TEXTE
39/2019

Expert Report on the Usability of Data from the POP Database in the Field of 'Technosphere'

Final report

German Environment Agency

Umwelt Bundesamt

TEXTE 39/2019

Project No. (FKZ) 97932
Report No. FB000074

Expert Report on the Usability of Data from the POP Database in the Field of 'Technosphere'

Final report

by

Heidelore Fiedler
Örebro University, School of Science and Technology, MTM Research Centre, Örebro, Sweden

On behalf of the German Environment Agency

Imprint

Publisher
Umweltbundesamt
Wörlitzer Platz 1
06844 Dessau-Roßlau
Tel: +49 340-2103-0
Fax: +49 340-2103-2285
info@umweltbundesamt.de
Internet: www.umweltbundesamt.de

/umweltbundesamt.de
/umweltbundesamt

Report performed by:
Örebro University
School of Science and Technology, MTM Research Centre
SE-701 82 Örebro
Sweden

Report completed in:
January 2019

Edited by:
Section IV2.1 Information Systems Chemical Safety
Dr. Thomas Gräff (statistics and R programming), Dr. Gerlinde Knetsch (specialist support), Dr. Ralph Stephanowitz (specialist support)

Publication as pdf:
http://www.umweltbundesamt.de/publikationen

ISSN 1862-4804

Dessau-Roßlau, April 2019

Abstract: Expert Report on Usability of Data from the POP Database in the Field of Technosphere
Expert Report on the Usability of Data from the POP Database in the Field of 'Technosphere'

The present report provides a quality-assured database for the technosphere compartment as part of the further development of content and technology of the congeners profile POP dioxin database into a knowledge-based system.

Based on the scientific know-how of a technical expert, sound proposals for the development of a flexible evaluation instrument were developed for a total of 278 samples from the technosphere (sources: research projects, EURL-Freiburg, evaluation of original scientific literature).

The evaluation focused on the inventory and classification of these samples in terms of primary and secondary sources and their classification into further subgroups taking into account chemical and physical characteristics, the collection of new data and the raising the data quality of the already acquired data. The classification scheme developed here also allows the classification of future samples.

For the recognition of typical patterns in the data, R programs for cluster analysis of the groups of polychlorinated dibenzo-*p*-dioxins and polychlorinated dibenzofurans (PCCD/PCDF), dioxin-like polychlorinated biphenyls (dl-PCB) and non-dioxin-like polychlorinated biphenyls (ndl-PCB) were developed and successfully used. These programs can also be used to compare profiles of contaminated samples with existing profiles in the database.

They are thus an important tool for investigating and solving dioxin and PCB problems.

Kurzbeschreibung: Gutachten zur fachlichen Nutzbarkeit der Daten aus der POP-Dioxindatenbank für den Bereich Technosphäre

Das vorliegende Gutachten stellt im Rahmen der inhaltlichen und technologischen Weiterentwicklung der Kongenerenprofil-POP-Dioxin-Datenbank hin zu einem wissensbasierten Systems eine qualitätsgesicherte Datenbasis für das Kompartiment Technosphäre bereit.

Basierend auf dem wissenschaftlichen Know-how einer Fachexpertin wurden für insgesamt 278 Proben aus der Technosphäre (Quellen: Forschungsprojekte, EURL-Freiburg, Auswertung wissenschaftlicher Originalliteratur) fundierte Vorschläge zur Entwicklung eines flexiblen Auswerteinstruments erarbeitet.

Schwerpunkte des Gutachtens waren die Bestandsaufnahme und Klassifizierung dieser Proben hinsichtlich primärer und sekundärer Quellen und deren Einordnung in weitere Untergruppen unter Berücksichtigung chemischer und physikalischer Merkmale, die Erfassung neuer Daten und die Hebung der Datenqualität der bereits akquirierten Datenbestände. Das hier entwickelte Klassifizierungsschema ermöglicht auch die Einordnung zukünftiger Proben.

Für die Erkennung typischer Muster in den Daten wurden R-Programme zur Clusteranalyse der Gruppen polychlorieret Dibenzo-*p*-dioxine und polychlorierte Dibenzofurane (PCDD/PCDF), dioxinähnliche polychlorierte Biphenyle (dl-PCB) und nicht-dioxinähnliche polychlorierte Biphenyle (ndl-PCB) entwickelt und erfolgreich eingesetzt.

Diese Programme können ebenfalls dazu genutzt werden, um Kongenerenprofile kontaminierter Proben mit bereits in der Datenbank vorhandenen Profilen zu vergleichen. Sie sind somit ein wichtiges Werkzeug zur Untersuchung und Aufklärung von Dioxin- und PCB-Skandalen.

Table of content

List of figures

List of tables

List of abbreviations

BfR	Federal Institute for Risk Assessment
BVL	Federal Office of Consumer Protection and Food Safety
dl	Dioxin-like
Cl	Chlorine
CNP	Chloronitrophen
d.m.	Dry matter
EURL	European Reference Laboratory (here for the European Union Reference Laboratory for Halogenated Persistent Organic Pollutants (POPs) in Feed and Food) located at the Chemisches und Veterinäruntersuchungsamt (CVUA) in Freiburg, Germany
HCA	Hierarchical cluster analysis
ndl	Non-dioxin-like
NRL	National reference laboratory
P&P	Pulp and paper
PCB	Polychlorinated biphenyls
PCDD/PCDF	Polychlorinated dibenzo-*p*-dioxins/polychlorinated dibenzofurans
PCP	Pentachlorophenol
POPs	Persistent organic pollutants
prim	Primary
RIKILT	Rijks Kwaliteitsinstituut voor Land-en Tuinbouwproducten (Imperial Quality Control of Agricultural and Horticultural Products), located in Wageningen, the Netherlands
sec	Secondary
TEF	Toxicity equivalency factor
TEQ	Toxic equivalent
UBA	Umweltbundesamt (German Environment Agency)
WHO	World Health Organization

List of congeners and homologs

Polychlorinated dibenzo-*p*-dioxins and polychlorinated dibenzofurans

Substitution	Homolog	Abbrev-iation1	Abbrev-iation2	Abbrev-iation3	WHO_{2005}-TEF
2,3,7,8	Tetrachlorodibenzo-*p*-dioxin	Cl_4DD	TCDD	TCDD	1
1,2,3,7,8	Pentachlorodibenzo-*p*-dioxin	Cl_5DD	PnCDD	PnCDD	1
1,2,3,4,7,8	Hexachlorodibenzo-*p*-dioxin	Cl_6DD	HxCDD	HxCDD1	0.1
1,2,3,6,7,8	Hexachlorodibenzo-*p*-dioxin	Cl_6DD	HxCDD	HxCDD2	0.1
1,2,3,7,8,9	Hexachlorodibenzo-*p*-dioxin	Cl_6DD	HxCDD	HxCDD3	0.1
1,2,3,4,6,7,8	Heptachlorodibenzo-*p*-dioxin	Cl_7DD	HpCDD	HpCDD	0.01
1,2,3,4,6,7,8,9	Octachlorodibenzo-*p*-dioxin	Cl_8DD	OCDD	OCDD	0.0003
2,3,7,8	Tetrachlorodibenzofuran	Cl_4DF	TCDF	TCDF	0.1
1,2,3,7,8	Pentachlorodibenzofuran	Cl_5DF	PnCDF	PnCDF1	0.03
2,3,4,7,8	Pentachlorodibenzofuran	Cl_5DF	PnCDF	PnCDF2	0.3
1,2,3,4,7,8	Hexachlorodibenzofuran	Cl_6DF	HxCDF	HxCDF1	0.1
1,2,3,6,7,8	Hexachlorodibenzofuran	Cl_6DF	HxCDF	HxCDF2	0.1
1,2,3,7,8,9	Hexachlorodibenzofuran	Cl_6DF	HxCDF	HxCDF3	0.1
2,3,4,6,7,8	Hexachlorodibenzofuran	Cl_6DF	HxCDF	HxCDF4	0.1
1,2,3,4,6,7,8	Heptachlorodibenzofuran	Cl_7DF	HpCDF	HpCDF1	0.01
1,2,3,4,7,8,9	Heptachlorodibenzofuran	Cl_7DF	HpCDF	HpCDF2	0.01
1,2,3,4,6,7,8,9	Octachlorodibenzofuran	Cl_8DF	OCDF	OCDF	0.0003

Dioxin-like PCB			
Substitution	**Homolog**	**Abbreviation**	**WHO_{2005}-TEF**
Non-ortho substituted PCB			
3,3',4,4'	Tetrachlorobiphenyl	PCB 77	0.0001
3,4,4',5	Tetrachlorobiphenyl	PCB 81	0.0003
3,3',4,4',5	Pentachlorobiphenyl	PCB 126	0.1

3,3',4,4',5,5'	Hexachlorobiphenyl	PCB 169	0.03
Mono-ortho substituted PCB			
2,3,3',4,4'	Pentachlorobiphenyl	PCB 105	0.00003
2,3,4,4',5	Pentachlorobiphenyl	PCB 114	0.00003
2,3',4,4',5	Pentachlorobiphenyl	PCB 118	0.00003
2',3,4,4',5	Pentachlorobiphenyl	PCB 123	0.00003
2,3,3',4,4',5	Hexachlorobiphenyl	PCB 156	0.00003
2',3,3',4,4',5	Hexachlorobiphenyl	PCB 157	0.00003
3,3',4,4',5,5'	Hexachlorobiphenyl	PCB 167	0.00003
2,3,3',4,4',5,5'	Heptachlorobiphenyl	PCB 189	0.00003

Non-dioxin-like PCB (ndl-PCB, six indicator PCB)

Substitution	**Homolog**	**Abbreviation**
2,4,4'	Trichlorobiphenyl	PCB 28
2,2',5,5'	Tetrachlorobiphenyl	PCB 52
2,2',4,5,5'	Pentachlorobiphenyl	PCB 101
2,2',3,4,4',5'	Hexachlorobiphenyl	PCB 138
2,2',4,4',5,5'	Hexachlorobiphenyl	PCB 153
2,2',3,4,4',5,5'	Heptachlorobiphenyl	PCB 180

Summary

This report was prepared to link the existing POP-Dioxin Database of the German Federal Environment with results from food and feed incidents in relation to 'dioxins and PCB' occurring in the European Union. The European Reference Laboratory for Halogenated Persistent Organic Pollutants (POPs) in Feed and Food is frequently asked to provide assistance with the identification of the source of the contamination; *e.g.* when limit values were exceeded.

An assessment was undertaken as to the usability for pattern recognition of chemical analytical data from the POP-Dioxin Database and from the datasets present at the food/feed control laboratories. In order to establish possible correlations between contaminated samples and their possible sources, datasets from the literature or published reports that had shown such linkages were included.

At both ends and before the start of this project, expert teams were established. The German Federal Environment Agency´s (UBA) team included experts from Section IV 2.1 Information Systems Chemical Safety with special expertise in POPs data management and a specialist on statistics and R programming. Within the network of the European and National Reference Laboratories for Food and Feed, a core working group (CWG) 'Dioxin pattern' was established. Both groups worked independently but in close cooperation between the different institutions and experts, including physical meetings between the UBA and the EURL and bilateral meetings with the technical expert. The work did undergo several iterations until the final version was found.

The first and very important step was the data selection and quality control as to the analytical data such as mass concentrations of the chemical compounds in the different samples and datasets. A common format to store and present each sample in one row was agreed. Rapidly, it became clear that not all information was available for all samples/datasets. Special attention was paid to strictly harmonizing the information.

Through this study, linkages were assessed that may exist between sources of dioxins and related compounds and environmental compartments but also with biota and humans. It is well known that polychlorinated dibenzo-*p*-dioxins (PCDD/PCDF) and polychlorinated dibenzofurans (PCDF) are man-made chemicals that do not serve a useful purpose but that their formation and release is associated with many activities. Most of the sources are anthropogenic and the majority are either from chemical synthesis processes or from thermal processes. A subset of the 'synthesis' are the naturally occurring processes in geogenic sources such as ball clay. Polychlorinated biphenyls (PCB) have been included into this assessment from a needs assessment for the following reasons:

- Limit values exist for dioxin-like PCB (dl-PCB) and non-dioxin-like PCB (ndl-PCB);
- The toxicity equivalent concept includes PCDD/PCDF and PCB;
- PCDD, PCDF and PCB can be formed as unintentional contaminants in the same processes (especially thermal ones);
- PCB are a source of PCDF.

Such information on sources of PCDD, PCDF and PCB were named 'primary sources' or 'prim' and many datasets were contributed from the POP Dioxin Database. These datasets did undergo further classification to include the fields of the technosphere such as technical products,

emissions or residues. An example for the classification into large Sources, Groups, and the Levels 1-3 is shown in Table S1 below.

Table S1: Example cases for multilevel classification of primary-source samples from the technosphere

#	Source	Group	Level_1	Level_2	Level_3	CaseContext	Sample
205	prim	Synthesis	Cl other	Aliphatics	Dye	Generic	Phthalocyanine copper
135	prim	Synthesis	Cl other	Elemental	P&P	Generic	non-wood pulp, PM4, after C-stage
144	prim	Synthesis	Cl other	Elemental	Residue Cl ind	Generic	Residue, Cl industry
45	prim	Synthesis	Cl-precursor	Cl-benzene	ClBz	Generic	Chlorobenzene, China
6	prim	Synthesis	Cl-precursor	Cl-benzene	PCB	Generic	Clophen A30
16	prim	Synthesis	Cl-precursor	Cl-benzene	PCB	Generic	Kanechlor KC-400
33	prim	Synthesis	Cl-precursor	Cl-benzene	PCB	Generic	Aroclor 1260
35	prim	Synthesis	Cl-precursor	Cl-phenol	PCP	Generic	PCP
54	prim	Synthesis	Cl-precursor	Cl-phenol	Phenoxy	Generic	2,4-D/2,4,5-T
246	prim	Thermal	Controlled	Small	Residue	Generic	Chimney ash
225	prim	Thermal	OBurn	Accident	Residue	DUS airport fire	Soot, DUS Terminal
240	prim	Thermal	OBurn	Fuel dep	Residue	Open burn	Ash, open burning
68	prim	Nat'l formation	Geogenic	Clay		Ball clay	Ball clay USA
71	prim	Nat'l formation	Geogenic	Clay		Ball clay	Mabele

The same approach was taken for the datasets of samples where the PCDD, PCDF or PCB content was transferred; these sources were named 'secondary sources' or 'sec' and many datasets for biotic environments were contributed from the NRLs or EURL or the literature where known 'dioxin accidents' were described. Further classification included abiotic, biotic and unknown matrices. It shall be noted that (almost all) food samples were biotic but feed samples can originate from abiotic environments; e.g. clay, minerals. Within the secondary sources, there were also samples detected that could not be assigned any Group or all Levels. An example is shown in Table S2 below.

Table S2: Example cases for multilevel classification of secondary-source samples from the technosphere

#	Source	Group	Level_1	Level_2	Level_3	CaseContext	Sample
270	sec	Abiotic	Commodity	Synthetic	Construction	UBA PCB report	Sealant
138	sec	Abiotic	Commodity	Synthetic	Dye	UBA PCB report	Paint ceiling tile
214	sec	Abiotic	Sink	Soil	Impacted	Generic	Soil
311	sec	Biotic	Animal	Food	Aquatic	Ball clay	Catfish
94	sec	Biotic	Animal	Food	Terrestrial	Ball clay	Turkey meat
338	sec	Biotic	Animal	Food	Terrestrial	Citrus pulp	Butter
319	sec	Biotic	Animal	Food	Terrestrial	Belgian PCB case	Chicken meat
73a	sec	Biotic	Animal	Humans	Tissue	Ball clay	Human milk
84	sec	UKN	UKN	Feed		Chile pork	ZnO mix
343	sec	UKN	UKN	Feed		Belgian chicken	Feed mix

With respect to the compounds (congeners), the following numeric information with respect to mass concentration was included into the datasets (if available). In total, there are 66 columns containing numeric information:

- **21 columns for PCDD/PCDF**
  17 2,3,7,8-substituted PCDD/PCDF congeners + 3 TEQs (WHO_{2005}-TEQ_{PCDD}, WHO_{2005}-TEQ_{PCDF}, WHO_{2005}-$TEQ_{PCDD/PCDF}$) + ΣPCDD/PCDF(17)
- **16 columns for dl-PCB**
  12 dl-PCB congeners + 3 TEQs (WHO_{2005}-$TEQ_{no\text{-}PCB}$, WHO_{2005}-$TEQ_{mo\text{-}PCB}$, WHO_{2005}-TEQ_{PCB},) + ΣPCB(12)
- **7 columns for ndl-PCB**
  6 indicator PCB congeners + ΣPCB(6)
- **11 columns for PCDD/PCDF homologs,**
  Homologs for PCDD/PCDF (*tetra* through *octa*, Cl_4 through Cl_8) + Σ(Cl_4-Cl_8)
- **11 columns for PCB homologs**
  Homologs for PCB (*mono* through *deca*, Cl_1 through Cl_{10}) + Σ(Cl_1-Cl_{10}).

Once the datasets are harmonized and have passed quality control, they can be assessed using univariate or multivariate statistical methods. Depending on the availability of quantitative information, different types of 'profiles' can be generated from the data as follows.

Profile 1: Contribution of congeners to the sum within a sub-group of POPs

Profile 1a: Seventeen PCDD/PCDF
Concentrations of the 17 PCDD/PCDF congeners are normalized to the sum of these 17 congeners and the contribution of each congener is processed.

Profile 1b: Twelve dioxin-like PCB
Concentrations of the 12 dl-PCB congeners are normalized to the sum of these 12 congeners and the contribution of each congener is processed.

Profile 1c: 29 dioxin-like POPs having a assigned a TEF
Concentrations of all 29 congeners contained in the TEF scheme are normalized to the sum of these 29 congeners.

Profile 1d: Six indicator PCB
Concentrations of the six indicator PCB congeners are normalized to the sum of these six congeners and the contribution of each congener is processed.

Profile 1e: Homologs of PCDD/PCDF
Concentrations of the five homologs from Cl_4 to Cl_8 of PCDD and PCDF for are normalized to the sum of the homologs and the contribution of each homolog is processed.

Profile 1f: Homologs of PCB
Concentrations of the ten homologs of PCB are normalized to the sum of the homologs and the contribution of each homologue to the sum is processed.

Profile 2: Contribution of PCDD/PCDF congener to its homolog

The concentration of each homolog is divided by the concentration of the respective 2,3,7,8-subsituted congener. There are a total of 15 relations (quotient for OCDD and OCDF equals 1 since the congener is also the homolog and therefore cannot be used for this profile).

Profile 3: Profiles using the TEF concept

Each congener is multiplied by its toxicity equivalency factor using WHO_{2005}-TEFs and the contribution of each weighted congener to the WHO_{2005}-TEQ is processed.

