



SHAPLEY ADDITIVE EXPLANATIONS OF INDICATOR PCB-138 DISTRIBUTION IN BREAST MILK

Andreja Stojić^{1,2*},
Marijana Matek Sarić³,
Snježana Herceg Romanić⁴

¹Institute of Physics Belgrade,
National Institute of the Republic of Serbia,
University of Belgrade,
Belgrade, Serbia

²Environment and Sustainable Development,
Singidunum University, Belgrade,
Serbia

³Department of Health Studies,
University of Zadar,
Zadar, Croatia

⁴Institute for Medical Research and
Occupational Health,
Zagreb, Croatia

Abstract:

Breastfeeding provides numerous health benefits for newborns by meeting the infants' nutritional needs and supporting associated immunological protection. Maternal milk is high in fat, and therefore, represents a very suitable media for the bioaccumulation of lipophilic pollutants such as organochlorine pesticides (OCPs) and polychlorinated biphenyls (PCBs). This makes breast milk the primary source of the infant's postnatal exposure to persistent toxic xenobiotics. In this study, we applied a novel SHapley Additive exPlanations (SHAP) method to investigate the key parameters that govern distribution of PCB-138, an indicator of non-dioxin congeners, in the mother's milk. According to the accuracy metrics, the eXtreme Gradient Boosting regression was successfully employed, with relative error below 20% and high correlation coefficient ($r=0.97$), for finding the relationships between PCB-138 and other non-dioxin congeners, mother's age, and number of births. According to the results, PCB-156, PCB-180, HCB, HCH and PCB-118 have the major, while PCB-28, PCB-52 and PCB-189 have the minor impact on PCB-138 distribution in breast milk. Similar behavior of contaminants, which belong to the both indicator congener group (-28, -52, -180) and toxicologically relevant PCBs (-118, -189), was recognized. The SHAP conclusions were only partially consistent with the results of correlation analysis suggesting that POPs exhibit non-linear dynamics and interrelations. Therefore, the current knowledge on contamination of complex biomatrices would benefit from further detailed analyses of pollutant intermittent relationships.

Keywords:

human biomonitoring, persistent organic pollutants (POPs), organochlorine pesticides (OCPs), polychlorinated biphenyls (PCBs), Shapley Additive exPlanations (SHAP).

1. INTRODUCTION

Organochlorine pesticides (OCPs) and polychlorinated biphenyls (PCBs) belong to a group of compounds known as Persistent Organic Pollutants (POPs). Due to their common properties such as persistence, bioaccumulation, toxicity, and ability to be transported via air over long distances, the production and use of POPs are eliminated and/or restricted by the Stockholm Convention. Additionally, due to their persistence and lipophilic properties, they tend to accumulate in the fatty tissues of humans and animals.

Correspondence:

Andreja Stojić

e-mail:

andreja.stojic@ipb.ac.rs



Research has shown that organochlorine compounds, as endocrine modulators, affect human health. The toxic potential of individual compounds is different, but the overall toxic effect is additive. According to epidemiological and laboratory studies, depending on the compound, these toxins can affect the development of cancer, cause allergy and hypersensitivity, as well as central and peripheral nervous system damage, reproductive disorders and immune system disorders. Humans and wildlife are exposed to PCBs and organochlorine pesticides in the polluted environment via air, water, sediment soil and food, by ingestion, inhalation or skin contact. Previous studies suggest that more of these compounds is introduced into the body by food intake than by inhalation or through the skin. It is also known that unborn children are exposed to organochlorine compounds that pass through the placenta. Biological monitoring of these compounds in the body requires invasive techniques referring to the surgical extraction of adipose tissue. The alternative monitoring approach and one of the mostly used non-invasive methods for the assessment of human exposure is the analysis of human milk POPs content.

