Multivariate statistical survey of PCDD/F–profiles and PCB–patterns of soil samples from an industrialized urban area (Linz, Upper Austria)

Weiss, P., Gschmeidler, E., Schentz, H.
Federal Environmental Agency, Spittelauer Lände 5, A–1090 Vienna, Austria

Introduction

In the industrialized (chemical and steel industries), densely populated urban area of Linz pooled soil samples from grassland sites were investigated for their contents of PCDD/F (13 sites) and PCB (26 sites). Their specific PCDD/F–profiles and PCB–patterns were then subjected to a multivariate statistical survey to obtain additional information about the different sources of pollution.

Material and methods

Detailed information on the chemical–analytical methods used and on the results is given in[1].
At first, correlation analyses were carried out in order to determine the correlation between the concentrations of the various investigated PCB–congeners (PCBs 28, 52, 101, 138, 153, 180), the PCDD/F–homologues and between these two groups of pollutants.
To allow a multivariate processing of the data, the contents of the various PCB–congeners and the PCDD/F–homologues were expressed in per cent of the sum of PCBs and the sum of PCDD/Fs, respectively, per site. These percentage values were then standardized, the mean was defined to be 0, the standard deviation 1.
In order to define groups of sites showing similar PCB–patterns or PCDD/F–profiles, the standardized data sets were subjected to a cluster analysis according to the average–linkage– and the ward–algorithm–method. The results given below were obtained by applying the average–linkage–method, which, compared to the ward algorithm, has the advantage of identifying freak values more clearly[2]. Furthermore, to assist this process, the squared Euclidian distance was used in the clustering process. The number of clusters to be formed was varied and determined according to the dendrograms and the respective F–values. In addition, the T–values were calculated to facilitate interpretation.
The clusters thus obtained were then subjected to a canonical discriminant analysis, a technique for dimension–reduction related to principal components.
For the PCDD/Fs further variants of the cluster analysis were calculated in order to get more information on the structure of the data. Based on the results of the correlation analysis, individual PCDD/F–homologues were excluded from the profiles and the cluster analysis calculated on the basis of the remaining incomplete profiles.

Results

The soil samples showed PCDD/F concentrations between 1.6 and 14.4 ng TE/kg (I–TEF).
On the whole, high correlations (significant at a 0.001 level) could be established between the PCDF–homologues, whereas only few significant correlations could be found between the PCDD–homologues. It is noteworthy that more significant correlations could be established between the PCDD– and PCDF–homologues than between the individual PCDD–homologues. O8CDD constituted a special case. Apart from H7CDD no correlation could be established with any other PCDD/F–homologue. This result indicated the strong influence of O8CDD on the difference between the PCDD/F–profiles, a fact which was subsequently confirmed by cluster analysis.
Fig. 1 shows the fusion of the individual site–specific PCDD–profiles by means of cluster analysis. Most of the sites of cluster 1 showed urban background concentrations. Their profiles were characterized by a comparatively high amount of O8CDD.