All of the above profiles can be visualized using commercially available programmes or other tools. We have included into this assessment a combination of heatmaps and dendrograms from hierarchical cluster analysis to identify the similarity of the profiles and of the samples. The graphical outputs were generated using the R programme and the R scripts are made available in the Appendix of this report.

Conclusions

A dataset of 278 samples has been selected and applied to quantitative results for PCDD/PCDF and PCB. It was found that a strict structure needs to be established and maintained in order to use the data in further applications including recognition of patterns and profiles of environmental and food/feed samples.

In this study, a structure to present the data is proposed containing:

- Three large source groups each as to primary or secondary sources complemented by three levels of further descriptive sample characterization;
- Inclusion of the references for each dataset;

- Organization of numeric data according to 17 congeners of PCDD/PCDF, 12 congeners of dioxin-like PCB and six congeners of indicator PCB (ndl-PCB);
- Homologs of PCDD/PCDF and PCB;
- Further quantitative information includes the calculated values for the toxic equivalents or the sums of groups of the chemicals (ΣPCDD/PCDF$_{17}$, ΣPCB$_{12}$, ΣPCB$_{6}$ or ΣPCDD/PCDF$_{Homologs}$ and ΣPCB$_{Homologs}$

The structures are flexible and can be amended by new datasets (samples) but also allow for further classification at all three levels; and even to include more chemicals.

Monovariate and multivariate statistics and visualization tools have been developed and applied using the quantitative information in the database. The datasets have been tested between each other but also with 'unknown' samples to 'match' the profiles for further assessment. Although very good correlations were found and the usefulness of the profile matching was shown, it will need expertise to interpret the results and a warning be expressed to use these correlations as the only proof of the hypothesis.

The use of this dataset will be for future assessments in two directions by the two authorities:

1. The German UBA :
   a) Introduce the proposed structure to the samples in the Dioxin POP database
   b) Optimize the definitions and expand or change the characterization if necessary;
   c) Apply to the samples of types other than sources and food/feed; *e.g.,* ambient air, soil, sediment;
   d) Explore the usefulness for further studies on transfers between environmental compartments and between environment and humans.
2. EURL/NRLs
   a) Optimization of the datasets and descriptions for food and feed incidents/accidents;
   b) Identify correlations between similar samples according to the profiles
   c) Assign sources that have contaminated the feed or food samples.

Zusammenfassung

Diese Studie wurde durchgeführt, um die POP-Dioxin Datenbank der Umweltbundesamtes (Deutschland) mit Daten und Ergebnissen von Lebensmittel und Futtermitteluntersuchungen in 'Dioxin- und PCB-Skandalen' in der Europäischen Union zu verbinden. Das EU-Referenzlabor für halogenierte persistente organische Schadstoffe (POP) in Futtermitteln und Lebensmitteln wird häufig um Unterstützung bei der Ermittlung der Quellen von Kontaminationen gebeten, z.B. wenn Grenzwerte überschritten wurden.

Es wurde eine Bewertung der Verwendbarkeit für die Mustererkennung von chemischen Analysendaten aus der POP-Dioxin-Datenbank mit den bei Futtermittel- und Lebensmittel-kontrolllabors vorhandenen Datensätzen vorgenommen. Um mögliche Korrelationen zwischen kontaminierten Proben und ihren möglichen Quellen festzustellen, wurden Datensätze aus der Literatur oder aus veröffentlichten Berichten, die solche Verknüpfungen aufgezeigt hatten, eingeschlossen.

Auf beiden Seiten und vor Beginn dieses Projekts wurden Expertenteams gebildet. Zum Team des Umweltbundesamtes (UBA) gehörten Experten aus dem Fachgebiet IV 2.1 Informations-systeme Chemikaliensicherheit mit besonderer Expertise im POP-Datenmanagement sowie ein Spezialist für Statistik und R-Programmierung. Innerhalb des Netzwerks der europäischen und nationalen Referenzlaboratorien für Lebens- und Futtermittel wurde eine Kernarbeitsgruppe (Core working group) ‚Dioxin pattern' (‚Dioxinprofile') eingerichtet. Beide Gruppen arbeiteten unabhängig voneinander, jedoch in enger Zusammenarbeit sowohl zwischen den verschiedenen Institutionen und als auch den Experten, einschließlich physischer Treffen zwischen dem UBA und EURL sowie bilateralen Treffen mit dem technischen Experten. Die Arbeit wurde mehreren Iterationen unterzogen, bis die endgültige Version gefunden wurde.

Der erste und ein sehr wichtige Schritt war die Auswahl der Daten und die Qualitätskontrolle hinsichtlich der analytischen Messwerte wie Massenkonzentrationen der chemischen Verbindungen in den verschiedenen Proben und Datensätzen. Es wurde ein gemeinsames Format zum Speichern und Präsentieren jeder Probe in einer Zeile vereinbart. Schnell wurde klar, dass nicht alle Informationen für alle Proben/Datensätze verfügbar waren. Besonderes Augenmerk wurde auf die strikte Harmonisierung der Informationen gelegt.

Durch diese Studie wurden Verknüpfungen bewertet, die möglicherweise zwischen Dioxin-quellen und verwandten Verbindungen und Umweltkompartimenten, aber auch zu Biota und Menschen bestehen. Es ist bekannt, dass es sich bei den polychlorierten Dibenzo-*p*-dioxinen (PCDD/PCDF) und den polychlorierten Dibenzofuranen (PCDF) um synthetische Chemikalien handelt, die keinen nützlichen Zweck erfüllen, deren Bildung und Freisetzung jedoch mit vielen Aktivitäten verbunden sind. Die meisten Quellen sind anthropogen und die Mehrheit stammt entweder aus chemischen Syntheseverfahren oder aus thermischen Prozessen. Eine Untermenge der "Synthese" sind die natürlich vorkommenden Prozesse in geogenen Quellen wie in Tonmineralien (‚ball clay', Kaolinit) gefunden wurden. Polychlorierte Biphenyle (PCB) wurden aus den folgenden praktischen Gründen in die Bewertung einbezogen:

- Es gibt Grenzwerte für dioxinähnliche PCB (dl-PCB) und nicht dioxinähnliche PCB (ndl-PCB);
- Das Konzept der Toxizitätsäquivalente umfasst PCDD/PCDF und PCB;
- PCDD, PCDF und PCB können als unbeabsichtigte Verunreinigungen in denselben (insbesondere thermischen) Prozessen gebildet werden;
- PCB sind eine Quelle für PCDF.

Die Informationen zu PCDD-, PCDF- und PCB-Quellen wurden als „Primärquellen" oder „prim" bezeichnet, und viele Datensätze wurden aus der POP-Dioxin-Datenbank übernommen. Diese Datensätze wurden einer weiteren Klassifizierung unterzogen, um die Bereiche der Technosphäre wie technische Produkte, Emissionen oder Rückstände zu berücksichtigen. Ein Beispiel für die Einteilung in große Quellen, Gruppen und die Stufen 1-3 ist in der folgenden Tabelle Z1 dargestellt.

Tabelle Z1: Beispiele für die gestufte Einteilung von Technosphäre-Proben aus primären Quellen

#	Source	Group	Level_1	Level_2	Level_3	CaseContext	Sample
205	prim	Synthesis	Cl other	Aliphatics	Dye	Generic	Phthalocyanine copper
135	prim	Synthesis	Cl other	Elemental	P&P	Generic	non-wood pulp, PM4, after C-stage
144	prim	Synthesis	Cl other	Elemental	Residue Cl ind	Generic	Residue, Cl industry
45	prim	Synthesis	Cl-precursor	Cl-benzene	ClBz	Generic	Chlorobenzene, China
6	prim	Synthesis	Cl-precursor	Cl-benzene	PCB	Generic	Clophen A30
16	prim	Synthesis	Cl-precursor	Cl-benzene	PCB	Generic	Kanechlor KC-400
33	prim	Synthesis	Cl-precursor	Cl-benzene	PCB	Generic	Aroclor 1260
35	prim	Synthesis	Cl-precursor	Cl-phenol	PCP	Generic	PCP
54	prim	Synthesis	Cl-precursor	Cl-phenol	Phenoxy	Generic	2,4-D/2,4,5-T
246	prim	Thermal	Controlled	Small	Residue	Generic	Chimney ash
225	prim	Thermal	OBurn	Accident	Residue	DUS airport fire	Soot, DUS Terminal
240	prim	Thermal	OBurn	Fuel dep	Residue	Open burn	Ash, open burning
68	prim	Nat'l formation	Geogenic	Clay		Ball clay	Ball clay USA
71	prim	Nat'l formation	Geogenic	Clay		Ball clay	Mabele

Der gleiche Ansatz wurde für die Datensätze von Proben verwendet, bei denen der PCDD-, PCDF- oder PCB-Gehalt übertragen wurde; diese Quellen wurden als „sekundäre Quellen" oder „sec" bezeichnet, und viele Datensätze für biotische Umgebungen wurden von den NRLs oder EURL oder der Literatur beigesteuert, wo bekannte „Dioxinunfälle" beschrieben wurden. Weitere Klassifizierungen umfassten abiotische, biotische und unbekannte Matrizen. Es sei darauf hingewiesen, dass (fast alle) Lebensmittelproben biotisch waren, Futtermittelproben jedoch aus abiotischen Umgebungen stammen können; z.B. Ton, Mineralien. In den Sekundärquellen wurden auch Stichproben gefunden, denen keine Gruppe oder alle Stufen zugeordnet werden konnten. Ein Beispiel ist in der folgenden Tabelle Z2 dargestellt.

Tabelle Z2: Beispiele für die gestufte Einteilung von Technosphäre-Proben aus sekundären Quellen

#	Source	Group	Level_1	Level_2	Level_3	CaseContext	Sample
270	sec	Abiotic	Commodity	Synthetic	Construction	UBA PCB report	Sealant
138	sec	Abiotic	Commodity	Synthetic	Dye	UBA PCB report	Paint ceiling tile
214	sec	Abiotic	Sink	Soil	Impacted	Generic	Soil
311	sec	Biotic	Animal	Food	Aquatic	Ball clay	Catfish
94	sec	Biotic	Animal	Food	Terrestrial	Ball clay	Turkey meat
338	sec	Biotic	Animal	Food	Terrestrial	Citrus pulp	Butter
319	sec	Biotic	Animal	Food	Terrestrial	Belgian PCB case	Chicken meat
73a	sec	Biotic	Animal	Humans	Tissue	Ball clay	Human milk
84	sec	UKN	UKN	Feed		Chile pork	ZnO mix
343	sec	UKN	UKN	Feed		Belgian chicken	Feed mix

In Bezug auf die Verbindungen (Kongenere) wurden die folgenden numerischen Informationen bezüglich der Massenkonzentration in die Datensätze aufgenommen (falls verfügbar). Insgesamt gibt es 66 Spalten mit numerischen Informationen:

- **21 Spalten für PCDD/PCDF**
  17 2,3,7,8-substituierte PCDD/PCDF Kongenere + 3 TEQs (WHO_{2005}-TEQ_{PCDD}, WHO_{2005}-TEQ_{PCDF}, WHO_{2005}-$TEQ_{PCDD/PCDF}$) + ΣPCDD/PCDF(17)
- **16 Spalten für dl-PCB**
  12 dl-PCB Kongenere + 3 TEQs (WHO_{2005}-$TEQ_{no\text{-}PCB}$, WHO_{2005}-$TEQ_{mo\text{-}PCB}$, WHO_{2005}-TEQ_{PCB},) + ΣPCB(12)
- **7 Spalten für ndl-PCB**
  6 Kongenere für Indikator PCB + ΣPCB(6)
- **11 Spalten für PCDD/PCDF Homologe**
  Homologe für PCDD/PCDF (von *tetra* bis *octa*, von Cl_4 bis Cl_8) + Σ(Cl_4-Cl_8)
- **11 Spalten für PCB Homologe**
  Homologe für PCB (von *mono* bis *deca*, von Cl_1 bis Cl_{10}) + Σ(Cl_1-Cl_{10}).

Sobald die Datensätze harmonisiert sind und die Qualitätskontrolle bestanden haben, können sie mit univariaten oder multivariaten statistischen Methoden bewertet werden. Abhängig von der Verfügbarkeit quantitativer Informationen können aus den Daten verschiedene Kategorien von ‚Profilen' wie folgt generiert werden.

Profil 1: Beitrag von Kongeneren zur Summe innerhalb einer Untergruppe von POPs

Profil 1a: Siebzehn PCDD/PCDF
Die Konzentrationen der 17 PCDD/PCDF-Kongenere werden auf die Summe dieser 17 Kongenere normiert und der Beitrag jedes Kongeneren zur Summe wird verarbeitet.

Profil 1b: Zwölf dioxin-ähnliche PCB
Die Konzentrationen der 12 dl-PCB-Kongenere werden auf die Summe dieser 12 Kongenere normiert und der Beitrag jedes Kongeners zur Summe wird verarbeitet.

Profil 1c: 29 dioxin-ähnliche POPs mit einem zugewiesenen TEF
Die Konzentrationen aller 29 im TEF-Schema enthaltenen Kongenere werden auf die Summe dieser 29 Kongenere normalisiert und der Beitrag jedes Kongeners zur Summe wird verarbeitet.

Profil 1d: Sechs Indikator-PCB
Die Konzentrationen der sechs Indikator-PCB-Kongenere werden auf die Summe dieser sechs Kongenere normiert und der Beitrag jedes Kongeners zur Summe wird verarbeitet.

Profil 1e: Homologe der PCDD/PCDF
Die Konzentrationen der fünf Homologen von Cl_4 bis Cl_8 von PCDD und PCDF werden auf die Summe der Homologen normalisiert und der Beitrag jedes Homologen zur Summe wird verarbeitet.

Profil 1f: Homologe der PCB
Die Konzentrationen der zehn Homologen werden auf die Summe der Homologen der PCB normalisiert und der Beitrag jedes Homologen zur Summe wird verarbeitet.

Profil 2: Beitrag von PCDD/PCDF zu seinem Homolog

Die Konzentration jedes Homologen wird durch die Konzentration des jeweiligen 2,3,7,8-substituierten Kongeners geteilt. Es gibt insgesamt 15 Beziehungen (Quotient für OCDD und OCDF ist gleich 1, da das Kongener auch das Homolog ist und daher für dieses Profil nicht verwendet werden kann).

Profil 3: Profile unter Verwendung des TEF Konzepts

Jedes Kongener wird multipliziert mit seinem Toxizitätsäquivalenzfaktor unter Verwendung des Schemas der Weltgesundheitsorganisation von 2005 (WHO_{2005}-TEFs) und der Beitrag jedes TEF-gewichteten Kongeners zum WHO_{2005}-TEQ wird verarbeitet.

Alle der obengenannten Profile können mithilfe handelsüblicher Programme oder anderen Tools visualisiert werden. Wir haben in diese Studie eine Kombination von Heatmaps und Dendrogrammen aus der hierarchischen Clusteranalyse aufgenommen, um die Ähnlichkeit der Profile und der Proben zu ermitteln. Die hier vorgestellten Graphiken wurden mit Hilfe des Programmes R erstellt und die R-Skripte sind im Anhang beigefügt.

Schlussfolgerungen

Ein Datensatz von 278 Proben wurde ausgewählt und auf quantitative Ergebnisse für PCDD/PCDF und PCB angewendet. Es wurde festgestellt, dass eine strikte Struktur festgelegt und aufrechterhalten werden muss, um die Daten für weitere Anwendungen einschließlich der Erkennung von Mustern und Profilen von Umwelt- und Lebensmittel-/Futtermittelproben zu verwenden.

In dieser Studie wird eine Struktur zur Darstellung der Daten vorgeschlagen, die Folgendes enthält:

- Drei große Quellgruppen, jede in Bezug auf primäre oder sekundäre Quellen, ergänzt durch drei Stufen der weiteren Beschreibung der beschreibenden Stichprobe.
- Aufnahme der Referenzen für jeden Datensatz;
- Organisation numerischer Daten nach 17 Kongeneren von PCDD/PCDF, 12 Kongeneren von dioxinähnlichen PCB und sechs Kongeneren von Indikator-PCB (ndl-PCB);
- Homologe von PCDD/PCDF und PCB;
- Weitere quantitative Angaben umfassen die berechneten Werte für die toxischen Äquivalente oder die Summen der Gruppen der Chemikalien ($\Sigma PCDD/PCDF_{17}$, ΣPCB_{12}, ΣPCB_6 or $\Sigma PCDD/PCDF_{Homologe}$ and $\Sigma PCB_{Homologe}$)

Die Strukturen sind flexibel und können durch neue Datensätze (Muster) ergänzt werden, sie ermöglichen auch eine erweiterte Klassifizierung auf allen drei Ebenen sowie die Einbindung weiterer Chemikalien(gruppen).

Es wurden monovariate und multivariate Statistik- und Visualisierungswerkzeuge entwickelt und unter Verwendung der quantitativen Informationen in der Datenbank angewendet. Die Datensätze wurden untereinander getestet, aber auch mit „unbekannten" Stichproben, um die Profile für die weitere Beurteilung anzupassen. Obwohl sehr gute Korrelationen gefunden wurden und die Nützlichkeit des Profilabgleichs gezeigt wurde, ist zur Interpretation der Ergebnisse Fachwissen erforderlich, und es wird davor gewarnt, diese Korrelationen als einzigen Beweis für die Hypothese zu verwenden.

Die Verwendung dieser Datensätze erfolgt für zukünftige Bewertungen in zwei Richtungen durch die beiden Behörden:

1. durch das UBA:

   d) Anwendung der vorgeschlagene Struktur in die Proben in der Dioxin POP-Datenbank

   e) Optimierung der Definitionen und Erweiterung oder Änderung der Parameter für die Charakterisierung der Proben (falls erforderlich)

   f) Übertragung der Struktur, Definitionen und Visualisierung auf Proben anderer Herkunft, z. B. Umgebungsluft, Boden, Sediment;

   g) Überprüfung der Nützlichkeit in weiteren Studien zum Transfer zwischen Umweltkompartimenten und zwischen Umwelt und Menschen.

2. durch EURL/NRL

   a) Optimierung der Datensätze und Beschreibungen für Zwischenfälle/Unfälle mit Lebensmitteln und Futtermitteln;

   b) Erkennung der Korrelationen zwischen ähnlichen Stichproben anhand der Profile

   c) Zuweisung von Quellen, die die Futter- oder Lebensmittelproben verunreinigt haben.

1 Introduction

Since 1991, the Germany Federal Environment Agency maintains a database of initially focused on results for polychlorinated dibenzo-*p*-dioxins and polychlorinated dibenzofurans (PCDD/PCDF) from monitoring and surveillance programs coordinated by the Federal States (Länder), the Federal Institute for Risk Assessment (BfR) and the Federal Office of Consumer Protection and Food Safety (BVL). The original 'dioxin database' has been expanded to include more persistent organic pollutants (POPs) in relevant matrices such as environmental samples including (stack) emissions and residues of sources of POPs, but also soils, sediments, ambient air and biota such as food and feed. The database is continuously updated through contributions from the official surveillance programs of the German agencies but also through research projects such as 'Identification of potentially POP-containing Wastes and Recyclates – Derivation of Limit Values' (Potrykus et al., 2015) or 'Analyse und Trendabschätzung der Belastung der Umwelt und von Lebensmitteln mit ausgewählten POPs und Erweiterung des Datenbestandes der POP-Dioxin-Datenbank des Bundes und der Länder mit dem Ziel pfadbezogener Ursachenforschung ' (Weber et al., 2015).

