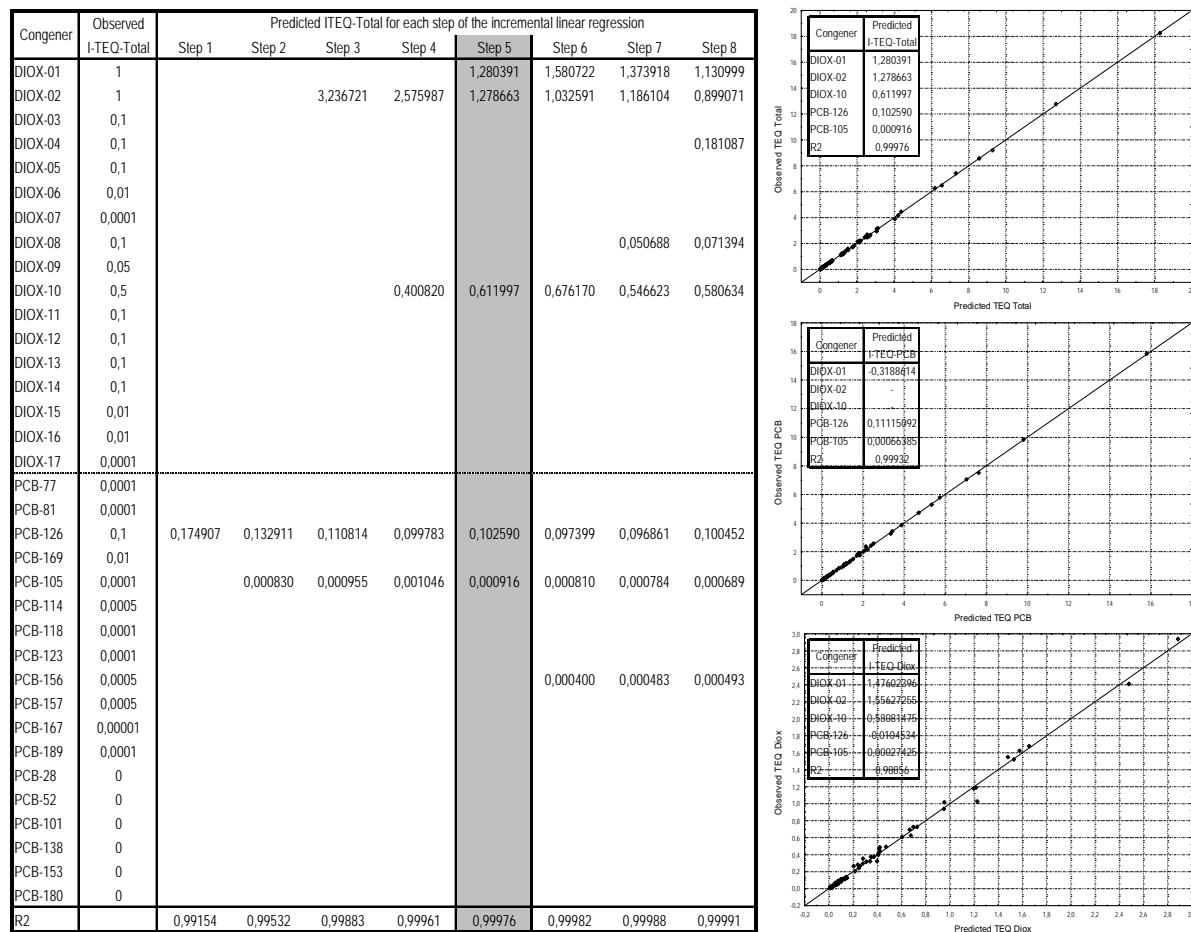


Figure 3: Result of the incremental step-by-step multiple linear regression analysis performed on the first data set (n=94 samples).

In order to validate the proposed predictive models, the obtained equations were applied to two independent set of data (n=60 and n=87 fish samples, respectively) collected in LABERCA in 2004 and 2005. In both cases, the results demonstrated an excellent correlation between the predicted and observed TEQ Total and TEQ PCB values ( $R^2 > 0.99$ ), and a very acceptable correlation for TEQ Dioxin value ( $R^2 > 0.98$ ). Figure 4 presents these results for the second validation data set. In order to precise the eventual error induced by this predictive model, all the trueness deviations between predicted and observed values were calculated. Figure 5 presents the results obtained for the second validation data set.