A number of literature sources reported on POPs interrelations in the breast milk or/and their dependence on parameters like mother's age and weight loss, childbirth, dietary habits and occupation [1, 2, 3, 4]. Our previous investigations were aimed at investigating the presence of OCPs and PCBs in human milk as well as their mutual interrelations and associations with the mother's age and parity using machine learning (ML) algorithms [5, 6, 7]. In this study, we applied the additive feature attribution method – Shapley Additive Explanations (SHAP), which offers uniquely consistent and locally accurate attribution values in comparison with the conventional attribution methods and has become increasingly popular as a tool for predicting environmental phenomena [8]. We aimed to obtain a detailed insight into the parameters including organochlorine pollutants, mother's age and parity, that shape the PCB-138 distribution in the mother's milk. We chose this congener since it has been considered among PCB₆ group as indicator for the PCB behavior patterns in various samples and it has been also taken as the most suitable target compound for evaluating non-dioxin-like PCBs. The classification of non-dioxin- and dioxin-like congeners is used in the European food and feed regulation where dioxin-like PCBs resemble biochemical and toxicological characteristics of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD).

2. MATERIAL AND METHODS

Sampling

Breast milk samples were collected from 152 healthy primiparae, secundiparae and multiparae (third child delivery), aged between 19 and 45, living in Zadar (Croatia). The mothers had no history of accidental or occupational exposure to the analyzed POPs. More details about sampling were described previously [7].

Chemical analysis of PCBs and OCPs

The analytical procedure was described in details previously [5]. Briefly, two subsamples of each unfrozen milk sample (5 g) were extracted twice with a mixture of chloroform and methanol and dried under a nitrogen flow. Subsequently, milk fat was dissolved in n-hexane and purification and fractionation were performed.

Seven OCPs [hexachlorobenzene (HCB), hexachlorocyclohexane isomers (α -, β -, and γ -HCH), 1,1,1-trichloro-2,2-di(4-chlorophenyl)ethane (p,p' -DDT), 1,1-dichloro-2,2-di(4-chlorophenyl)ethylene (p,p' -DDE), 1,1-dichloro-2,2-di(4-chlorophenyl)ethane (p,p' -DDD)], 17 PCB congeners [PCB-28, PCB-52, PCB-101, PCB-138, PCB-153, PCB-180 (six indicator congeners), PCB-105, PCB-114, PCB-118, PCB-123, PCB-156, PCB-157, PCB-167, PCB-189 (eight mono ortho congeners), PCB-60, PCB-74 and PCB-170] were analysed. High-resolution gas chromatography with electron capture detector (s) was performed on a CLARUS 500 chromatograph using two capillary columns (Restek, Bellefonte, PA, USA) simultaneously: (1) 60m \times 0.25 mm, Rtx-5 film thickness of 0.25 μ m, and (2) 30m \times 0.25 mm, Rtx-1701 film thickness of 0.25 μ m. The LODs for the analyzed compounds were 0.5 ng g⁻¹ milk fat for PCB congeners, 0.1 ng g⁻¹ milk fat for α -HCH and HCB, 0.2 ng g⁻¹ milk fat for p,p' -DDE, 0.3 ng g⁻¹ milk fat for β -HCH, γ -HCH and p,p' -DDD, and 0.6 ng g⁻¹ milk fat for p,p' -DDT. The average recoveries for PCBs ranged between 58% and 86% and for organochlorine pesticides between 59% and 92%. Method reproducibility expressed as relative standard deviation was between 6% and 22%, and 7% and 24% for PCBs and OCPs, respectively.

Data analysis

The relationships between PCB-138 in breast milk and all other measured parameters were modeled by using XGBoost regression. The details on the method are given elsewhere [4]. In this study, we used Python XGBoost implementation. The dataset was split into training (80%) and validation (20%) sets. Hyperparameter tuning



was implemented using a brute-force grid search and 10-fold stratified cross-validation. The best performing hyperparameter values were used for the final model.

The explainability of the produced XGBoost model that operates with high-dimensional input data in a non-linear fashion was obtained by using explainable artificial intelligence method SHapley Additive exPlanations (SHAP) [5]. Based on the game theory, the Shapley explanations represent the only possible locally accurate and globally consistent feature attribution values. The captured attributed importance of a feature is visually presented as SHAP summary plot.

3. RESULTS AND DISCUSSION

The concentrations of organochlorines found in the milk samples of primiparae, secundiparae and multiparae aged between 19 and 45 (with an average age of 30 years) are given in Fig. 1. The most abundant pollutants were the by-products of lindane (γ -HCH and HCB), the organochlorine pesticide metabolites (p,p' -DDE and p,p' -DDD), and heavy hexa- and hepta-chloro non-dioxin like congeners (PCB-138, PCB-153, PCB-180 and PCB-170).