At Dioxin 2017 in Vancouver, Canada, Malisch and coworkers (Malisch et al., 2017) presented the initial results from a collaboration between the Core Working group 'Dioxin Pattern' of the network of the European Union Reference Laboratory (EURL) with its National Reference Laboratories (NRLs) and the POP-Dioxin database maintained at the German Federal Environment Agency (UBA) (Knetsch and Gärtner, 2011). The EURL and the NRLs are tasked with the surveillance and control of the EU limit values for feed and food and thus, have accumulated an abundance of measured data from official food and feed controls. Of special interest are feed or food samples that exceed the action levels for PCDD/PCDF (on TEQ basis), dioxin-like polychlorinated biphenyls (dl-PCB; on TEQ basis) or polychlorinated biphenyls (PCB; sum of six indicator PCB) and often are related to feed or food accidents. Such accidents have happened in the past and resulted in contamination 'cases' or 'crisis' inside and outside of the European Union (Büchert et al., 2001; Malisch, 2017). Such accidents can have very different origin and some of the prominent examples include

1. The citrus pulp case from Brazil where lime from industrial use has been used in compound feed pellets for ruminants (Malisch et al., 1999);
2. The Belgian dioxin crisis was caused by a feed additive heavily contaminated with improperly discharged PCB (Covaci et al., 2002; van Larebeke et al., 2001),
3. Guar gum from India contaminated with sodium pentachlorophenate,
4. Several accidents involving ball clay or kaolinite as feed additive, a 'natural' dioxin source (Ferrario et al., 2007; Ferrario et al., 2000; Fiedler et al., 1998), and
5. Numerous incidents caused by improperly operated thermal processes such as open burning, drying or contaminated fuels (Solorzano-Ochoa et al., 2012; Zhang et al., 2011).

In this project by Örebro University, Sweden (#97932, Z6-92 950/10), an attempt is made to join relevant datasets for PCDD/PCDF and polychlorinated biphenyls (PCB) from primary sources, where these chemicals are formed and released, with the secondary matrices such as feed and food or technical materials (paints, sealants), where contamination was discovered. These datasets have been selected to establish typical patterns and profiles for the identification of sources of PCDD/PCDF and/or PCB contamination and thus, serve as a starting point for expert tools to assess sources, transfers for rapid responses in cases of feed or food contaminations or impacts on the environment.

The activities under this project included the following:

1. Classification of samples/datasets
2. Revision of the names of profiles/datasets
3. Classification of analytes (including a structure for sum parameters)
4. Assignment of original units as shown in the references and recalculation to harmonized units
5. Inventory of existing datasets and classification of the datasets
6. Inclusion of references/sources of information for all datasets (original data)
7. Acquisition of new data
8. Development of typical pattern
9. Quality control
10. Corrections within existing datasets and inclusion into the new structure

2 Structure of the database

The database for this joint project is composed of 'technosphere' datasets from the Federal Environment Agency and amended by new datasets from literature found relevant for source identification and datasets from the food and feed controls provided by the EURL and the NRLs. The database is developed and maintained in MsExcel®. Each sample is characterized by its unique sample number and descriptive qualitative information for identification and characterization as well as units and their basis (lipid, dry matter, fresh weight) and numeric results as mass concentrations for the following congeners and sum parameters.

It was found very important to harmonize all information – qualitative and quantitative – and therefore, a rigid structure was developed and applied.

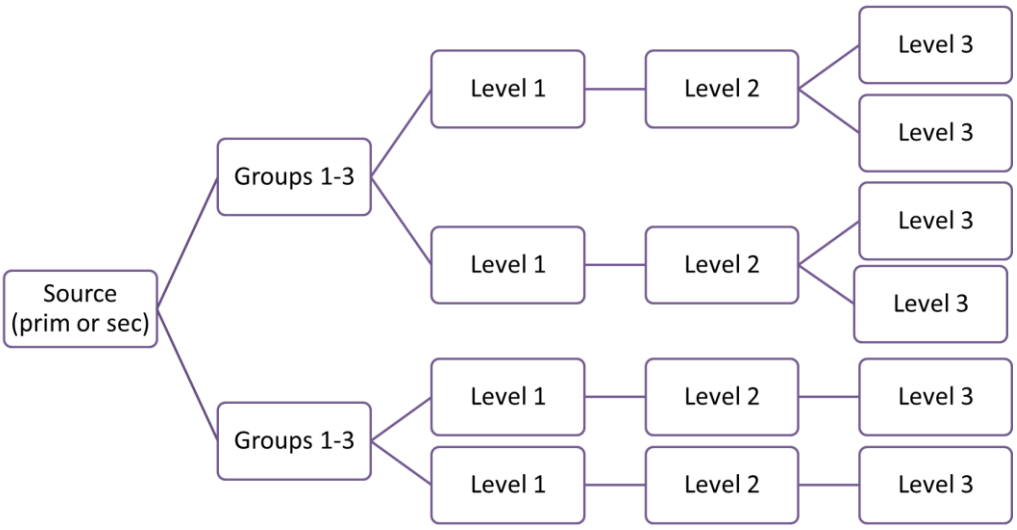
The database is maintained in MsExcel® as one sample per row and contains descriptive qualitative and quantitative information (including units and basis (fat, dry matter (d.m.) as well as literature references

The references - for each dataset - are maintained in EndNote X7® (Clarivate Analytics, Boston, MA, USA). The references used in this study are listed in the Appendix section 9.3 References for the samples included into the UBA/EURL dataset.

2.1 Hierarchical classification of the samples (*per* row)

It was found important to characterize the datasets/samples systematically. A hierarchical but flexible approach was chosen and Figure 1 shows the generic format.

Figure 1: Generic structure to set-up qualitative information for the classification of samples/



Own diagram, Örebro University

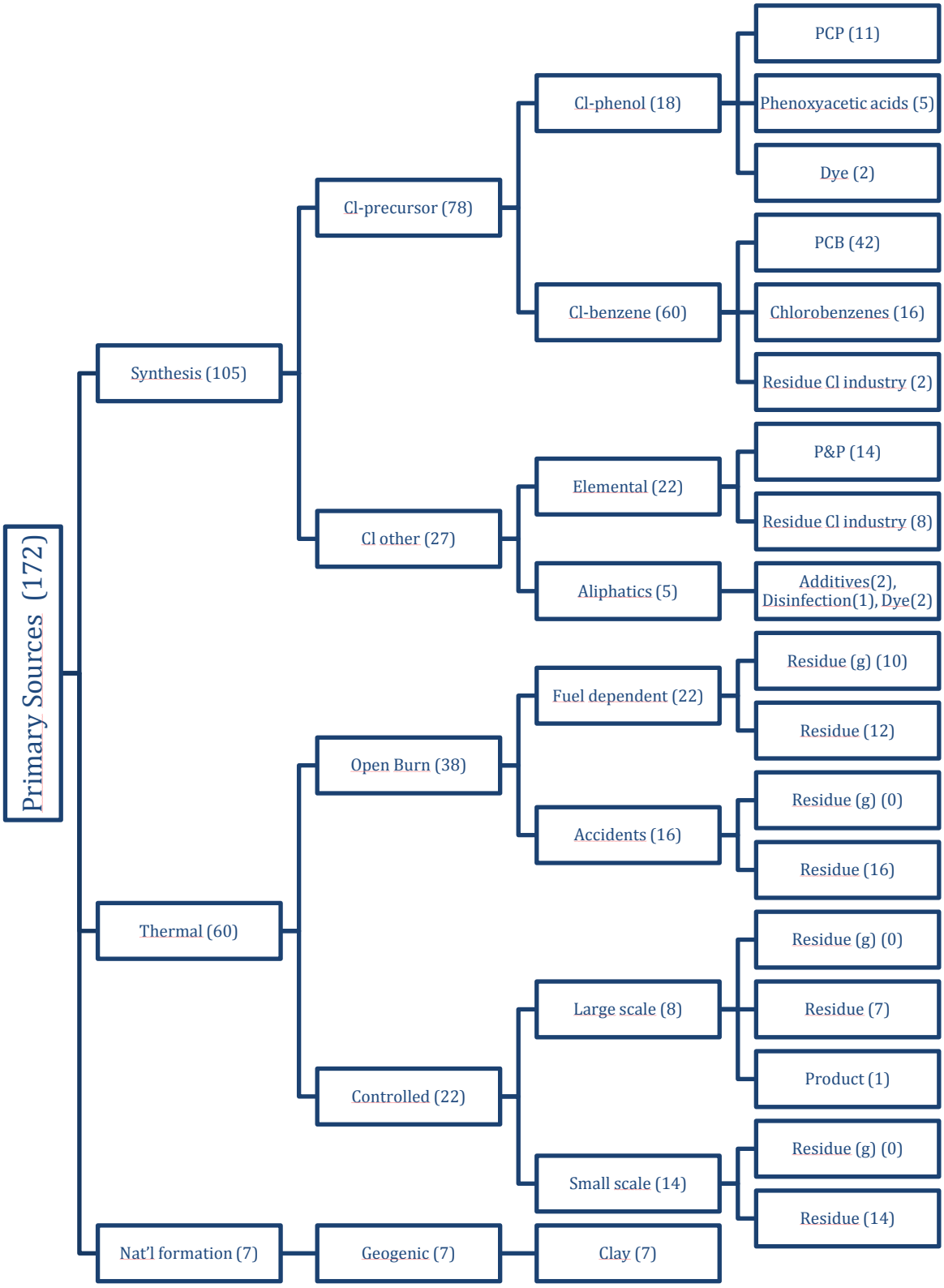
All samples were divided into primary sources ('prim') and secondary sources ('sec'). 'Primary sources' means that PCDD/PCDF but also PCB are formed at the sources and 'secondary sources' means that PCDD/PCDF and PCB have been transferred from a primary source (known or unknown) to another matrix. The structure includes assignment trees for both large groups as are shown in Figure 2 and Figure 3. Each structure has three large groups followed by further differentiation into three levels.

In Figure 2, commonly used abbreviations are used as listed below:

- Cl-phenol chlorophenol
- Cl-benzene chlorobenzene
- Cl-precursor chlorine-containing precursor compound
- PCP pentachlorophenol
- P&P pulp and paper
- Residue (g) gas-phase residues, emissions in the gas phase (*e.g.* stack emissions)
- Residue Residue solid (*e.g.* ashes, sludges)

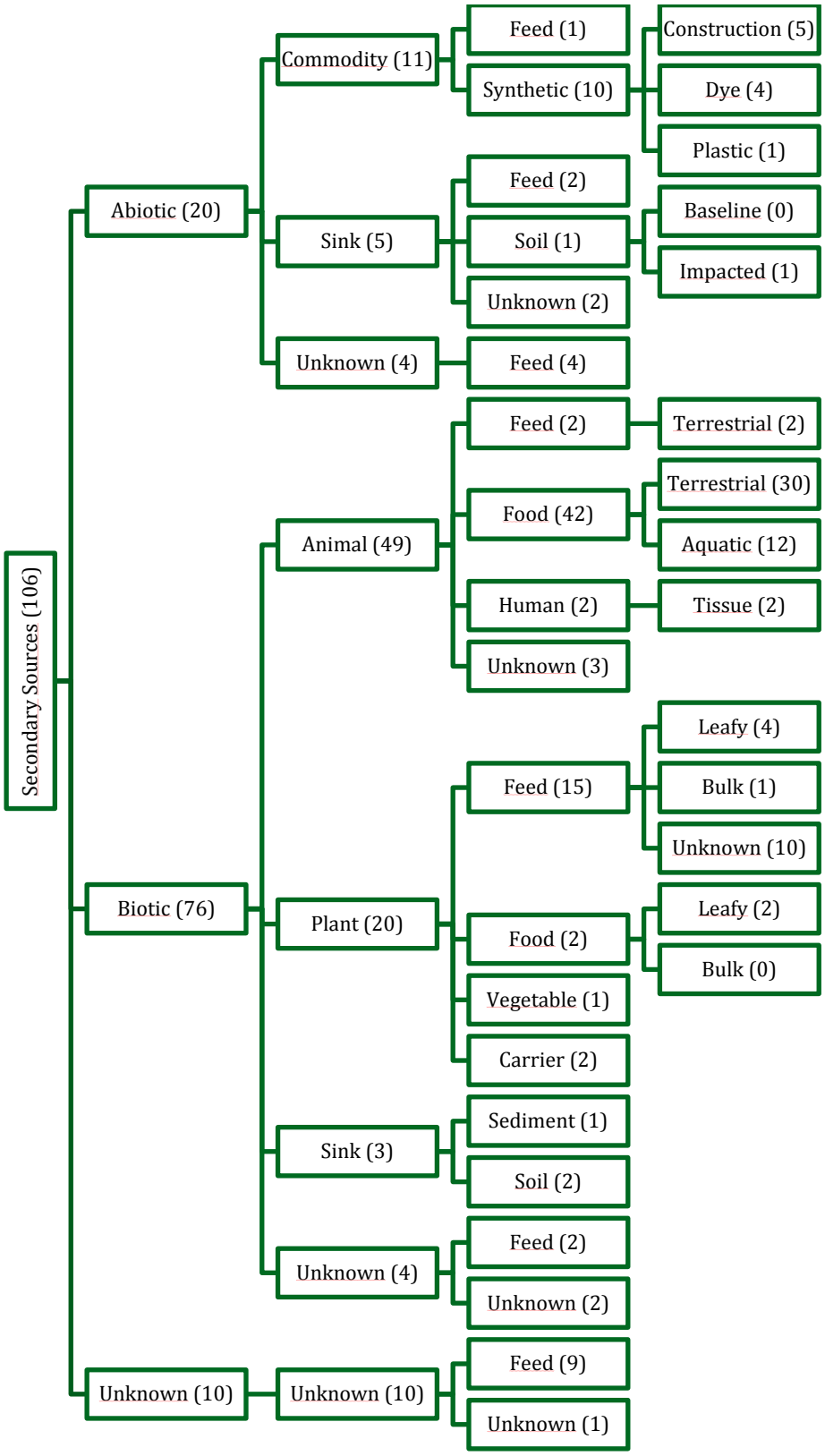
In order to better understand the chemistry under these classifications, the structural formulae of the chemicals that are listed in the database are shown in the Appendix 9.1 - Some Chemical Molecules/Structures referenced in this study.

Figure 2: **Structure for primary sources with the three large groups 'Synthesis', 'Thermal' and 'Natural formation' followed by the three levels. The number in parenthesis indicates the number of samples contained in each classification level**



Own diagram, Örebro University

Figure 3: **Structure for secondary sources with the three large groups 'Abiotic', 'Biotic' and 'UKN (Unknown)' followed by the three levels. . The number in parenthesis indicates the number of samples contained in each classification level**



Own diagram, Örebro University

2.2 Numeric information of the samples (*per* row)

The numeric information includes concentrations for the

- **17 PCDD**/PCDF **congeners**
  amended by WHO_{2005}-TEQ_{PCDD}, WHO_{2005}-TEQ_{PCDF}, WHO_{2005}-$TEQ_{PCDD/PCDF}$, and ΣPCDD/PCDF(17)
- **12 dl-PCB congeners**
  amended by WHO_{2005}-$TEQ_{no\text{-}PCB}$, WHO_{2005}-$TEQ_{mo\text{-}PCB}$, WHO_{2005}-TEQ_{PCB}, and ΣPCB(12)
- **6** indicator **PCB congeners** amended by ΣPCB(6)
- Homologs **for PCDD/PCDF** (*tetra* through *octa*, Cl_4 through Cl_8) and Σ(Cl_4-Cl_8)
- Homologs **for PCB** (*mono* through *deca*, Cl_1 through Cl_{10}) and Σ(Cl_1-Cl_{10}).

The numeric information is compiled in 66 columns

- 17+4 for PCDD/PCDF,
- 12+4 for dl-PCB,
- 6+1 for indicator PCB,
- 10+1 for PCDD/PCDF homologs, and
- 10+1 for PCB homologs

It shall be noted that some columns had to be duplicated such as for example Cl_8DD (octachlorodibenzo-*p*-dioxin) or Cl_8DF (octachlorodibenzofuran), which are congeners in the section containing the 17 2,3,7,8-substituted PCDD/PCDF but are also homologs in the section containing the ten homologs of PCDD/PCDF.

3 Corrections/amendments to the database

3.1 General

A 'Master table' has been generated in MsExcel® from the extract of the UBA Dioxin POP-database. The basis table has been amended by columns to harmonize the information for the:

- Homologs of PCDD/PCDF:

  Units (it shall be noted that the units for the homologs differ from the unit the 17 PCDD/PCDF congeners were reported)
  OCDD und OCDF

- Homologs of PCB:

  Units (it shall be noted that the units for the homologs differ from the unit the six indicator PCB were reported)
  mono- and *di*chlorinated PCB as well as *deca*-PCB

All datasets having the wrong (IUPAC) sequence for the hexachlorinated dibenzofurans have been changed to be in the correct sequence: Cl_6DF3 (1,2,3,7,8,9-HxCDF) followed by Cl_6DF4 (2,3,4,6,7,8-HxCDF).

3.2 Datasets contained in the UBA database

The majority of the datasets from the UBA database were contributions from a recent report (Weber et al., 2015); however, not all datasets could be retrieved or not all were included in the database. This resulted in addition of more datasets but also elimination of existing datasets. Further, a number of datasets were corrected. Subsequently, the datasets contained in the UBA database were denominated as "Profil" (not to be mixed up with "profile", which is used to characterize the composition of the individual datasets/samples according to the analytes of PCDD/PCDF or PCB; see section 5).

3.2.1 UBA #154 – Branntkalk

PCDD/PCDF concentrations were all below LOD (NWG); therefore, they could not be used and were eliminated. For dl-PCB, no units were found; hence, the unit „ng/kg" was added (the same as for the PCDD/PCDD) and all numerical values for concentration were divided by 1000.

3.2.2 UBA #155 – Styropor

Renamed to "polystyrene" for international applications. Almost all PCDD/PCDF data were below LOD (NWG); therefore, they could not be used and were eliminated. For dl-PCB, no units were found; hence, the unit „ng/kg" was added (the same as for the PCDD/PCDD) and all numerical values for concentration were divided by 1000.

3.2.3 UBA #156 and UBA #157

The two originals (LUFA-Prüfbericht LV4-RU-0010-2012 (UBA # 156) and LUFA-Prüfbericht LV4-RU-0011-2012 (UBA # 157) could not be retrieved (as part of (Weber et al., 2015). It was considered that the numerical values for dl-PCB concentrations were too high by a factor of 1000; hence, the same procedure was followed as for the previous samples 3.2.1 and 3.2.2.