Compared to the other sites, cluster 2 showed high amounts of T4CDF, P5CDF and H6CDF (fig. 1). The profile of the sites 2 and 16 is broadly similar to the PCDD/F–emission profile of a sintering plant. The sites of this cluster are situated in the industrial area (site 2) and on an exposed (i.e. emission receiving) slope which faces the industrial area (sites 15, 16) and, thus, exhibit a higher amount of pollutants. Evaluations of the emissions of various industrial plants in the Linz area clearly identified the sintering plant as the main source of PCDD/F–emissions. Apart from higher amounts of T4CDF, P5CDF and H6CDF, the sites 8 and 11 (cluster 3) also showed higher O8CDD values within the profiles (fig. 1). The profile of this cluster represents a transitional form between cluster 1 and cluster 2. Had four instead of five clusters been formed, both sites would have been attributed to cluster 2. Situated near a former hospital waste incineration plant, site 8 ranked among the comparatively more polluted sites; site 11 is situated in the industrialized area.
Since the sites 5 and 14 represent real exceptions on account of their profile, they were separated by the cluster analysis. Site 14, which, in comparison, shows an average degree of pollution, is situated outside the urban, industrialized area. Its profile exhibited a high amount of T4CDD which indicates another source of pollution (fig. 1).
Site 5 (cluster 5), which is comparatively less polluted, showed a homogenous profile with similar amounts of H7CDD to O8CDF (fig. 1).
The ten dimensions given by the homologues were reduced to three dimensions (canonical variables, "can" in fig. 2) by means of linear combinations of the canonical discriminant analysis. As the three–dimensional diagramme (fig. 2) reflects 98 % of the variation of the individual data, the data structure is rendered almost completely.
No significant positive correlation could be established between the PCDD/F–homologues and the PCB–congeners. P5CDD even showed a slight negative correlation with two PCB congeners and the sum of PCBs. This hints at different sources of pollution for the investigated soils. High correlations (significant at a 0.001 level) could be established between the higher chlorinated PCBs (101, 138, 153, 180). PCB 28 showed only low correlations with PCB 138 and 153. Since PCB 52 was below the detection limit in most of the samples, no statement about correlation can be made.
The pattern of the sites of cluster 1 which are spread over the whole investigation area exhibit comparatively low amounts of PCB 138 and comparatively high amounts of PCB 153 (fig. 4). All five sites were only slightly polluted.
Compared to other clusters, the eleven sites of cluster 2 which were only slightly polluted and which are spread over the whole investigation area, too, showed neither very high nor very low amounts of individual PCBs (fig. 4).
The comparatively higher polluted sites of cluster 3 exhibited lower amounts of PCB 28 and PCB 101 and higher amounts of PCB 180 (fig. 4). Sites 5 and 11 are close to each other and adjacent to or within the industrial area, whereas site 34 is further away from these two sites.
Cluster 4 contains higher polluted sites (situated closer to the industrial area) as well as lower polluted sites, their patterns being characterized by low amounts of PCB 101 and high amounts of PCB 138 (fig. 4). All six sites are partly adjacent to and west of the industrial area. Their contamination might therefore be due to the same sources.
On account of the high amount of PCB 52 in the pattern, the reference site (cluster 5) which is situated far away from the conurbation was separated by cluster analysis (fig. 4).
The six dimensions given by the investigated PCB–congeners were reduced to three dimensions (canonical variables, "can" in fig. 3) by means of linear combinations of the canonical discriminant analysis. As 95 % of the variation of the individual data could be explained by means of the three linear combinations, the data structure is rendered almost completely. The multivariate test for differences between clusters was significant at the 0.0001 level.

[1] Weiss P., Riss, A. Schadstoffe im Raum Linz, Teil B: Bodenuntersuchung von Grünland und Waldböden im Raum Linz auf Schwermetalle, polyzyklische aromatische Kohlenwasserstoffe und chlorierte Kohlenwasserstoffe. Umweltbundesamt Wien, Monographien, Bd. 20, 1992

[2] Backhaus, K., Erichson, B., Plinke, W., Weiber, R. Multivariate Analysenmethoden. 6th ed., Springer Verlag, Berlin, 1990