Significant linear correlation coefficients ($r \geq 0.90$) were found between the following pairs of the investigated variables: PCB-170-PCB-138; PCB-170-PCB-153; PCB-170-PCB-180; PCB-153-PCB-180 and PCB-153-PCB-138 (Fig. 2).

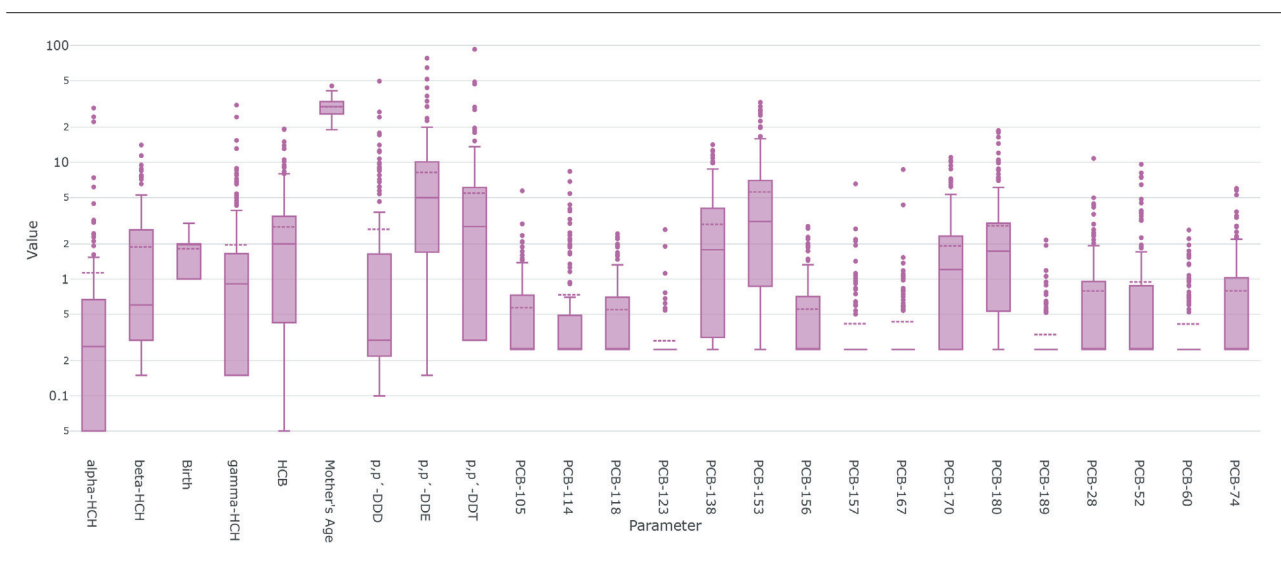


Fig. 1. Measured parameter box plots.

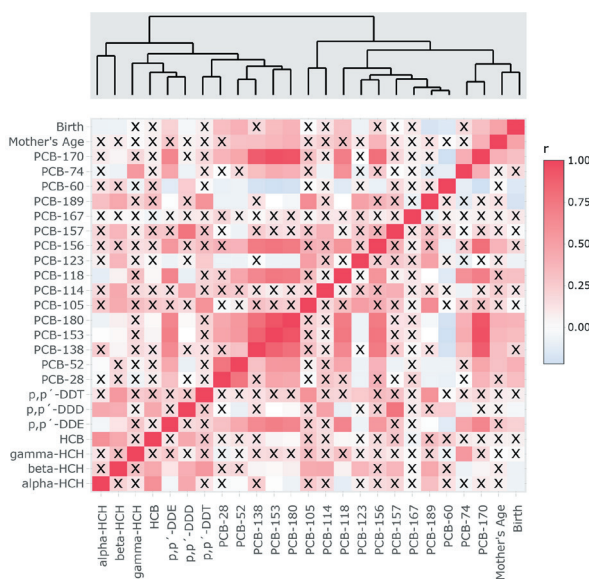


Fig. 2. Obtained parameter correlation matrix.