3.2.4 From UBA #171 to UBA #176

The concentrations listed could not be identified in the referenced publication (Liu et al., 2004). The datasets were substantially manipulated and presented in the report in such a way that concentrations were calculated for not analyzed congeners. In addition, there were too many gaps for the PCB homologs and the unit was wrong (should be ng/g). These datasets were eliminated due to manipulation and gaps.

3.2.5 UBA #380

For the Chinese graphite sludge (Xu et al., 2000), the unit was too low by a factor of 1000 (Note: The heading of the table in the original publication reads "ng/kg x 10^3".)

3.2.6 UBA #813, UBA #814, UBA #819 and UBA #815

The Reference was corrected from "Wang (2014)" (Wang et al., 2014) to "Wang (2012)" (Wang et al., 2012). The sequence for the HxDF was corrected. The report for the Italian sample was not found (UBA #813).

For UBA #813 and UBA #819 (chlorine bleach of non-wood fibers), the data for effluent and sludge data were added as well as for the paper sample PM4. The reference was corrected to "Wang (2012)" (Wang et al., 2012).

3.2.7 UBA #811

Additional samples were added from Santl H, Gruber L, Meisburger K, Wolz G Identifizierung und Quantifizierung dioxin-relevanter Einträge in das zur Papierherstellung eingesetzte Altpapier. Abschlussbericht. Dezember 1995 in Weber UBA 114/2015 (Weber et al., 2015). This entry still needs a check if the sequence for HxDF is correct. The original report was not available.

3.2.8 From UBA #669 to UBA #675

Several results from pesticides reported by Holt et al. (Holt et al., 2010) were too high by a factor of 1000; hence, corrections were done. The listed concentration for the sample 2,4-D was wrong, since it was not considered that the compound was present in aqueous solution. Concentration values for analytes in UBA #673 (2,4,5-T/2,4-D) were given in ng/mL and were wrong by a factor of 1000. Values for PCDD/PCDF homologs were corrected as well as values for OCDD concentrations in UBA #669 and UBA #670.

3.2.9 From UBA #664 to UBA #668

Numerical values for were provided as 0.5 × LOD for values below LOD, as taken from a publication by Huang et al. (Huang et al., 2015). It is considered that setting values to 0.5 × is more appropriate; hence, listed values were changed accordingly. Some of the samples were identified as duplicates. Assignment of samples needs a final check as to raw and final product (note: final products have higher contaminations than the raw products).

3.2.10 From UBA #753 to UBA #756

These are data from the report by Volland and Neuwirth (Volland and Neuwirth, 2005). The names of the samples may be misleading: samples come from ceiling tiles (Schallschutzdeckenplatten, 0-2 mm) found at the University of Tübingen. It is assumed that they were covered with PCB-contaminated paints. The reported concentrations were corrected from the units originally reported (these were, e.g., µg/kg for PCDD/PCDF or mg/kg for all PCB).

UBA #753 is presented as a sample from building B in Tübingen; actually, two different locations B1 and B2 were merged into one in the UBA sample (see report page 21 vs. page 23). Note: The sealants from these buildings were not available; they have thus been added as new datasets (Raum E; page 222), since it was not possible to relate the PCB and the PCDD/PCDF data.

3.2.11 From UBA #121 to UBA #125

Congeners and homologs were available in the referenced publication and have been added to the existing datasets (Thoma, 1988)

From Theisen et al. (Theisen et al., 1989): the results from real accidents were added and homolog data added to the laboratory experiments.

3.3 Datasets changed or commented

Many of the publications analyzing PCB on congener-specific basis do not provide mass concentrations, but percentages (%). Instead. The database contains several datasets where information is provided as percentages.

Publications by Frame often do not contain quantitative information on technical mixtures of PCB (on congener-specific basis), but rather retention indices. Quantitative data (%) were found in the Chemosphere publication (Frame et al., 1996).

Takasuga (Takasuga et al., 2006): All numbers are in % and were retained as such.

The thesis by Peter Behnisch does not contain numbers for PCB 180 (Behnisch, 1997).

Kim et al. (Kim et al., 2011): The numbers for the samples from the Chile pork accident were provided as partial TEQs. These numbers were recalculated to the respective mass concentration (by dividing by the TEF; here WHO1998-TEFs). Some of the data in the database were removed, since they contained "expert-judged" numbers by the author of the UBA Report (Weber et al., 2015).

Ni et al. (Ni et al., 2005): The Publication year was corrected from 2004 to 2005.

PCP and PCP-Na results were duplicates from the sources (a) doctoral thesis of Hermann Brunner and (b) publication in Chemosphere by Hagenmaier und Brunner (Hagenmaier and Brunner, 1987). Slight differences were identified. The data from the publication were retained; the ones from the doctoral thesis were eliminated.

EURL #93 was eliminated by A. Schächtele.

The reference Grochowalski et al. (Grochowalski et al., 2015) for samples EURL #119 and EURL #120 was completed and the data corrected for the agreed sequence within the HxCDF..

3.4 Additional datasets

Additional datasets have been entered, which were found to be relevant as they characterize potential sources for transfer of PCDD/PCDF or PCB contaminations into biotic or abiotic environments, into products or into the foodchain. These datasets covered the sources PCB, pulp and paper, pesticides, and open burning processes.

Furthermore, datasets were added from reports on food/feed accidents, belonging to the group of secondary sources.

3.4.1 PCB

Aroclor data for dl-PCB found in the publication by Rushneck et al. (Rushneck et al., 2004) were added (the datasheet was presented from R. Weber).

PCB-3 and PCB-5 from China were added; the originally used publication by (Huang et al., 2011); however data are from (Jiang et al., 1997). The EURL data did not have numbers after the decimal; these have been added (#10 and #26).

3.4.2 Pesticides

Masunaga et al. (2001) (Masunaga et al., 2001): Some more pesticides were added; units corrected. Dl-PCB were not included since the authors report coelution of PCB with PCDD/PCDF (see Table 3 of the publication).

A more recent PCP sample from the USA (2010) by Tondeur et al. was added (Tondeur et al., 2010).

3.4.3 Pulp and Paper

New publications from China (RCEES) on non-wood fibers and chlorine bleach were found but unfortunately did not have numeric results and could not be used (Fang et al., 2017; Yang et al., 2017; Zheng et al., 2001). Similarly, papers by Beck et al. had too many gaps in the congeners (Beck et al., 1989; Beck et al., 1988).

From Publications by Santl and coworkers, concentrations from dyestyffs were added for PCDD/PCDF congeners (Santl et al., 1994). At this stage, the homolog data have not been included but can be calculated by adding the results from the 2,3,7,8-substituted congeners to the homologs that were characterized as „non-2,3,7,8" homologs.

3.4.4 Open Burning

A number of air and ash/soot samples from open burning experiments (Solorzano-Ochoa et al., 2012; Zhang et al., 2011) or accidents (Lindert and Fiedler, 1999) have been added since they were found more relevant as a local source of contamination than the large combustion or thermal plants (the latter ones are typically well controlled).

3.4.5 Food/Feed Accidents

New samples #321, #322, #323, #324 – from Belgian chicken accident: Results by Covaci et al. (Covaci et al., 2002) were amended by dl-PCB. The data were reported as partial TEQ using the WHO1998-TEFs. In the dataset now, the mass concentrations have been reclcuated using the WHO2005-TEF scheme; as a consequence the concentrations for TEQs are much lower now.

It is not clear if the results reported by Covaci et al. (Covaci et al., 2002) and van Larebeke (sample 3-15-320) (van Larebeke et al., 2001).

Additional samples from Germany were added (CVUA, 1999a, b)

A sample name "toxic fat" from RIKILT was eliminated since it was taken from a publication by Hayward et al. (Hayward et al., 1999). Further, mistakes with the units were detected (pg/g and not ng/g); homologs from the original reference were missing in the Rikilt dataset.

Data from the citrus pellet accident were added (CVUA, 1998).

4 Composition and content of the database

The database now contains 278 samples; of these, 51 were from the UBA POP Dioxin database; 227 were added. It shall be noted that none of the datasets was complete with all of the congeners or homologs being available as displayed in the database structure. As can be seen from Table 1, there are 238 datasets that contain PCDD/PCDF congeners; of these, 95 datasets have all congeners quantified (no result <LOD). There are only 19 datasets from PCB or chlorobenzenes that have information on the ten PCB homologs and of these, only seven have quantitative results for all ten homologs.

Table 1: Coverage of data according to groups of PCDD/PCDF and PCB in the dataset

	Σ PCDD/PCDF$_{(17)}$	Σ dl-PCB$_{(12)}$	Σ PCB$_{(6)}$	ΣPCDD/PCDF$_{(10)}$	ΣPCB$_{(10)}$
n>0	238	75	63	98	19
n=max	95	21	51	73	7

The classification of the datasets has already been depicted in Figure 2 for the source group of the primary sources and in Figure 3 for the secondary sources. The following tables (Table 2 and Table 3) show the composition and content of the large source group and the split into primary and secondary sources. It can be seen that the overall 278 datasets include 172 datasets describing primary sources and 106 datasets that describe secondary sources. Within the primary sources, there are 105 samples assigned 'Synthesis', 60 samples assigned 'Thermal' and 7 samples assigned 'Nat'l formation'. Within the secondary sources, there 20 samples assigned 'Abiotic', 76 assigned 'Biotic' and 10 'Unknown' (UKN).

From these tables it can be seen also that the classification 'UKN' occurs also at the lower levels; *e.g.*, level 2 or Level 3. It is used when no specific classification can be made at present.

In terms of sources of information (References), 83 references have contributed to these datasets; 45 for the primary sources and 44 for the secondary sources. This demonstrates that there is only small overlap in publications providing full congener information on both, primary and secondary sources.

Table 2: Primary and secondary sources and corresponding classifications up to Level 2

Source group	**Primary sources**					**Secondary sources**								**Grand Total**
Grand Total	**172**					**106**								**278**
Group	Synthesis		Thermal		Nat'l for-mation	Abiotic			Biotic				Un-known	
Subtotal	**105**		**60**		**7**	**20**			**76**				**10**	**278**
Level 1	Cl-pre-cursor	Cl other	Open burning	Controlled	Geo-genic	Comm odity	Sink	UKN	Animal	Plant	Sink	UKN	UKN	
Subtotal	**78**	**27**	**38**	**22**	**7**	**11**	**5**	**4**	**49**	**20**	**3**	**4**	**10**	**278**
Level 1	Cl-phenol	Elemental	Accidents	Large scale	Clay	Feed	Feed	Feed	Feed	Feed	Soil	Feed	Feed	
Level 2	Cl-benzene	Aliphatics	Fuel-dependent	Small scale		Synthe tics	Soil		Food	Food	Sedim ent	UKN	UKN	
							UKN		Humans	Vegeta ble				
									UKN	Carrier				

Table 3: Matrix of datasets at Level 1 *vs.* Level 2

Level 1 vs. Level 2	Cl-phenol	Cl-benzene	Elemental	Aliphatics	Fuel-	Accidents	Large scale	Small scale	Clay	Feed	Synthetics	Feed	Soil	UKN	Feed	Feed	Food	Humans	UKN	Feed	Food	Vegetable	Carrier	Soil	Sediment	Feed	UKN	Feed	UKN
Prim. – Synthesis	105																												
Cl precursor	18	60																											
Cl other			22	5																									
Prim. – Thermal					60																								
Open burning					22	16																							
Controlled							8	14																					
Prim. – Nat'l form.									7																				
Geogenic									7																				
Sec. – Abiotic										20																			
Commodity										1	10																		
Sink												2	1	2															
UKN															4														
Sec. – Biotic																76													
Animal																2	42	2	3										
Plant																				15	2	1	2						
Sink																								1	2				
UKN																										2	2		
Sec. – UKN																												10	
UKN																												9	1

5 Profiles

An important factor to consider are the data generators. It has been shown that despite highest qualifications including accredited laboratories, there are substantial differences between their mandates or research objectives. Among these are the following: Historically, laboratories when analyzing abiotic matrices for the technosphere such as chemicals, ashes, industrial chemicals, contaminated soils) typically report 2,3,7,8-substituted PCDD/PCDF congeners but also homologs and other non-2,3,7,8-substituted PCDD/PCDF congeners, whereas laboratories analyzing biotic matrices such as foods of animal origin or humans 'only' analyze the 17 2,3,7,8-substituted PCDD/PCDF. Today – due to enforcement – the analytical spectrum often is limited to 17 PCDD/PCDF congeners that have been assigned a toxicity equivalency factor (TEF) by a WHO/UNEP expert group (van den Berg et al., 2006) and therefore, important information as to the share of the 2,3,7,8-substituted congeners within a PCDD or PCDF homolog is lost. This lack of information for source identification seems to be less relevant for food of animal origin (meat, dairy, eggs) but can be severe for feedstuffs or mollusks.

Another factor to consider is the very wide range of concentrations reported from mg/kg to fg/kg or fg/m^3.

Data have been entered using the original unit. Since the scale of the concentrations may play a role in certain applications, the worksheet 'ppt' was created. Therein, all concentrations were unified to 'kg *per* sample' for abiotic matrices and 'g *per* sample' for biotic matrices available for

- PCDD/PCDF (congeners and homologs including the sum parameters or TEQs) or dl-PCB (including sum parameters or TEQs)
  - All abiotic samples were assigned the concentration ng/kg
  - All biotic samples were assigned the concentration pg/g
- The concentrations of the six indicator PCB (or ndl-PCB) were converted into units three orders of magnitude higher, which corresponds to 'ppb', *i.e.,*
  - All abiotic samples were assigned the concentration unit µg/kg
  - All biotic samples were assigned the concentration unit ng/g

In the following the term 'profile' is used to characterize the samples according to contributions of congeners to a sum parameter (summation to the total number of relevant congeners). The following profiles are proposed:

5.1 Profile 1: Contribution of congeners to the sum within a sub-group of POPs

5.1.1 Profile 1a: Seventeen PCDD/PCDF

Concentrations of the 17 PCDD/PCDF congeners are normalized to the sum of these 17 congeners and the contribution of each congener is processed. The approach follows the following equation:

$$\frac{\text{2,3,7,8-Cl}_4\text{DD}}{\text{PCDD+PCDF}} + \ldots + \frac{\text{Cl}_8\text{DD}}{\text{PCDD+PCDF}} + \frac{\text{2,3,7,8-Cl}_4\text{DF}}{\text{PCDD+PCDF}} + \ldots + \frac{\text{Cl}_8\text{DF}}{\text{PCDD+PCDF}}$$

5.1.2 Profile 1b: Twelve dioxin-like PCB

Concentrations of the 12 dl-PCB congeners are normalized to the sum of these 12 congeners and the contribution of each congener is processed.

5.1.3 Profile 1c: 29 dioxin-like POPs having with an assigned TEF

Although not used in scientific publications, there is an option to assess the datasets that have numeric values for the 29 dl-POPs; *i.e.* 7 PCDD, 10 PCDF and 12 PCB. However, it shall benoted that the sum will be dominated by the *mono-ortho* PCB. In order to obtain reasonable information, a division factor may be applied to the eight *mono-ortho* PCB.

5.1.4 Profile 1d: Six indicator PCB

Concentrations of the six indicator PCB congeners are normalized to the sum of these six congeners and the contribution of each congener is processed.

5.1.5 Profile 1e: Homologues of PCDD/PCDF

Concentrations of the five homologs from Cl_4 to Cl_8 of PCDD and PCDF for are normalized to the sum of the homologs and the contribution of each homolog congener is processed.

5.1.6 Profile 1f: Homologues of PCB

With the present dataset, an assessment does not make sense since there are too few compounds. However, in a larger database, the homologs can be normalized to the contribution of each homolog to the sum of the homologs.

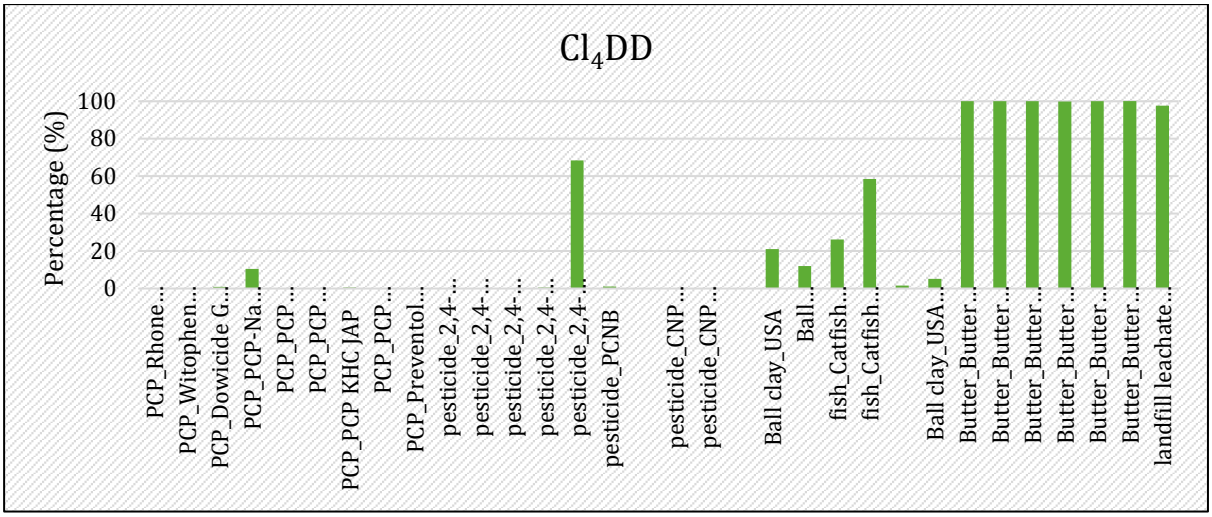
5.2 Profile 2: Contribution of PCDD/PCDF congener to its homolog

For samples where 2,3,7,8-substituted congeners and homologs are available, the contribution of the 2,3,7,8-substituted congener within the homolog can be used. The following two figures show the profile for Cl_4DD, namely the contribution of 2,3,7,8-Cl_4DD to the homolog Cl_4DD. It shall be noted that the theoretical value is ca. 4% since there are 22 isomers within the homolog. It can be seen that in the combustion profile, the contributions from 2,3,7,8-Cl_4DD is closer to the theoretical value indicating a de novo formation than in the synthesis processes. It can be seen, that in the butter sample (corresponding to human or mammal type), all non-2,3,7,8-substituted congeners are metabolized and only the 2,3,7,8-substituted transferred.

The concentration of each 2,3,7,8-substituted congener is divided by the concentration of the respective homolog. There are a total of 15 relations (quotient for OCDD and OCDF equals 1 since the congener is also the homolog).