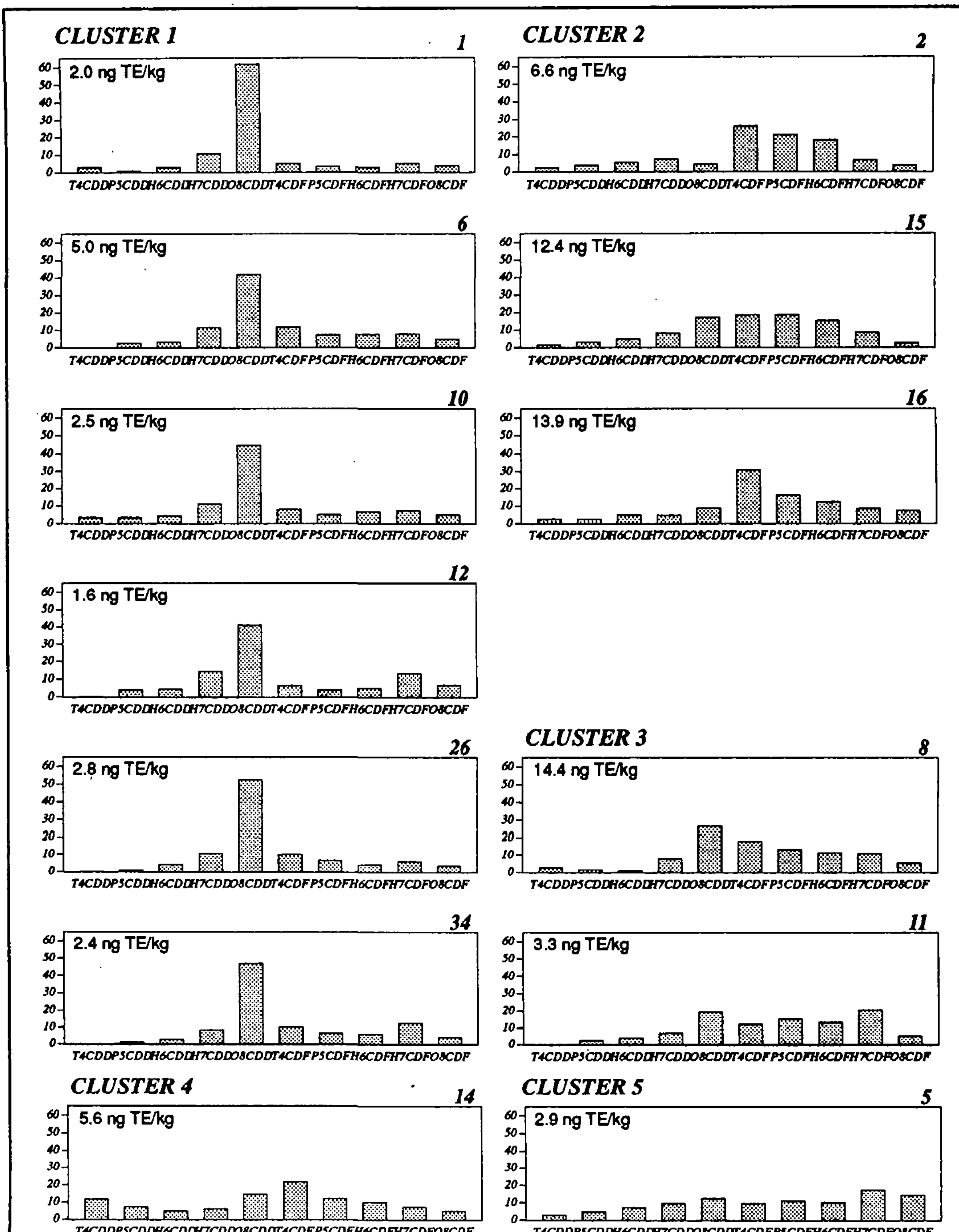

fig. 1.: PCDD/F–profiles of grassland soils in the Linz area (Upper Austria)
Sites arranged according to clusters.
Values given in % of the sum of PCDDs and PCDFs/site.