## Levels in feed and food

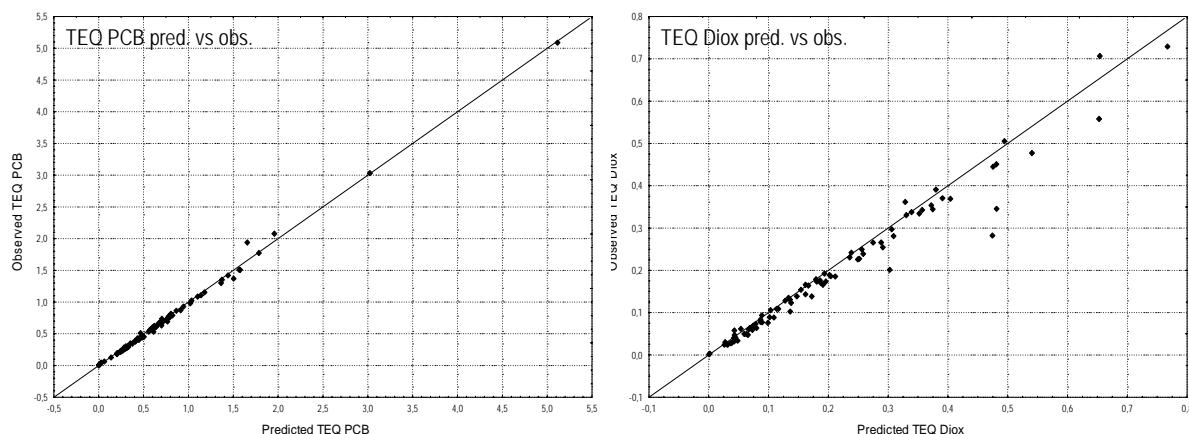


Figure 4: Correlation between observed and predicted TEQ PCB and TEQ Diox values obtained for the second validation data set (n=87 fish samples) on the basis of the model elaborated by the multiple linear regression.

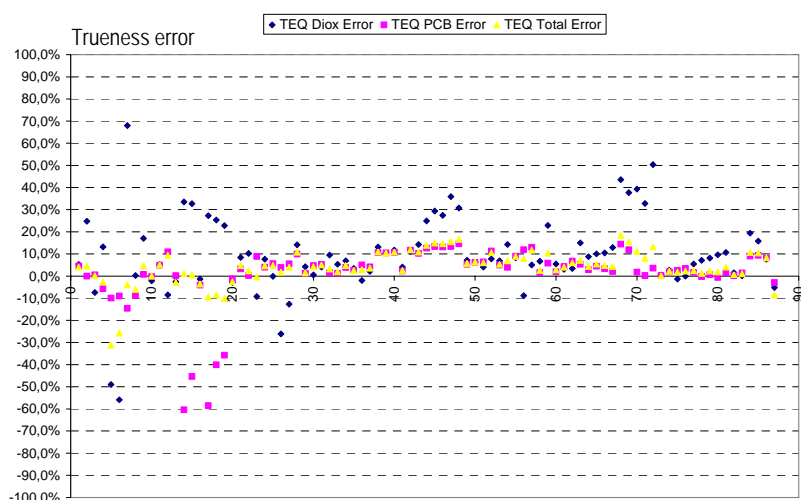


Figure 5: Trueness errors calculated between predicted and observed TEQ Total, TEQ PCB, and TEQ Diox values for the second validation data set (n=87 fish samples) on the basis of the model elaborated by the multiple linear regression.

Globally, all the error for Total TEQ and PCB TEQ were found to be lower than  $\pm 12\%$ . For Diox TEQ, these error were comprised between  $-10\%$  and  $+30\%$ . Regarding common analytic criteria in use in the case of screening methods, these results appeared very promising. Finally, this in-house validation do confirm the suitability of the proposed predictive model.

In a final validation step, 3 samples analysed within the framework of interlaboratory studies (IS) – including the international IS organised by the Norwegian Institute of Public Health in 2005 with more than 70 participants were considered. The consensual values determined for each sample after the IS were compared to the values predicted by the proposed model. The results (Table 1) demonstrated once again a very good efficiency of the predictive method, with error deviations lower than 10%.

Table 1: Comparison of the consensual and predicted TEQ values obtained for 3 samples analysed during interlaboratory studies.

Sample	Consensual Values			Predicted Values			Error (%)		
	TEQ Diox	TEQ PCB	TEQ Total	TEQ Diox	TEQ PCB	TEQ Total	TEQ Diox	TEQ PCB	TEQ Total
Trout EIL 2004	1,1	3,5	4,6	1,2	3,5	4,7	9,8%	-1,2%	1,5%
Hearing EIL 2005	0,8	0,9	1,7	0,8	0,9	1,7	-4,6%	0,7%	-1,8%
Cod Oil EIL 2005	2,0	11,8	13,8	2,2	12,9	15,1	7,6%	9,3%	9,1%

## Conclusion

A method was proposed and validated, which makes it to predict the Total TEQ, PCB TEQ and Dioxin TEQ values determined in the field of PCDD/PCDF/PCB analysis in fish, on the basis of only 5 concentration measurements (2 PCB and 3 PCDD/F congeners). This method is expected to reduce significantly the time and the cost of such analysis, almost for screening purpose, considering the possibility to purify the considered congeners in a single fraction and the time saved in term of data analysis (from 39 to 5 ion chromatograms).