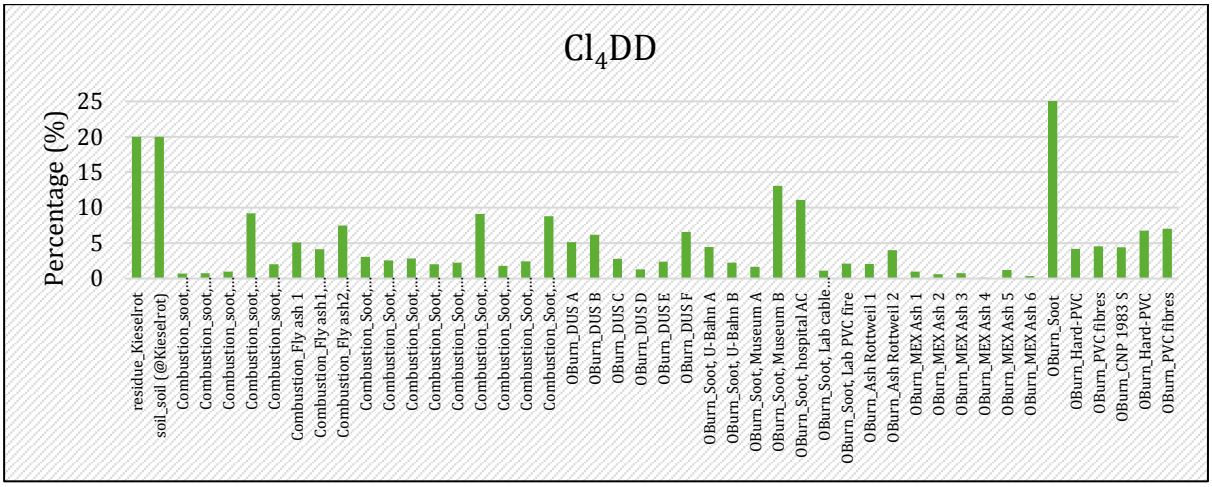
$$\frac{\text{2,3,7,8-Cl}_4\text{DD}}{\text{Cl}_4\text{DD}}, \frac{\text{1,2,3,7,8-Cl}_5\text{DD}}{\text{Cl}_5\text{DD}}, \ldots, \frac{\text{1,2,3,4,6,7,8-Cl}_7\text{DF}}{\text{Cl}_7\text{DF}}, \frac{\text{1,2,3,4,7,8,9-Cl}_7\text{DD}}{\text{Cl}_7\text{DD}}$$

Figure 4: Contribution of 2,3,7,8-Cl_4DD to the homolog Cl_4DD for the set of 'Synthesis' samples



Own diagram, Örebro University

Figure 5: Contribution of 2,3,7,8-Cl_4DD to the homolog Cl_4DD for the set of 'Thermal' samples



Own diagram, Örebro University

5.3 Profile 3: Profiles using the TEF concept

Each congener multiplied by its toxicity equivalency factor using WHO_{2005}-TEFs (van den Berg et al., 2006) and the contribution to the WHO_{2005}-TEQ is processed

$$\frac{2,3,7,8\text{-}Cl_4DD \bullet TEF_{2378\text{-}Cl4DD}}{WHO_{2005}\text{-}TEQ} + ... + \frac{Cl_8DD \bullet TEF_{Cl8DD}}{WHO_{2005}\text{-}TEQ} + \frac{2,3,7,8\text{-}Cl_4DFD \bullet TEF_{2378\text{-}Cl4DF}}{WHO_{2005}\text{-}TEQ} + ... + \frac{Cl_8DF \bullet TEF_{Cl8DF}}{WHO_{2005}\text{-}TEQ}$$

The above can be applied depending on availability of data to

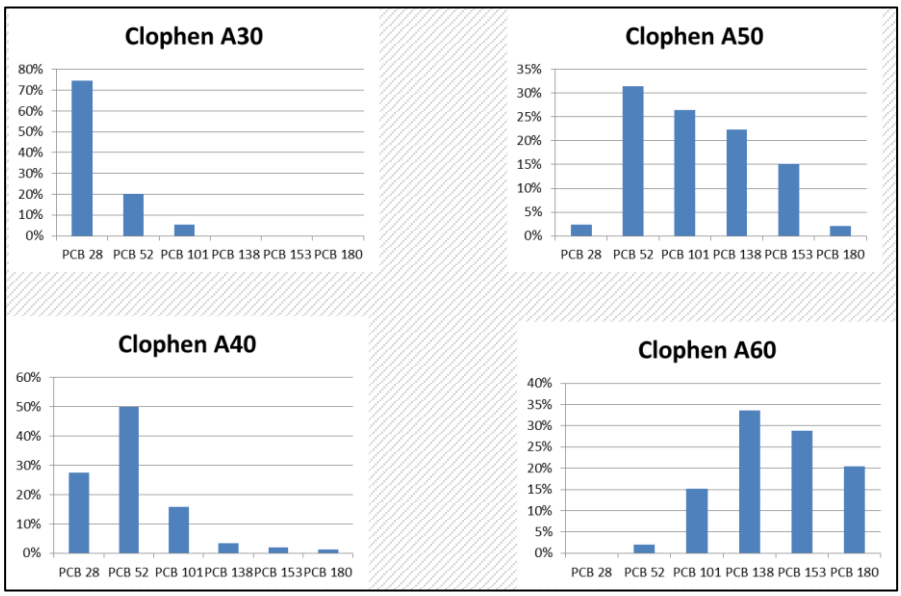
- 17 congeners of PCDD and PCDF whereupon these can be divided into 7 congeners of PCDD and 10 congeners of PCDF ⇒ **Profile 3a**
- 12 congeners of dl-PCB whereupon these can be subdivided into 4 congeners of non-*ortho* PCB and 8 *mono-ortho* substituted PCB ⇒ **Profile 3ba**, and
- All 29 congeners that have a TEF assigned ⇒ **Profile 3c**.

6 Visualization

There are several 'simple' ways such as bar graphs to visualize the profiles/data.

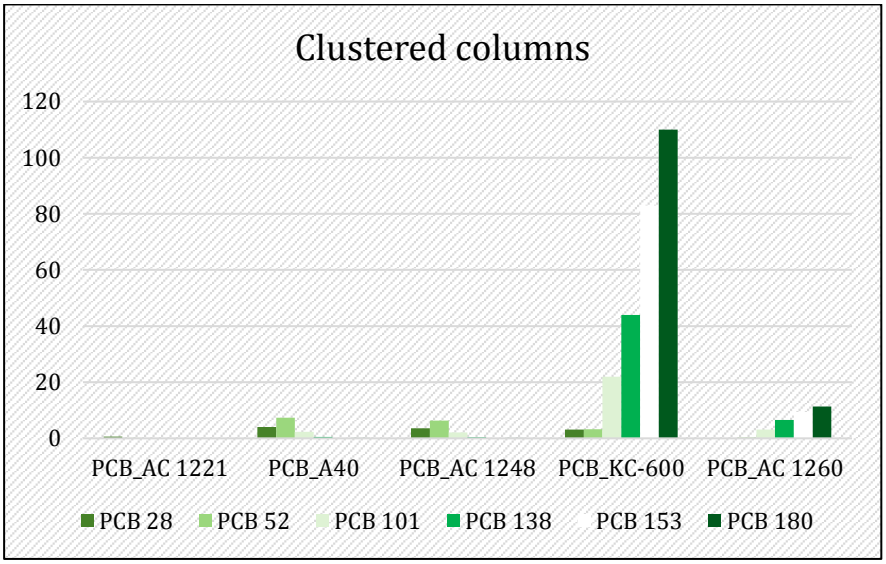
Simple bar graphs can be applied to demonstrate the distribution for a sample; *e.g.,* for the six indicator PCB either separately for each sample (Figure 6) or combined into one graph (Figure 7). Better comparison is obtained when normalizing to 100% and showing percentage distributions of the six ndl-congeners (Figure 8).

Figure 6: Bar diagram of four technical samples of PCB showing the concentrations of the ndl-PCB



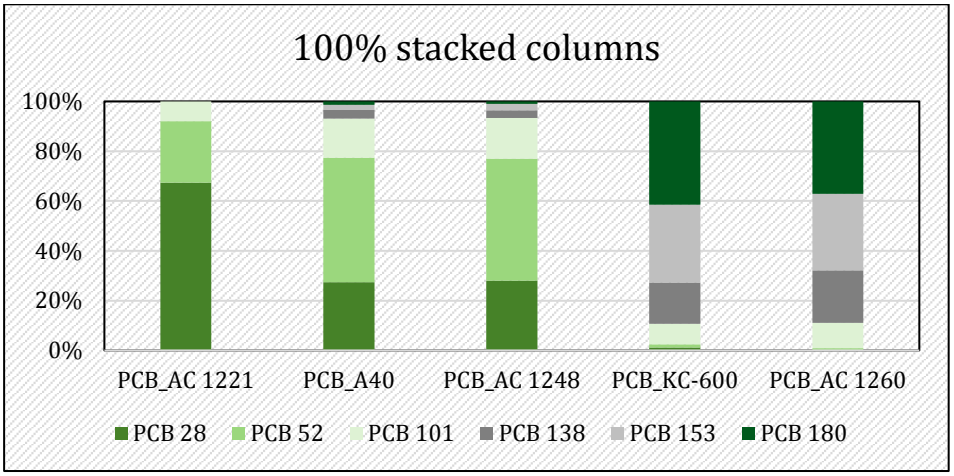
Own diagram, Örebro University

Figure 7: Bar diagram for six ndl-PCB (clustered columns) (concentration for KC-600 divided by 1000)



Own diagram, Örebro University

Figure 8: Bar diagram for six indicator PCB (100% stacked columns)



Own diagram, Örebro University

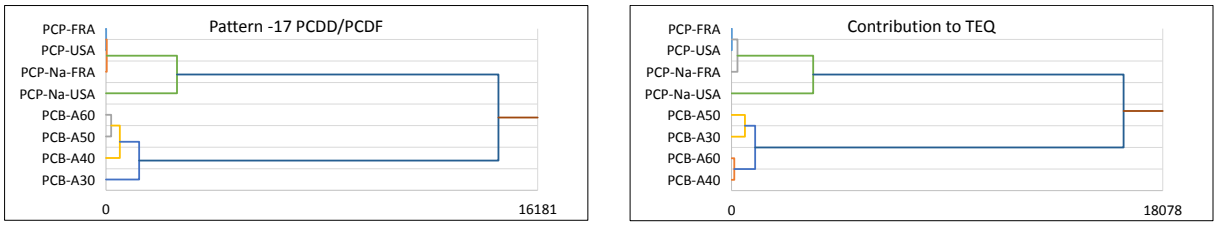
Heat maps can be used to obtain a rapid color-coded picture for the distribution of congeners using the 'conditional formatting' option in MsExcel®. In Figure 9 the three colors were set to signify contributions to the sum of the six indicator PCB as less than 5% (green), between 5% and 30% (yellow) and greater than 30% (red). It can be seen that the lower-chlorinated PCB mixtures (AC for Aroclor, A for Clophen A and KC for Kanechlor products) have quite similar compositions with PCB-28 dominating the sum of the six indicator PCB with shares of more than 60%. In contrast, the higher-chlorinated products such as Aroclor 1262, Aroclor 1260 and Kanechlor 600 have more than half of the sum of the six congeners from PCB-153 and PCB-180. Interestingly, within the chemicals having at around 40% of chlorine, two commercial products from US and German production – Clophen A40 and Aroclor 1248 have a distinct but very similar compositions (PCB A40 and PCB AC 1248) and are distinct from the other Aroclor (AC 1242) and the Kanechlor product (KC 400).

Figure 9: Heat map for six ndl-PCB (red >30%, yellow 5%-30%, green <5%)

	PCB 28	PCB 52	PCB 101	PCB 138	PCB 153	PCB 180
PCB_AC 1221	67	25	8	-	-	-
PCB_AC 1232	63	29	5	1.0	0.8	0.2
PCB_KC-200	78	16	3.2	1.4	1.0	0.4
PCB_AC 1016	65	35	0.3	-	-	-
PCB_A30	75	20	5	-	-	-
PCB_KC-300	78	16	3.2	1.6	1.1	0.5
PCB_AC 1242	61	31	6	0.9	0.5	-
PCB_A40	27	50	16	3.4	2.1	1.4
PCB_AC 1248	28	49	16	3.1	2.6	0.9
PCB_KC-400	43	37	14	3.3	2.4	1.1
PCB_AC 1254	0.7	4.5	38	33	20	3.1
PCB_AC 1254	1.9	31	30	22	11	4.2
PCB_A50	2.4	31	26	22	15	2.1
PCB_KC-1000	2.3	16	32	26	20	4.3
PCB_KC-500	2.5	15	31	27	20	4.1
PCB_A60	-	2.0	15	34	29	20
PCB_AC 1262	0.4	0.6	4.5	11	28	56
PCB_KC-600	1.2	1.2	8	17	31	41
PCB_AC 1260	0.1	0.8	10	21	31	37

Multivariate methods can be applied to further assess similarities and differences between datasets. Hierarchical cluster analysis (HCA) can be applied: HCA attempts to find a natural grouping (clustering) of a data set to achieve less variation (greater similarity) within a group (cluster) and more variation (less similarity) between the groups (clusters). The data in a set are grouped into clusters of great(er) similarity to form a dendrogram. Figure 10 shows a dendrogram from HCA (Software: StatistiXL in MsExcel® using Squared Euclidian distance and Ward method) for four PCB and four PCP samples.

Figure 10: Hierarchical cluster analysis, dendrograms for contribution of the 17 PCDD/PCDF to the sum of 17 PCDD/PCDF and the contribution of the TEF-weighted 17 PCDD/PCDF congeners to the TEQ



Own diagram, Örebro University

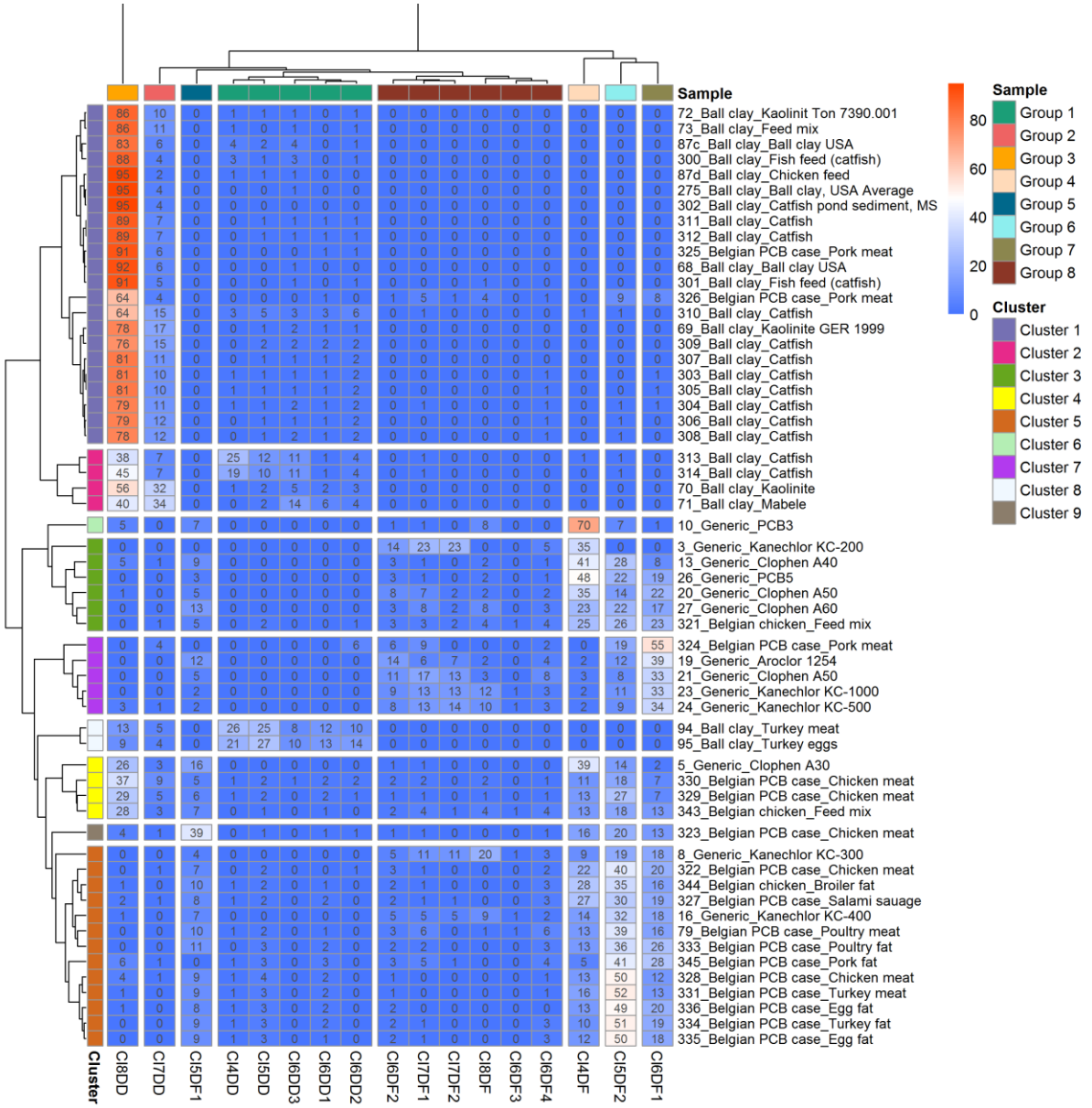
The following figures show the combination of a heatmap with dendrograms from hierarchical cluster analysis. When examining the various outputs, it was found that such visualization presents the most comprehensive picture as to the similarity and correlations between samples. The following Figure 11 through Figure 13 show the graphics for the contribution of congeners to the sum of the respective congeners as detailed in section 5.1.

Figure 11 shows the heat map dendrogram obtained from the contribution of the 17 PCDD/PCDF congeners to Σ17 PCDD/PCDF as detailed for Profile 1a in section 5.1.1. The numbers in the cells provide the contribution of the specific congener (shown on bottom of the graph). The congeners are organized according to decreasing contribution from left to right. Eight groups and nine clusters have been generated. Cluster 1 is dominated by OCDD with a contribution of 64% to 95% to the Σ17 PCDD/PCDF. Therein are found food and feed samples related to the Ball clay/Kaolinit contamination but also some of the original ball clay samples (from natural formation, Samples 72, 69). The same type of samples are clustered together in Cluster 2 having a lower but still significant contribution from OCDD. In the lower part of the graph are the food samples reported in the context of the Belgian chicken crisis in 1999. The clusters are in the neighbourhood of the PCB technical mixtures 'Generic PCB…'.

Figure 12 shows the heatmap dendrogram obtained for a set of open burn and PCB samples from the contribution of the 12 dl-PCB to Σ12 dl-PCB. In general, the profile is dominated by PCB-118. Five distinct groups and five clusters are formed. However, care needs to be taken upon interpretation since the Kanechlor and the human milk samples cluster close together. Further, it can be seen that from the 12 dl-PCB alone, no differentiation between a 'combustion source' (samples containing 'open burning') and technical PCB mixtures (sample containing 'PCB') can be made.