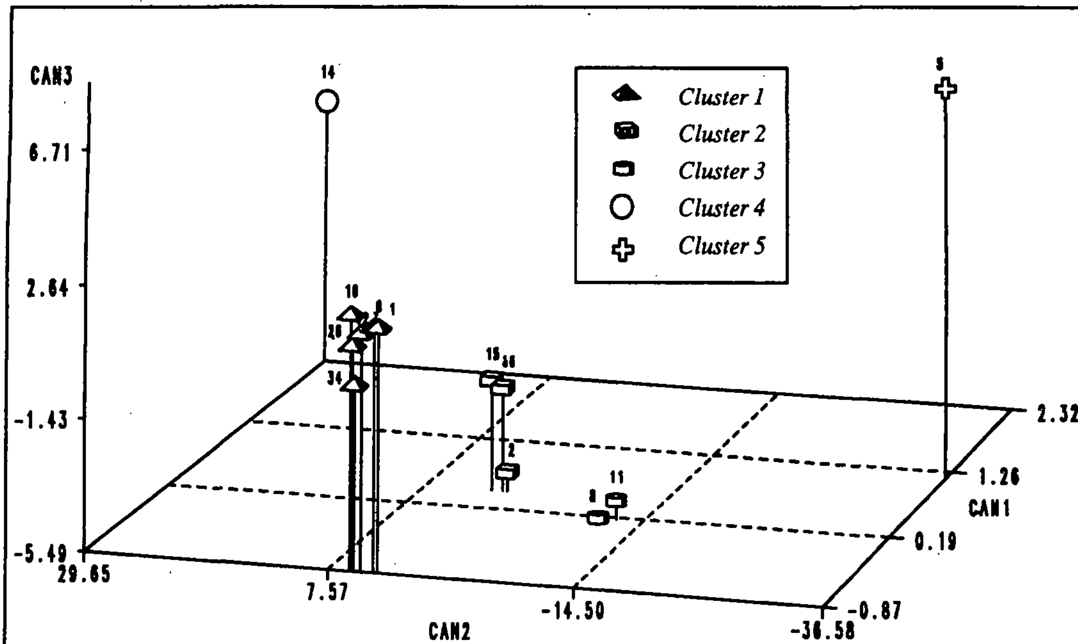

fig. 2: Site–specific PCDD/F–profiles joined to clusters in three–dimensional representation resulting from canonical discriminant analysis

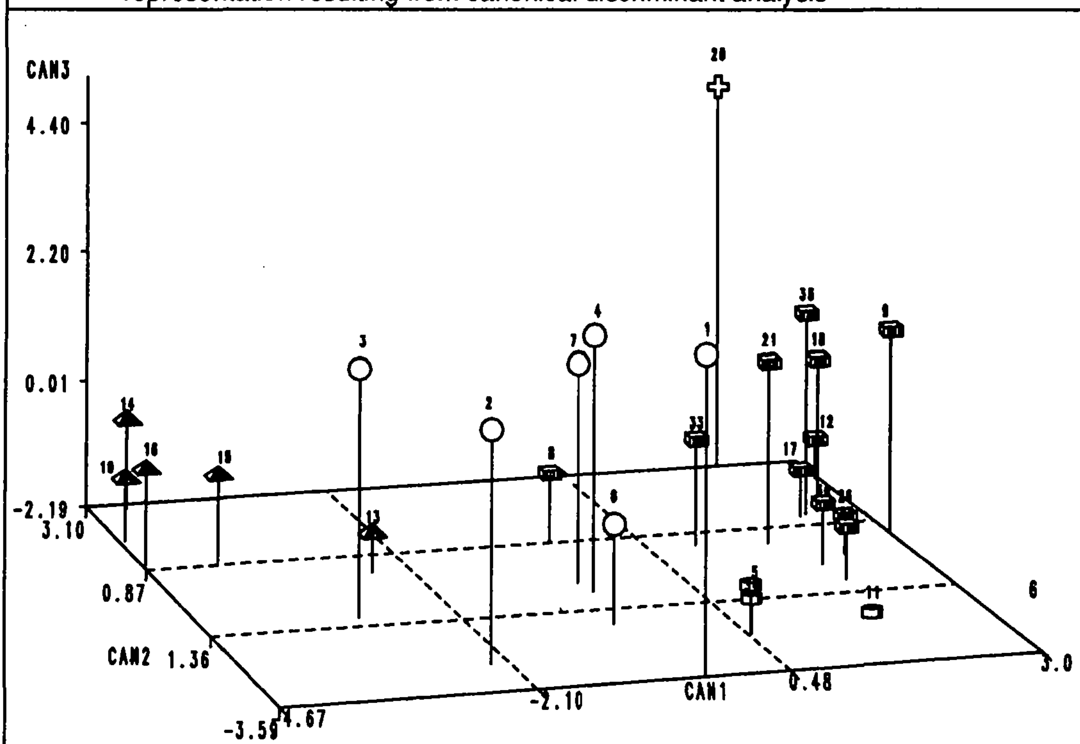

fig. 3: Site–specific PCB–patterns joined to clusters in three–dimensional representation resulting from canonical discriminant analysis