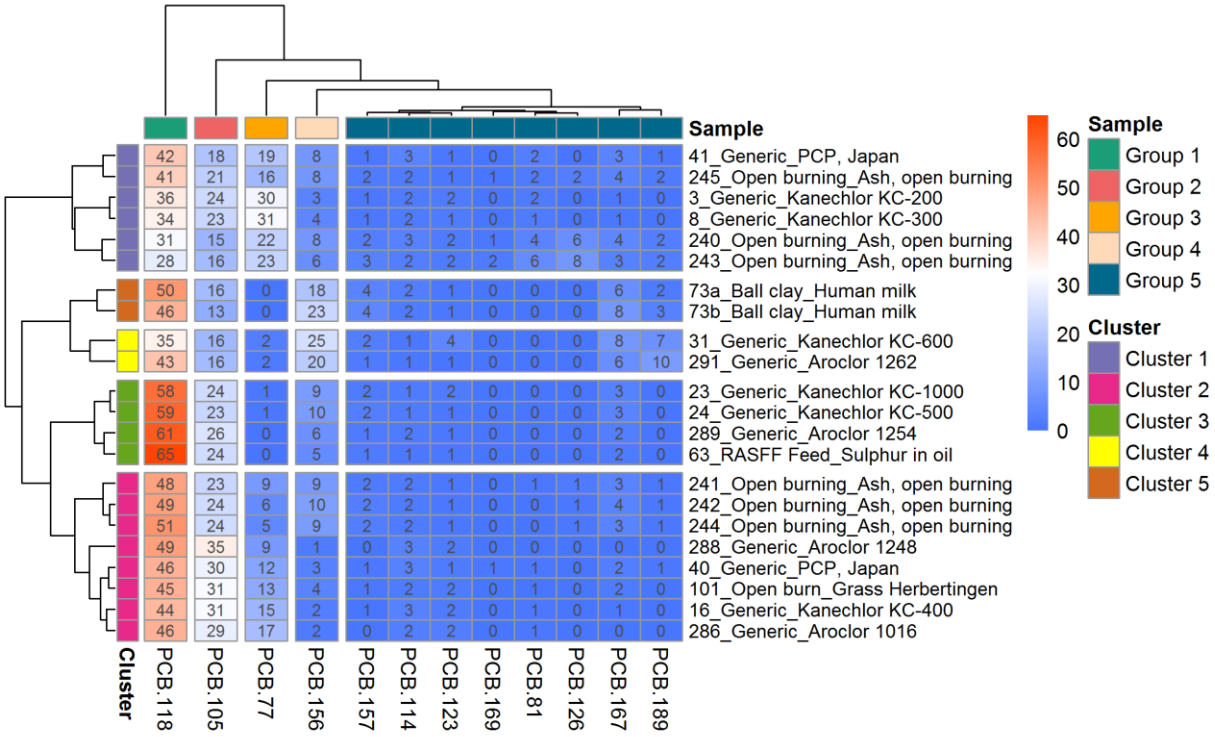
Finally, Figure 13 shows the heatmap dendrogram obtained for a set of PCB technical mixtures, ball clay and Belgian chicken accident samples from the contribution of the six indicator PCB to the Σ6 ndl-PCB. There are five clusters and only two groups. The grouping separates the low chlorinated PCB (PCB-28, PCB-52, PCB-101) from the higher chlorinated PCB (PCB-138, PCB-153, PCB-180). The group of the higher chlorinated PCB (Group 2, red color) contains the higher contributions from higher chlorinated PCB such as in Clusters 3 and 4and the PCB technical mixtures (Aroclor 1254, KC-500, KC-1000, *etc.)* as well as the Belgian chicken case food samples and paints at ceiling tiles. On the other hand, the PCB-containing sealants are found in Clusters 1 and 2, with a higher abundance of the lower chlorinated PCB (Group 1, green color).

Figure 11: Heatmap with dendrogram for subset of PCDD/PCDF for contribution of the 17 PCDD/PCDF to their sum



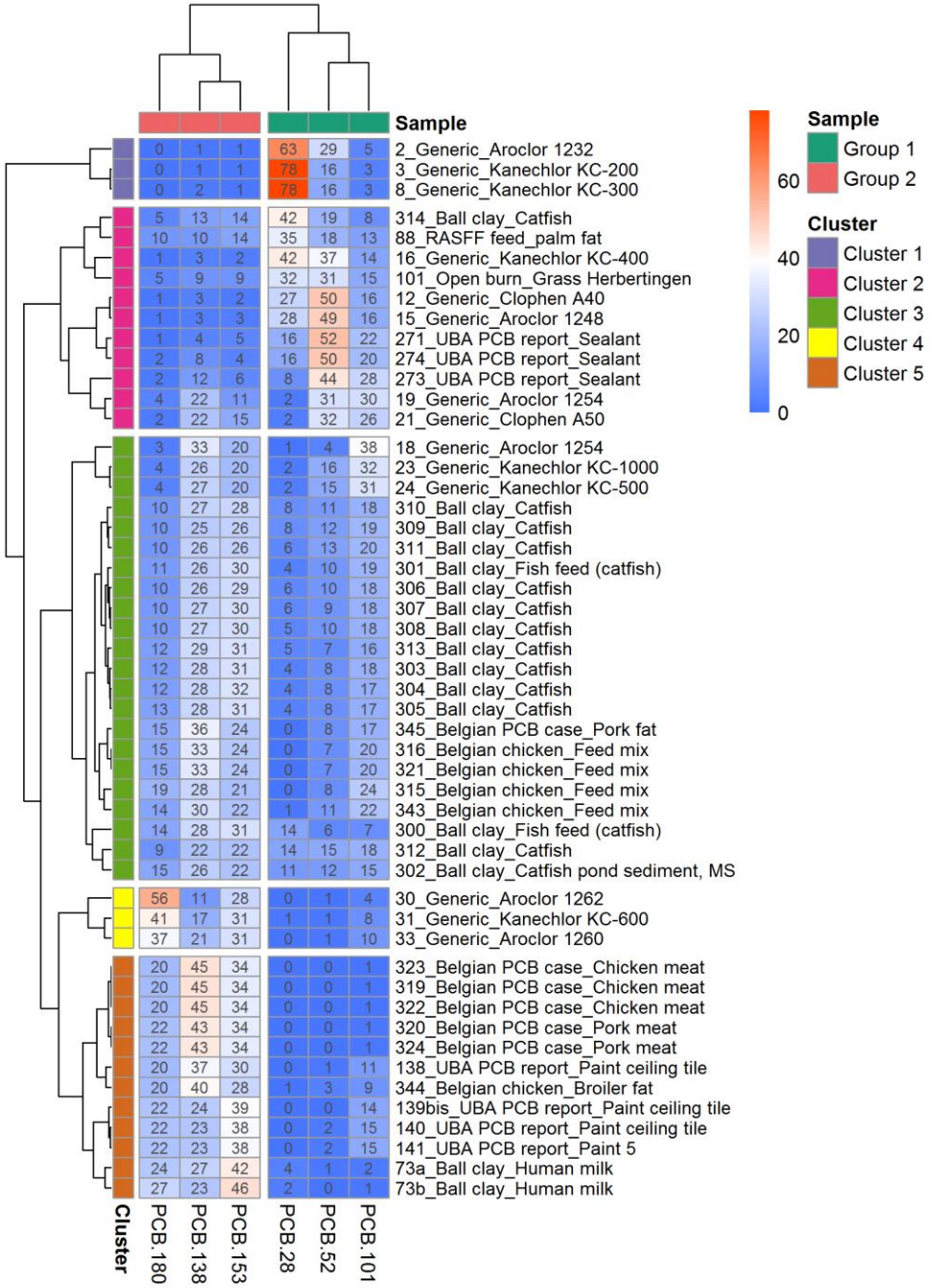
Own diagram, Örebro University

Figure 12: Heatmap with dendrogram for subset of dl-PCB for contribution of the 12 dl-PCB to their sum



Own diagram, Örebro University

Figure 13: Heatmap with dendrogram for subset of indicator PCB (ndl-PCB) for contribution of the six indicator PCB to their sum



Own diagram, Örebro University

The R-scripts are annexed to this report and can be found in the Appendix, section 9.2

7 Conclusions

A dataset of 278 samples has been selected and analyzed to produce quantitative results on congener patterns and profiles of PCDD/PCDF and PCB. It was found that a strict structure needs to be established and maintained in order to use the data in further applications including recognition of patterns and profiles of environmental and food/feed samples.

In this study, a structure to keep and organize the data is proposed containing:

- Four Levels of sample characterization as to primary or secondary sources;
- Inclusion of the references;
- Organization of numeric data according to 17 congeners of PCDD/PCDF, 12 congeners of dioxin-like PCB and six congeners of indicator PCB (ndl-PCB);
- Non-dioxin-like congeners of PCDD/PCDF were not included; although it is recognized that important information such as the contribution of a 2,3,7,8-substituted congener to the total of a homolog is lost;
- Homologs of PCDD/PCDF and PCB are included since the percentage of 'toxic' congeners within the respective homolog group can provide valuable information;
- Further columns include the calculated values for the toxic equivalents or the sums of groups of the chemicals (ΣPCDD/PCDF$_{17}$, ΣPCB$_{12}$, ΣPCB$_{6}$ or ΣPCDD/PCDF$_{Homologs}$ and ΣPCB$_{Homologs}$

Simple monovariate statistics and visualization tools normalizing the concentrations of congeners or homologs to the respective sums can be applied to generate profile and multivariate statistics such as hierarchical cluster analysis or principal component analysis can be applied to identify similarities between samples. The datasets have been tested between each other but also with 'unknown' samples to 'match' the profiles for further assessment. Although very good correlations were found and the usefulness of the profile matching was shown, it will need expertise to interpret the results and a warning be expressed to use these correlations as the only proof of the hypothesis.

The use of this dataset will be for future assessments in two directions by the two authorities:

1. The German UBA :
   a) Introduce the proposed structure to the samples in the Dioxin POP database
   b) Optimize the definitions and expand or change the characterization if necessary;
   c) Apply to the samples of types other than sources and food/feed; *e.g.,* ambient air, soil, sediment;
   d) Explore the usefulness for further studies on transfers between environmental compartments and between environment and humans.
2. EURL/NRLs
   a) Optimization of the datasets and descriptions for food and feed incidents/accidents;
   b) Identify correlations between similar samples according to the profiles;
   c) Assign sources that have contaminated the feed or food samples.

8 List of references

Büchert, A., Cederberg, T., Dyke, P., Fiedler, H., Fürst, P., Hanberg, A., Hosseinpour, J., Hutzinger, O., Kuenen, J.G., Malisch, R., Needham, L.L., Olie, K., Päpke, O., Aranda Rivera, J., Thanner, G., Umlauf, G., Vartiainen, T., van Holst, C. (2001). ESF Workshop. Environmental Science and Pollution Research 8, 84-88.

Covaci, A., Jake Ryan, J., and Schepens, P. (2002). Patterns of PCBs and PCDD/PCDFs in chicken and pork fat following a Belgian food contamination incident. Chemosphere 47, 207-217.

Ferrario, J., Byrne, C., and Schaum, J. (2007). Concentrations of polychlorinated dibenzo-p-dioxins in processed ball clay from the United States. Chemosphere 67, 1816-1821.

Ferrario, J.B., Byrne, C.J., and Cleverly, D.H. (2000). 2,3,7,8-Dibenzo-p-dioxins in Mined Clay Products from the United States: Evidence for Possible Natural Origin. Environmental Science & Technology 34, 4524-4532.

Fiedler, H., Cooper, K., Bergek, S., Hjelt, M., Rappe, C., Bonner, M., Howell, F., Willett, K., and Safe, S. (1998). PCDD, PCDF, and PCB in farm-raised catfish from southeast United States - Concentrations, sources, and CYP1A induction. Chemosphere 37, 1645-1656.

Malisch, R. (2017). Incidents with Dioxins and PCBs in Food and Feed-Investigative Work, Risk Management and Economic Consequences. Journal of Environmental Protection 8, 744-785.

Malisch, R., Berger, B., and Verstraete, F. (1999). Lime as Source for PCDD/F-Contamination of Brazilian Citrus Pulp Pellets (CPPs). Organohalogen Compounds 41, 51-53.

Malisch, R., Bernsmann, T., Ceci, R., Diletti, R., Eppe, G., Fernandes, A., Fiedler, H., Hart, J., Hoogenboom, R., Hove , H., Knetsch, G., Leondiadis, L., A. Lüth, A., Marchand, P., Maszewski, S., Rose, M., Schaechtele, A., Vanderperren, H., van Leeuwen, S., van Raamsdonk, L., Vassiliadou, I. (2017). Interactive data base of PCDD/F and PCB congener patterns to aid identification of contamination sources in feed and food. Organohalogen Compd 79, 545-547.

Solorzano-Ochoa, G., de la Rosa, D.A., Maiz-Larralde, P., Gullett, B.K., Tabor, D.G., Touati, A., Wyrzykowska-Ceradini, B., Fiedler, H., Abel, T., and Carroll, W.F. (2012). Open burning of household waste: Effect of experimental condition on combustion quality and emission of PCDD, PCDF and PCB. Chemosphere 87, 1003-1008.

van Larebeke, N., Hens, L., Schepens, P., Covaci, A., Baeyens, J., Everaert, K., Bernheim, J.L., Vlietinck, R., and De Poorter, G. (2001). The Belgian PCB and dioxin incident of January-June 1999: exposure data and potential impact on health. Environmental Health Perspectives 109, 265-273.

Van den Berg, M., Birnbaum, L.S., Denison, M., De Vito, M., Farland, W., Feeley, M., Fiedler, H., Hakansson, H., Hanberg, A., and Haws, L. (2006). The 2005 World Health Organization reevaluation of human and mammalian toxic equivalency factors for dioxins and dioxin-like compounds. Toxicological sciences 93, 223-241.

Weber, R., Hollert, H., Kamphues, J., Ballschmiter, K., Blepp, M., and Herold, C. (2015). Analyse und Trendabschätzung der Belastung der Umwelt und von Lebensmitteln mit ausgewählten POPs und Erweiterung des Datenbestandes der POP-Dioxin-Datenbank des Bundes und der Länder mit dem Ziel pfadbezogener Ursachenforschung. In UBA Dokumentation, N. Umweltforschungsplan des Bundesministeriums für Umwelt, Bau und Reaktorsicherheit, ed. (Fachgebiet IV 2.1 Internationales Chemikalienmanagement).

9 Appendix

9.1 Some chemical molecules/structures referenced in this study

Figure 14: **Polychlorinated dibenzo-*p*-dioxins (PCDD) (left) and polychlorinated dibenzofurans (PCDF) (right)**

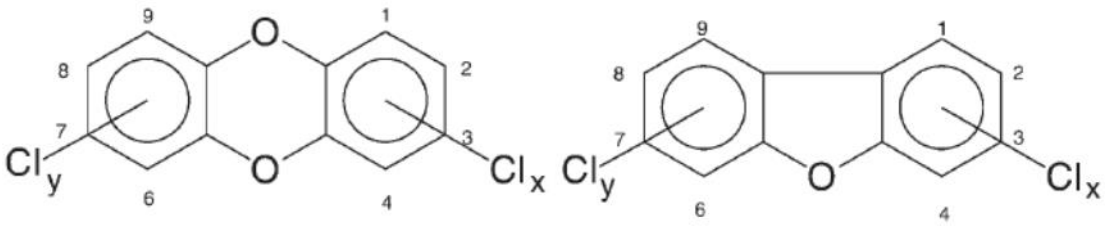
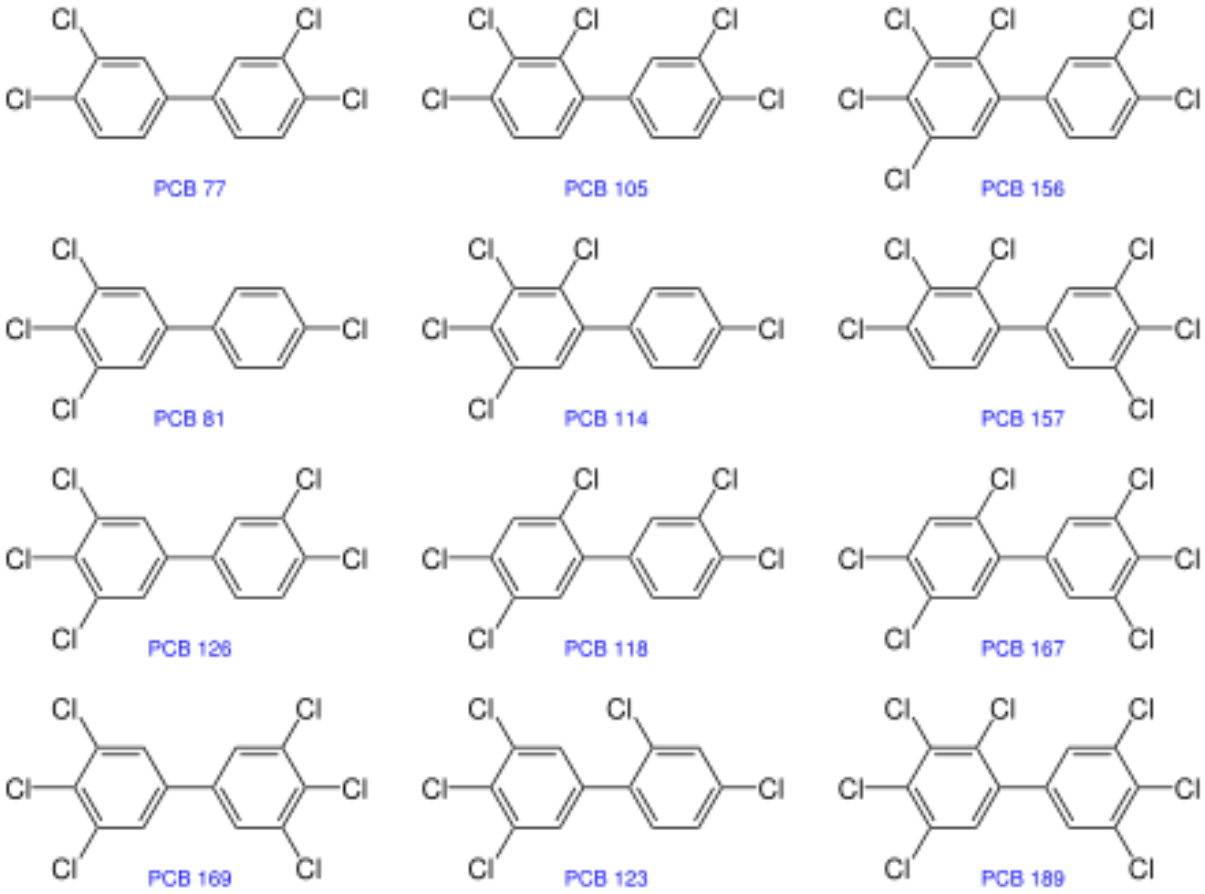
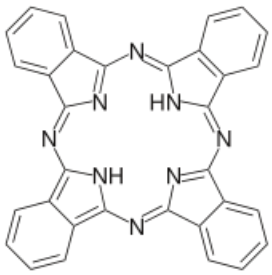


Figure 15: **Dioxin-like PCB (column left: non-*ortho* substituted PCB; 2nd and 3rd column *mono-ortho*-substituted PCB)**

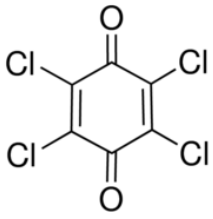


Ref: https://en.wikipedia.org/wiki/Polychlorinated_biphenyl

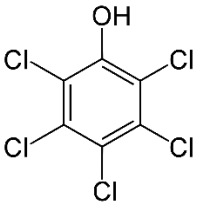
Figure 16: Structures of other relevant compounds



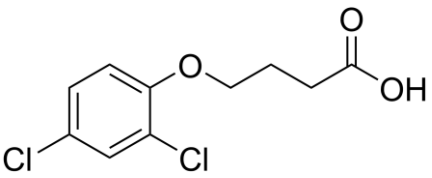
Phthalocyanine



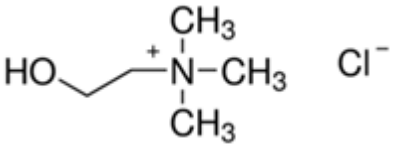
p-Chloranil (tetrachloro-1,4-benzoquinone)



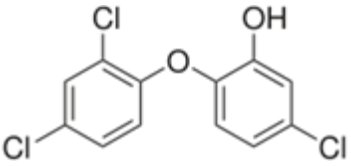
Pentachlorophenol (PCP)



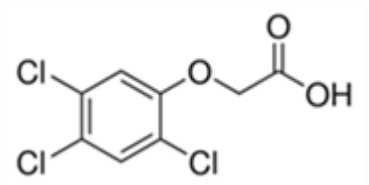
2,4-Dichlorophenoxybutyric acid (2,4-DB)



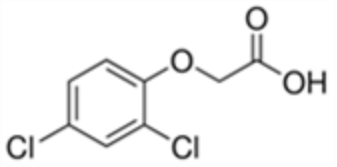
Choline chloride; vitamin B4



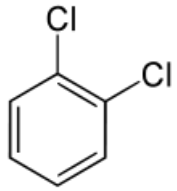
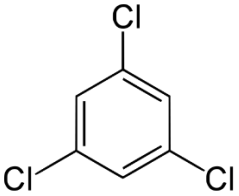
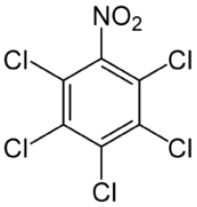
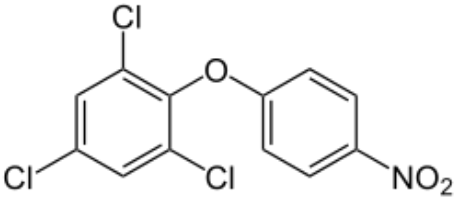
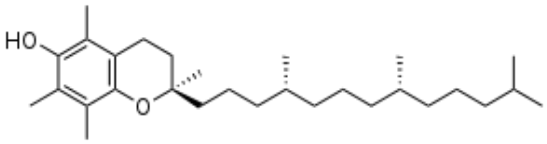
Triclosan (5-chloro-2-(2,4-dichlorophenoxy) phenol)



2,4,5-Trichlorophenoxyacetic acid (2,4,5-T)



2,4-Dichlorophenoxyacetic acid (2,4,-D)

 1,2-Dichlorobenzene (DCBz)	 1,3,5-Trichlorobenzene (TCBz)
 Pentachloronitrobenzene (PCNB)	 Chloronitrophene (CNP) a PCP replacement in Japan; esp. in rice fields
 α-Tocopherol	

9.2 R-scripts for cluster analysis

Cluster analysis serves to identify groups (clusters) and hitherto unknown, hidden structures in higher-dimensional datasets. The groups forming during the calculation should be as similar as possible to each other, but clearly different from other groups.

Another aspect is still very important in cluster analysis. The method is also a tool that allows a contaminated sample to be quickly compared to an existing pool of data by well-established mathematical methods. In particular for the investigation of damage cases (dioxin and PCB scandals), the group of possible causes is thereby narrowed down quickly and efficiently.

In the cluster analysis presented here, therefore, both the variables (congeners) and the samples are analyzed. In each case, a matrix of substances is used that can provide a complete data set for the corresponding definition (PCCD/PCDF, dl-PCB and ndl-PCB).

The values in the cells are the rounded percentages based on the sums of the chemical group in the specific sample. This may result in less than 100 in total in some cases due to rounding, but improves readability by omitting the decimal places.

The color scale is also based on the percentages. The more red the color, the higher the proportion; the deepest blue represents the lowest contribution.

The groups refer to the regression tree of PCCD/PCDF, dl-PCB and ndl-PCB. The clusters refer to the regression tree of the samples.

The modules were developed using the R programming language. The original version comes from the following source [https://github.com/rasbt/R_snippets/tree/master/heatmaps] and has been specially adapted to the needs of the dioxin and PCB calculations.

The modules for the different groups are slightly different. In order to be re-used and further developed, the R-source texts for each group are shown separately.

9.2.1 R-Module 1 - Cluster analysis for dioxins and furans (group PCDD/PCDF)

Result see Figure 11

```
library(readxl)
library(pheatmap)
library(dendextend)

DATA <- read_excel("UBA-EURL datafile_hardcoded DF.xlsx",
                   col_names = TRUE, na = c("NA", "ND"),
                   range = "A3:AH62")
colnames(DATA) <- DATA[1,]
DATA <- DATA[-1,]
rownames(DATA) <- as.vector(as.matrix(DATA[,1]))

#extract Dioxins and Furans
DATA$NAME <- paste(rownames(DATA), DATA$CaseContext, DATA$Sample, sep = "_")
DATAX <- DATA[,c(grep("NAME", names(DATA)),
                 c(grep("Cl4DD", names(DATA)): grep("Cl8DF", names(DATA))))]
#eliminate rows which consist only NA values
DATAXX <- as.data.frame(apply(DATAX[,-1], 2, FUN = function(X) as.numeric(X)))


DATAXXX <- data.frame(DATAX[which(apply(DATAXX, 1, sum, na.rm = TRUE) > 0),])


# transform detection limits
for (j in 2: length(DATAXXX))
{
  i <- grep("<", as.vector(DATAXXX[, j]))
  DATAXXX[i,j] <- as.numeric(sub("<", "",DATAXXX[i,j]))/2
  DATAXXX[,j] <- as.numeric(DATAXXX[,j])
}

names(DATAXXX)[1] <- "Sample"
# percentage values
SUM <- apply(DATAXXX[,-1], 1,
             FUN = function(X) sum(as.numeric(X), na.rm = TRUE))
PER <- data.frame(Sample = DATAXXX$Sample,
                  apply(DATAXXX[,-1], 2,
                        FUN = function(X) as.numeric(X)/SUM) * 100)

#########################################################
### customizing and plotting heatmap
#########################################################
PER_M <- as.matrix(round(PER[,-1], digits = 0))

rownames(PER_M) <- PER[,1]
# replace NA values with zeros
PER_M[is.na(PER_M)] <- 0

#exclude samples without measurement of the Dioxine/Furan
PER_M2 <- PER_M[which(is.finite(apply(PER[,-1], 1, sum))),]


my_hclust_DIOXINE <- hclust(dist(PER_M2), method = "complete")
as.dendrogram(my_hclust_DIOXINE) %>%
  plot(horiz = TRUE)

(my_col_DIOXINE9 <- cutree(tree = as.dendrogram(my_hclust_DIOXINE), k = 9))


my_col_DIOXINE9 <- data.frame(Cluster = paste("Cluster",
                                              my_col_DIOXINE9, sep = " "))

rownames(my_col_DIOXINE9) <- rownames(my_col_DIOXINE)
```

```
#vertical
vertical_DIOXINE <- hclust(dist(data.frame(t(PER_M2))), method = "complete")
as.dendrogram(vertical_DIOXINE) %>%
  plot(horiz = FALSE)

(vertical_col_DIOXINE8 <- cutree(tree = as.dendrogram(vertical_DIOXINE), k = 8))

vertical_col_DIOXINE8 <- data.frame(Sample = paste("Group",
                                                   vertical_col_DIOXINE8,
                                                   sep = " "))

rownames(vertical_col_DIOXINE8) <- rownames(vertical_col_DIOXINE)


# Specify colors for annotation
ann_colors <- list(
  Sample = c('Group 1' = "#1B9E77", 'Group 2' = "indianred2",
             'Group 3' = "orange", 'Group 4' = "peachpuff", 'Group 5' =
"deepskyblue4",
             'Group 6' = "darkslategray2", 'Group 7' = "khaki4",
             "Group 8" = "tomato4"),
  Cluster = c('Cluster 1' = "#7570B3", 'Cluster 2' = "#E7298A",
              'Cluster 3' = "#66A61E", 'Cluster 4' = "yellow",
              'Cluster 5' = "chocolate",
              'Cluster 6' = "darkseagreen2", 'Cluster 7' = "darkorchid2",
              'Cluster 8' = "aliceblue", "Cluster 9" = "bisque4")
)

png("Dioxines_heatmap.png",
    width = 10*300,        # 5 x 300 pixels
    height = 3*1000,
    res = 300,            # 300 pixels per inch
    pointsize = 8)        # smaller font size
pheatmap(PER_M2, cellwidth = 20, cellheight = 10, fontsize_row = 9,
         display_numbers =TRUE, number_format = "%.0f",
         annotation_row = my_col_DIOXINE9,
         annotation_col = vertical_col_DIOXINE8,
         cutree_rows = 9,
         cutree_cols = 8,
         color = colorRampPalette(c("royalblue1", "white", "orangered"))(100),
         annotation_colors = ann_colors)
dev.off()
```

9.2.2 R-Module 2 - Cluster analysis for dl-PCB

Result see Figure 12

```
library(readxl)
library(pheatmap)
library(dendextend)

DATA <- read_xlsx("UBA-EURL datafile_hardcoded DF.xlsx",
                  col_names = TRUE, na = c("NA", "ND"),
                  range = "Data_2018-11-16!AU3:BF282")

colnames(DATA) <- DATA[1,]
DATA <- DATA[-1,]

DATA_names <- read_xlsx("UBA-EURL datafile_hardcoded DF.xlsx",
                        col_names = TRUE, na = c("NA", "ND"),
                        range = "Data_2018-11-16!H4:J282")
DATA_ID <- read_xlsx("UBA-EURL datafile_hardcoded DF.xlsx",
                     col_names = TRUE, na = c("NA", "ND"),
                     range = "Data_2018-11-16!A4:A282")

rownames(DATA) <- paste(as.vector(as.matrix(DATA_ID[,1])),
                        as.vector(as.matrix(DATA_names$CaseContext)),
                        as.vector(as.matrix(DATA_names$Sample, DATA_names[,3])),
                        sep="_")

#extract DL PCB's
DATAXX <- as.data.frame(apply(DATA, 2, FUN = function(X) as.numeric(X)))

rownames(DATAXX) <- rownames(DATA)

# transforme detection limits
for (j in 2: length(DATAXX))
{
  i <- grep("<", as.vector(DATA[, j]))
  #ii <- which(!grepl("<", DATAXXX[, j]))
  DATAXX[i,j] <- as.numeric(sub("<", "",DATA[i,j]))/2
  DATAXX[,j] <- as.numeric(DATAXX[,j])

}

#eliminate rows with only NA values in the complete row
DATA_DL_PCB <- DATAXX[which(apply(DATAXX, 1, sum, na.rm = TRUE) > 0),]
# percentage values
SUM <- apply(DATA_DL_PCB, 1,
             FUN = function(X) sum(as.numeric(X), na.rm = TRUE))
PER <- data.frame(Sample = rownames(DATA_DL_PCB),
                  apply(DATA_DL_PCB, 2,
                        FUN = function(X) as.numeric(X)/SUM) * 100)

rownames(PER) <- rownames(DATA_DL_PCB)
#########################################################
### customizing and plotting heatmap
#########################################################
PER_M <- as.matrix(round(PER[,-1], digits = 1))

rownames(PER_M) <- rownames(PER)
# replace Na values with zeros, is not very scientific
PER_M[is.na(PER_M)] <- 0

#exclude samples without measurement of the PCB
PER_M2 <- PER_M[which(is.finite(apply(PER[,-1], 1, sum))),]
#find duplicates
rownames(PER_M2)[anyDuplicated(rownames(PER_M2))] <-
paste(rownames(PER_M2)[anyDuplicated(rownames(PER_M2))],
                                                        "B", sep = " ")
```

```
my_hclust_PCB <- hclust(dist(PER_M2), method = "complete")
as.dendrogram(my_hclust_PCB) %>%
  plot(horiz = TRUE)

(my_col_PCB5 <- cutree(tree = as.dendrogram(my_hclust_PCB), k = 5))

my_col_PCB5 <- data.frame(Cluster = paste("Cluster", my_col_PCB5, sep = " "))

head(my_col_PCB5)
rownames(my_col_PCB5) <- rownames(PER_M2)

my_sample_col <- data.frame(Sample = c("Group 3", "Group 5", "Group 5", "Group 5",
                                       "Group 2", "Group 5",
                                       "Group 1", "Group 5",
                                       "Group 4", "Group 5", "Group 5", "Group 5"))
row.names(my_sample_col) <- colnames(PER_M2)

# Specify colors for annotation
ann_colors = list(
  Sample = c('Group 1' = "#1B9E77", 'Group 2' = "indianred2",
             'Group 3' = "orange", 'Group 4' = "peachpuff", 'Group 5' =
"deepskyblue4"),
  Cluster = c('Cluster 1' = "#7570B3", 'Cluster 2' = "#E7298A",
              'Cluster 3' = "#66A61E", 'Cluster 4' = "yellow", 'Cluster 5' =
"chocolate")
)
png("DL_heatmap.png",
    width = 8*300,        # 5 x 300 pixels
    height = 6*300,
    res = 300,            # 300 pixels per inch
    pointsize = 8)        # smaller font size
pheatmap(PER_M2, cellwidth = 20, cellheight = 10, fontsize_row = 9,
         display_numbers =TRUE, number_format = "%.0f",
         cutree_rows = 5,
         cutree_cols = 5,
         annotation_row = my_col_PCB5, annotation_col = my_sample_col,
         color = colorRampPalette(c("royalblue1", "white", "orangered"))(100),
         annotation_colors = ann_colors)
dev.off()
```

9.2.3 R-Module 3 - Cluster analysis for indicator PCB (Group ndl-PCB)

Result see Figure 13

```
library(readxl)
library(pheatmap)
library(dendextend)


DATA <- read_xlsx("UBA-EURL datafile_hardcoded DF.xlsx",
                  col_names = TRUE, na = c("NA", "ND"),
                  range = "Data_2018-11-16!BN3:BS282")
head(DATA)
colnames(DATA) <- DATA[1,]
DATA <- DATA[-1,]

DATA_names <- read_xlsx("UBA-EURL datafile_hardcoded DF.xlsx",
                        col_names = TRUE, na = c("NA", "ND"),
                        range = "Data_2018-11-16!H4:J282")
DATA_ID <- read_xlsx("UBA-EURL datafile_hardcoded DF.xlsx",
                     col_names = TRUE, na = c("NA", "ND"),
                     range = "Data_2018-11-16!A4:A282")

rownames(DATA) <- paste(as.vector(as.matrix(DATA_ID[,1])),
                        as.vector(as.matrix(DATA_names$CaseContext)),
                        as.vector(as.matrix(DATA_names$Sample, DATA_names[,3])),
                        sep="_")

DATAXX <- as.data.frame(apply(DATA, 2, FUN = function(X) as.numeric(X)))

rownames(DATAXX) <- rownames(DATA)

# transform detection limits
for (j in 2: length(DATAXX))
{
  i <- grep("<", as.vector(DATA[, j]))
  #ii <- which(!grepl("<", DATAXXX[, j]))
  DATAXX[i,j] <- as.numeric(sub("<", "",DATA[i,j]))/2
  DATAXX[,j] <- as.numeric(DATAXX[,j])

}

#eliminate rows with only NA values in the complete row
DATA_DL_PCB <- DATAXX[which(apply(DATAXX, 1, sum, na.rm = TRUE) > 0),]
# percentual values
SUM <- apply(DATA_DL_PCB, 1,
             FUN = function(X) sum(as.numeric(X), na.rm = TRUE))
PER <- data.frame(Sample = rownames(DATA_DL_PCB),
                  apply(DATA_DL_PCB, 2,
                        FUN = function(X) as.numeric(X)/SUM) * 100)

rownames(PER) <- rownames(DATA_DL_PCB)
#########################################################
### customizing and plotting heatmap
#########################################################
PER_M <- as.matrix(round(PER[,-1], digits = 1))

rownames(PER_M) <- rownames(PER)
# replace Na values with zeros, is not very scientific
PER_M[is.na(PER_M)] <- 0

#exclude samples without measurement of the PCB
PER_M2 <- PER_M[which(is.finite(apply(PER[,-1], 1, sum))),]
#find duplicates
rownames(PER_M2)[anyDuplicated(rownames(PER_M2))] <-
paste(rownames(PER_M2)[anyDuplicated(rownames(PER_M2))],
                                                        "B", sep = " ")
```

```
my_hclust_PCB <- hclust(dist(PER_M2), method = "complete")
as.dendrogram(my_hclust_PCB) %>%
  plot(horiz = TRUE)


(my_col_PCB5 <- cutree(tree = as.dendrogram(my_hclust_PCB), k = 5))

my_col_PCB5 <- data.frame(Cluster = paste("Cluster", my_col_PCB5, sep = " "))

rownames(my_col_PCB5) <- rownames(PER_M2)


#vertical
vertical_DIOXINE <- hclust(dist(data.frame(t(PER_M2))), method = "complete")
as.dendrogram(vertical_DIOXINE) %>%
  plot(horiz = FALSE)

(vertical_col_DIOXINE <- cutree(tree = as.dendrogram(vertical_DIOXINE), k = 2))

vertical_col_DIOXINE <- data.frame(Sample = ifelse(test = vertical_col_DIOXINE ==
1,
                                                   yes = "Group 1", no = "Group
2"))

ann_colors = list(
  Sample = c('Group 1' = "#1B9E77", 'Group 2' = "indianred2"),
  Cluster = c('Cluster 1' = "#7570B3", 'Cluster 2' = "#E7298A",
              'Cluster 3' = "#66A61E", 'Cluster 4' = "yellow", 'Cluster 5' =
"chocolate")
)

png("NDL_heatmap.png",
    width = 8*300,        # 5 x 300 pixels
    height = 10*300,
    res = 300,            # 300 pixels per inch
    pointsize = 8)        # smaller font size
pheatmap(PER_M2, cellwidth = 20, cellheight = 10, fontsize_row = 9,
         display_numbers =TRUE, number_format = "%.0f",
         cutree_rows = 5,
         cutree_cols = 2,
         annotation_row = my_col_PCB5, annotation_col = vertical_col_DIOXINE,
         color = colorRampPalette(c("royalblue1", "white", "orangered"))(100),
         annotation_colors = ann_colors)
dev.off()
```

9.3 References for the samples included into the UBA/EURL dataset

Behnisch, P. (1997). Isomerenspezifische Untersuchungen über Eintrag, Verbleib und Risikoabschätzung der nicht-, mono- und di-ortho-chlorierten Biphenyle in der Umwelt - pp 293/294. In Fakultät für Chemie und Pharmazie (Tübingen, Germany: Eberhard-Karls-Universität Tübingen).