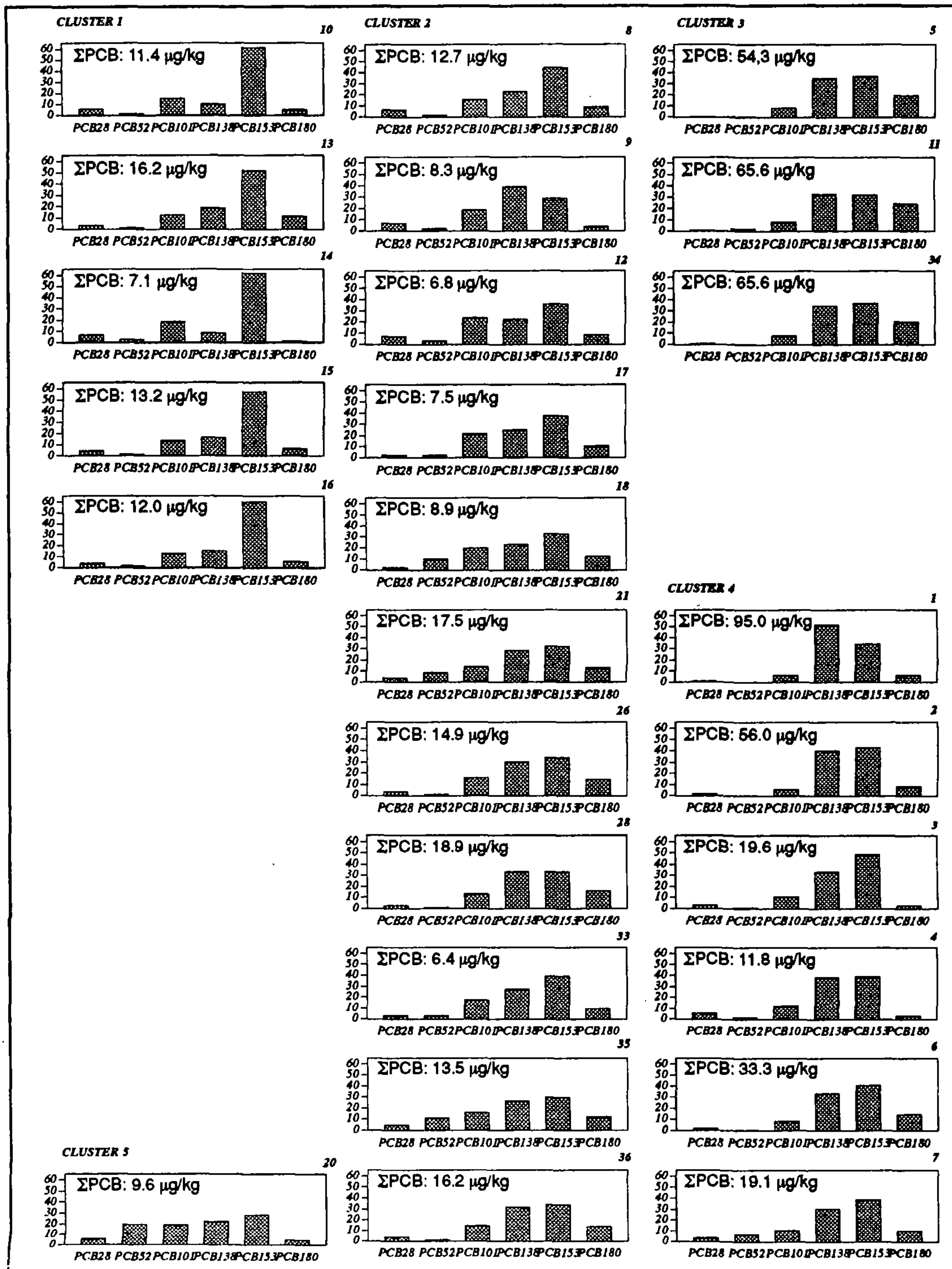

fig. 4: PCB–pattern of grassland soils in the Linz area (Upper Austria)
Sites arranged according to clusters. Values given in % of the sum of PCBs/site