Brunner, H. (1990a). Untersuchungen zu Herkunft und Vorkommen polychlorierter Dibenzodioxine und Dibenzofurane in der Umwelt - p 93 (Tübingen, Germany).

Brunner, H. (1990b). Untersuchungen zu Herkunft und Vorkommen polychlorierter Dibenzodioxine und Dibenzofurane in der Umwelt - p 103 (Tübingen, Germany).

Brunner, H. (1990c). Untersuchungen zu Herkunft und Vorkommen polychlorierter Dibenzodioxine und Dibenzofurane in der Umwelt - p 116 (Tübingen, Germany).

Carroll, W.F., Berger, T.C., Borrelli, F.E., Garrity, P.J., Jacobs, R.A., Ledvina, J., Lewis, J.W., McCreedy, R.L., Smith, T.P., Tuhovak, D.R., et al. (2001). Characterization of emissions of dioxins and furans from ethylene dichloride, vinyl chloride monomer and polyvinyl chloride facilities in the United States. Consolidated report. Chemosphere 43, 689-700.

Covaci, A., Jake Ryan, J., and Schepens, P. (2002). Patterns of PCBs and PCDD/PCDFs in chicken and pork fat following a Belgian food contamination incident. Chemosphere 47, 207-217.

CVUA (1998). Butterproben 1998 (butter samples) - CPP context, F. Chemisches und Veterinäruntersuchungsamt, Germany, ed. (Freiburg, Germany).

CVUA (1999a). Analysis of meat samples in 1999 - Belgian chicken accident context; Dioxin2004, F. Chemisches und Veterinäruntersuchungsamt, Germany, ed. (Freiburg, Germany).

CVUA (1999b). Untersuchung eines Truthahn-Gewürzbratens aus Belgien auf Dioxine (Expertise on dioxin contamination in turkey) - Belgian chicken crisis context, F. Chemisches und Veterinäruntersuchungsamt, Germany, ed. (Freiburg, Germany).

CVUA (2007). Grass - Herbertingen/Germany after scrap fire, F. Chemisches und Veterinäruntersuchungsamt, Germany, ed. (Freiburg, Germany).

Duinker, J.C., and Hillebrand, M.T.J. (1983). Characterization of PCB components in Clophen formulations by capillary GC-MS and GC-ECD techniques. Environmental Science & Technology 17, 449-456.

EU Commission (2012). Dried feed (results of SGS_IAC12-06326).

EU Commission (2013). Mixed vegetable oil, E. Commission, ed.

Ferrario, J., Byrne, C., and Schaum, J. (2007). Concentrations of polychlorinated dibenzo-p-dioxins in processed ball clay from the United States. Chemosphere 67, 1816-1821.

Ferrario, J., Byrne, C., Winters, D., Boone, T., Vigo, C., and Dupuy, A. (2003). Chlorinated dioxins and furans from kelp and copper sulfate: Initial investigations of dioxin formation in mineral feed supplements. Organohalogen Compd 63, 183-186.

Ferrario, J.B., Byrne, C.J., and Cleverly, D.H. (2000). 2,3,7,8-Dibenzo-p-dioxins in Mined Clay Products from the United States:? Evidence for Possible Natural Origin. Environmental Science & Technology 34, 4524-4532.

Fiedler, H., Cooper, K., Bergek, S., Hjelt, M., Rappe, C., Bonner, M., Howell, F., Willett, K., and Safe, S. (1998). PCDD, PCDF, and PCB in farm-raised catfish from southeast United States - Concentrations, sources, and CYP1A induction. Chemosphere 37, 1645-1656.

Frame, G.M., Cochran, J.W., and Bøwadt, S.S. (1996). Complete PCB congener distributions for 17 Aroclor mixtures determined by 3 HRGC systems optimized for comprehensive, quantitative, congener-specific analysis. Journal of Separation Science 19, 657-668.

Grochowalski, A., Wegiel, M., Maslanka, A., Chrzaszcz, R., and Chyc, M. (2015). Dioxin Catalytic Formation in irepolaces and Coal Heating Stoves in Poland. Organohalogen Compd 77, 202-205.

Hagenmaier, H., and Berchtold, A. (1986). Analysis of waste from production of Na-pentachlorophenolate for polychlorinated dibenzodioxins (PCDD) and dibenzofurans (PCDF). Chemosphere 15, 1991-1994.

Hagenmaier, H., and Brunner, H. (1987). Isomerspecific analysis of pentachlorophenol and sodium pentachlorophenate for 2, 3, 7, 8-substituted PCDD and PCDF at sub-ppb levels. Chemosphere 16, 1759-1764.

Hayward, D.G., Nortrup, D., Gardner, A., and Clower, M. (1999). Elevated TCDD in Chicken Eggs and Farm-Raised Catfish Fed a Diet with Ball Clay from a Southern United States Mine. Environmental Research 81, 248-256.

Heres, L., Hoogenboom, R., Herbes, R., Traag, W., and Urlings, B. (2010). Tracing and analytical results of the dioxin contamination incident in 2008 originating from the Republic of Ireland. Food Additives & Contaminants: Part A 27, 1733-1744.

Holt, E., Weber, R., Stevenson, G., and Gaus, C. (2010). Polychlorinated Dibenzo-p-Dioxins and Dibenzofurans (PCDD/Fs) Impurities in Pesticides: A Neglected Source of Contemporary Relevance. Environmental Science & Technology 44, 5409-5415.

Hoogenboom, L.A.P., Kan, C.A., Bovee, T.F.H., van der Weg, G., Onstenk, C., and Traag, W.A. (2004a). Residues of dioxins and PCBs in fat of growing pigs and broilers fed contaminated feed. Chemosphere 57, 35-42.

Hoogenboom, L.A.P., Van Eijkeren, J.C.H., Zeilmaker, M.J., Mengelers, M.J.B., Herbes, R., Immerzeel, J., and Traag, W.A. (2007). A novel source for dioxins present in recycled fat from gelatin production. Chemosphere 68, 814-823.

Hoogenboom, R., Bovee, T., Portier, L., Bor, G., van der Weg, G., Onstenk, C., and Traag, W. (2004b). The German bakery waste incident; use of a combined approach of screening and confirmation for dioxins in feed and food. Talanta 63, 1249-1253.

Huang, J., Matsumura, T., Yu, G., Deng, S., Yamauchi, M., Yamazaki, N., and Weber, R. (2011). Determination of PCBs, PCDDs and PCDFs in insulating oil samples from stored Chinese electrical capacitors by HRGC/HRMS. Chemosphere 85, 239-246.

Huang, J., Gao, J., Yu, G., Yamazaki, N., Deng, S., Wang, B., and Weber, R. (2015). Unintentional formed PCDDs, PCDFs, and DL-PCBs as impurities in Chinese pentachloronitrobenzene products. Environmental Science and Pollution Research 22, 14462-14470

Jiang, K., Li, L., Chen, Y., and Jin, J. (1997). Determination of PCDD/Fs and dioxin-like PCBs in Chinese commercial PCBs and emissions from a testing PCB incinerator. Chemosphere 34, 941-950.

Kim, M., Kim, D.-G., Choi, S.-W., Guerrero, P., Norambuena, J., and Chung, G.-S. (2011). Formation of polychlorinated dibenzo-p-dioxins/dibenzofurans (PCDD/Fs) from a refinery process for zinc oxide used in feed additives: A source of dioxin contamination in Chilean pork. Chemosphere 82, 1225-1229.

Lindert, M., and Fiedler, H. (1999). Auftreten von PCDD/PCDF in Brandfällen. UWSF - Z Umweltchem Ökotox 11, 39-47.

Liu, W., Zheng, M., Wang, D., Xing, Y., Zhao, X., Ma, X., and Qian, Y. (2004). Formation of PCDD/Fs and PCBs in the process of production of 1,4-dichlorobenzene. Chemosphere 57, 1317-1323.

Llerena, J.J., Abad, E., Caixach, J., and Rivera, J. (2003). Cholin chloride congeners in: An episode of dioxin contamination in feedingstuff: the choline chloride case. In Chemosphere, E.A.f.H.F.i.A. 2018, ed., pp. 679-683.

Malisch, R. (2000). Increase of the PCDD/F-contamination of milk, butter and meat samples by use of contaminated citrus pulp. Chemosphere 40, 1041-1053.

Masunaga, S., Takasuga, T., and Nakanishi, J. (2001). Dioxin and dioxin-like PCB impurities in some Japanese agrochemical formulations. Chemosphere 44, 873-885.

Ni, Y., Zhang, Z., Zhang, Q., Chen, J., Wu, Y., and Liang, X. (2005). Distribution patterns of PCDD/Fs in chlorinated chemicals. Chemosphere 60, 779-784.

Rappe, C., Kjeller, L.-O., Kulp, S.-E., de Wit, C., Hasselsten, I., and Palm, O. (1991). Levels, profile and pattern of PCDDs and PCDFs in samples related to the production and use of chlorine. Chemosphere 23, 1629-1636.

RASFF 2004/327 (2004). Fatty acid in complementary feedingstuff for production of compound feed for livestock, R.A.N. 2004/327, ed.

RASFF 2004/662 (2004). Palm fat (Spain), R.A. Notification, ed.

RASFF 2009/210 (2009). Ground dried flowers of Tagetes erecta (Marigold), R.A. Notification, ed.

RASFF 2009/1284 (2009). Sulphur in oil (Poland), R.A. Notification, ed. (LUFA Kiel_ 834537).

RASFF 2010/346 (2010). Feed premixes and supplements on basis of vitamin A palmitat (origin: China), R.A. Notification, ed.

RASFF 2010/0519 (2010). RASFF Alert Notification: 2010/0519 - eggs, R.A. Notification, ed.

Reeuwijk, N.M., Talidda, A., Malisch, R., Kotz, A., Tritscher, A., Fiedler, H., Zeilmaker, M.J., Kooijman, M., Wienk, K.J.H., Traag, W.A., et al. (2013). Dioxins (polychlorinated dibenzo-p-dioxins and polychlorinated dibenzo-furans) in traditional clay products used during pregnancy - Additional information: clay and human milk congeners. Chemosphere 90, 1678-1685.

RIKILT (1998). RIKILT database "Citrus pulp, Brazil".

RIKILT (1999). RIKILT database ""Kaolinitic clay used premix for feed production (1999)" (Germany).

RIKILT (2011a). RIKILT database "Beet pulp - Germany".

RIKILT (2011b). RIKILT database "Fatty acids, deinking of paper - Germany".

RIKILT (2013). RIKILT database "Chicken meat contaminated by unknown source (Chile)".

RIKILT (No Year-a). Rikilt database "Corn from Ukraine".

RIKILT (No Year-b). RIKILT database no 133877; potato peels.

RIKILT (No year-c). RIKILT databse (not confirmed) "Feed for pork - Chile".

Rushneck, D.R., Beliveau, A., Fowler, B., Hamilton, C., Hoover, D., Kaye, K., Berg, M., Smith, T., Telliard, W.A., Roman, H., et al. (2004). Concentrations of dioxin-like PCB congeners in unweathered Aroclors by HRGC/HRMS using EPA Method 1668A. Chemosphere 54, 79-87.

Santl, H., Gruber, L., and Stöhrer, E. (1994). Some new sources of polychlorinated dibenzodioxins (PCDDs) and dibenzofurans (PCDFs) in waste papers and recycled pulps. Chemosphere 29, 1995-2003.

Schecter, A., Jiang, K., Päpke, O., Fürst, P., and Fürst, C. (1994). Comparison of dibenzodioxin levels in blood and milk in agricultural workers and others following pentachlorophenol exposure in China. Chemosphere 29, 2371-2380.

Schulte, E., and Malisch, R. (1983). Berechnung der wahren PCB-Gehalte in Umweltproben. Fresenius' Journal of Analytical Chemistry 314, 545-551.

Slobs, W., Olling, M., Derks, H.J.G.M., and de Jong, A.P.J.M. (1995). Congener-specific bioavailability of PCDD/Fs and coplanar PCBs in cows: Laboratory and field measurements. Chemosphere 31, 3827-3838.

Solorzano-Ochoa, G., de la Rosa, D.A., Maiz-Larralde, P., Gullett, B.K., Tabor, D.G., Touati, A., Wyrzykowska-Ceradini, B., Fiedler, H., Abel, T., and Carroll, W.F. (2012). Open burning of household waste: Effect of

experimental condition on combustion quality and emission of PCDD, PCDF and PCB. Chemosphere 87, 1003-1008.

Takasuga, T., Kumar, K.S., Noma, Y., and Sakai, S. (2005). Chemical characterization of polychlorinated biphenyls,-dibenzo-p-dioxins, and-dibenzofurans in technical Kanechlor PCB formulations in Japan. Archives of environmental contamination and toxicology 49, 385-395.

Takasuga, T., Senthilkumar, K., Matsumura, T., Shiozaki, K., and Sakai, S. (2006). Isotope dilution analysis of polychlorinated biphenyls (PCBs) in transformer oil and global commercial PCB formulations by high resolution gas chromatography-high resolution mass spectrometry. Chemosphere 62, 469-484.

Theisen, J., Funcke, W., Balfanz, E., and König, J. (1989). Determination of PCDFs and PCDDs in fire accidents and laboratory combustion tests involving PVC-containing materials. Chemosphere 19, 423-428.

Theisen, J., Maulshagen, A., and Fuchs, J. (1993). Organic and inorganic substances in the copper slag "Kieselrot". Chemosphere 26, 881-896.

Thoma, H. (1988). PCDD/F-concentrations in chimney soot from house heating systems. Chemosphere 17, 1369-1379.

Tondeur, Y., Nestrick, T., Silva, H.A., Vining, B., and Hart, J. (2010). Analytical procedures for the determination of polychlorinated-p-dioxins, polychlorinated dibenzofurans, and hexachlorobenzene in pentachlorophenol. Chemosphere 80, 157-164.

UBA (1995). Profil 811 - Wellpappe (Italien) in Santl H, Gruber L, Meisburger K, Wolz G Identifizierung und Quantifizierung dioxin-relevanter Einträge in das zur Papierherstellung eingesetzte Altpapier. Abschlussbericht. Dezember 1995.

Université de Liège, C. (2013). Fiche d'Analyse, sample 26327BDAH-26328BQAH; assigned tocopherol from China (?).

van Larebeke, N., Hens, L., Schepens, P., Covaci, A., Baeyens, J., Everaert, K., Bernheim, J.L., Vlietinck, R., and De Poorter, G. (2001). The Belgian PCB and dioxin incident of January-June 1999: exposure data and potential impact on health. Environmental Health Perspectives 109, 265-273.

VDI (2003). Emission measurement Determination of polychlorinated dibenzo-p-dioxins (PCDDs) and dibenzofurans (PCDFs) - Dilution method, V. Richtlinien, ed.

Volland, G., and Neuwirth, A. (2005a). Datasets 753, 754, 755 from "Dioxinähnliche polychlorierte Biphenyle (PCB) und polychlorierte Dioxine/Furane (PCDD/F) im Innenraum - Modelluntersuchungen zum Einfluss PCB-belasteter Bauteile und Baustoffe auf die Innenraumluft in Abhängigkeit von baulichen Einrichtungen und klimatischen Randbedingungen".

Volland, G., and Neuwirth, A. (2005b). Tabelle II (Fugendichtung, Bau E, UBA 761) in: Dioxinähnliche polychlorierte Biphenyle (PCB) und polychlorierte Dioxine/Furane (PCDD/F) im Innenraum - Modelluntersuchungen zum Einfluss PCB-belasteter Bauteile und Baustoffe auf die Innenraumluft in Abhängigkeit von baulichen Einrichtungen und klimatischen Randbedingungen.

Wang, P., Zhang, Q., Lan, Y., Shiai Xu, S., Gao, R., Li, G., Zhang, H., Shang, H., Ren, D., Zhu, C., et al. (2014). Dioxins contamination in the feed additive (feed grade cupric sulfate) tied to chlorine industry. Scientific Reports 4: 5975 | 1-5.

Wang, X., Ni, Y., Zhang, H., Zhang, X., and Chen, J. (2012). Formation and Emission of PCDD/Fs in Chinese Non-Wood Pulp and Paper Mills. Environmental Science & Technology 46, 12234-12240.

Weber, R., Hollert, H., Kamphues, J., Ballschmiter, K., Blepp, M., and Herold, C. (2015). Anhänge - Analyse und Trendabschätzung der Belastung der Umwelt und von Lebensmitteln mit ausgewählten POPs und Erweiterung

des Datenbestandes der POP-Dioxin-Datenbank des Bundes und der Länder mit dem Ziel pfadbezogener Ursachenforschung

- Anhang 1: PCB im Bausektor und daraus freigesetzte Emissionen - eine Bestandsaufnahme und Neubewertung

- Anhang 2: Forschungsbedarf und Handlungsbedarf zur Verringerung und Vermeidung der Schadstoffbelastung

- Anhang 3: Fallauswertung von Dioxin- und PCB-Belastungen in Eiern aus Freilandhaltungen

- Anhang 4: Präsentation PCB-und Dioxin-Belastung der Umwelt und von Lebensmitteln

- Anhang 5: Erweiterung des Datenbestandes der Dioxin-Datenbank des Bundes und der Länder

- Anhang 6: Tagungsbeiträge auf dem 34th International Symposium on Halogenated Persistent Organic Pollutants. In UBA Dokumentation, N. Umweltforschungsplan des Bundesministeriums für Umwelt, Bau und Reaktorsicherheit, ed. (Fachgebiet IV 2.1 International

Xu, Y., Zhang, Q., Wu, W., and Li, W. (2000). Patterns and levels of PCDD/F in a Chinese graphite electrode sludge. Chinese Science Bulletin 45, 1471.

Zhang, T.T., Fiedler, H., Yu, G., Ochoa, G.S., Carroll, W.F., Gullett, B.K., Marklund, S., and Touati, A. (2011). Emissions of unintentional persistent organic pollutants from open burning of municipal solid waste from developing countries. Chemosphere 84, 994-1001.