The results indicated that the listed pollutants have similar molecular structure and metabolic pathways, but we assumed that the advanced methods apart from commonly applied correlation matrices could be employed to deepen the current understanding of PCB-138 behaviour patterns in breast milk.

For investigating the non-linear relationships between PCB-138 and other congeners, mother's age, and number of births, the XGBoost regression analysis was successfully employed, with a relative error below 20% and high correlation coefficient ($r=0.97$) (Fig. 3). As shown by the highest positive (up to 4) and negative (up to -1.5) SHAP values, the most important variables that shaped PCB-138 behaviour pattern in the examined milk samples were PCB-170 and PCB-153 (Fig. 4).

The strongest influences were observed to be related to the elevated concentrations of the listed pollutants suggesting that mono-chlorine congeners are more prone to bioaccumulation in human milk compared to other PCBs.

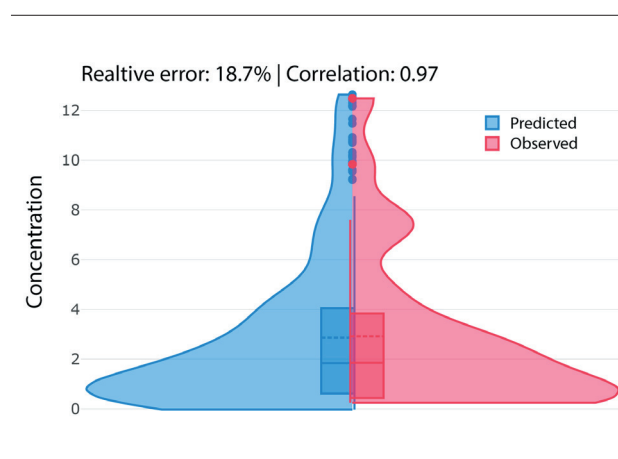


Fig. 3. XGBoost evaluation.

Position of the halogen substituting biphenyl ring of PCB-153, PCB-170 and PCB-153 provide such structure and rigidity of the molecules that facilitate PCBs' ability to pass from blood to breast milk. The other molecular properties including lipophilicity, molecular diameter and weight, and the number of attached halogens appeared to be less determinative for the PCB partitioning between blood and milk [9].

The SHAP analyses revealed less important impacts of PCB-156, PCB-180 and PCB-118 on the PCB-138 behaviour patterns in milk samples. The pollutants belong to non-dioxin-like/indicator congener group (-138, -153 and -180) and toxicologically relevant PCBs (-118 and -156), which elicit aryl hydrocarbon receptor-mediated

biochemical and toxic responses and resist in the food chain. Minor negative functional dependency was observed between PCB-153 and the elevated levels of low-chlorinated congeners (PCB-52 and PCB-28), which are more volatile and susceptible to metabolic breakdowns and excretion.

Numerous studies reported that POPs are eliminated from the body during the breastfeeding, which results in the lower levels of the pollutants in the breast milk of multipara. In addition, it has been usually reported that the milk of elder mothers contained higher concentrations of organochlorine xenobiotics [10]. As indicated by low SHAP values approaching zero, in this study, no significant influences of the mother's age and parity on the PCB-153 patterns was recorded (Fig. 3). For prediction of the POPs bioaccumulation in breast milk, the exposure routes including dietary habits, residential and working environment as well as the mother's childhood nursing history should deserve more attention in future research.

4. CONCLUSION

The health burden of organochlorines in breast milk represents an issue of global concern because of the POPs adverse impacts on human health, particularly sensitive sub-populations such as women and children. In this study, we presented promising explainable artificial intelligence methodology (XGBoost and SHAP) with the aim to investigate the organochlorine patterns as well as the PCB-138 dependence on the mother's age and child delivery. According to the results, similarly structured molecules, which belong to both non-dioxin-like/indicator congener group (-28, -52, -180) and toxicologically relevant PCBs (-118, -189, -156) as well as *p,p'*-DDT metabolite (*p,p'*-DDE) have impacts on PCB-138 distribution. No significant functional dependencies of the PCB-138 patterns and the maternal age and parity was observed suggesting that the impacts of dietary habits and health burden POPs in the residential and working environment should be more investigated in the future. The herein applied analyses could serve as a promising methodology for future epidemiological investigations and human health protection.

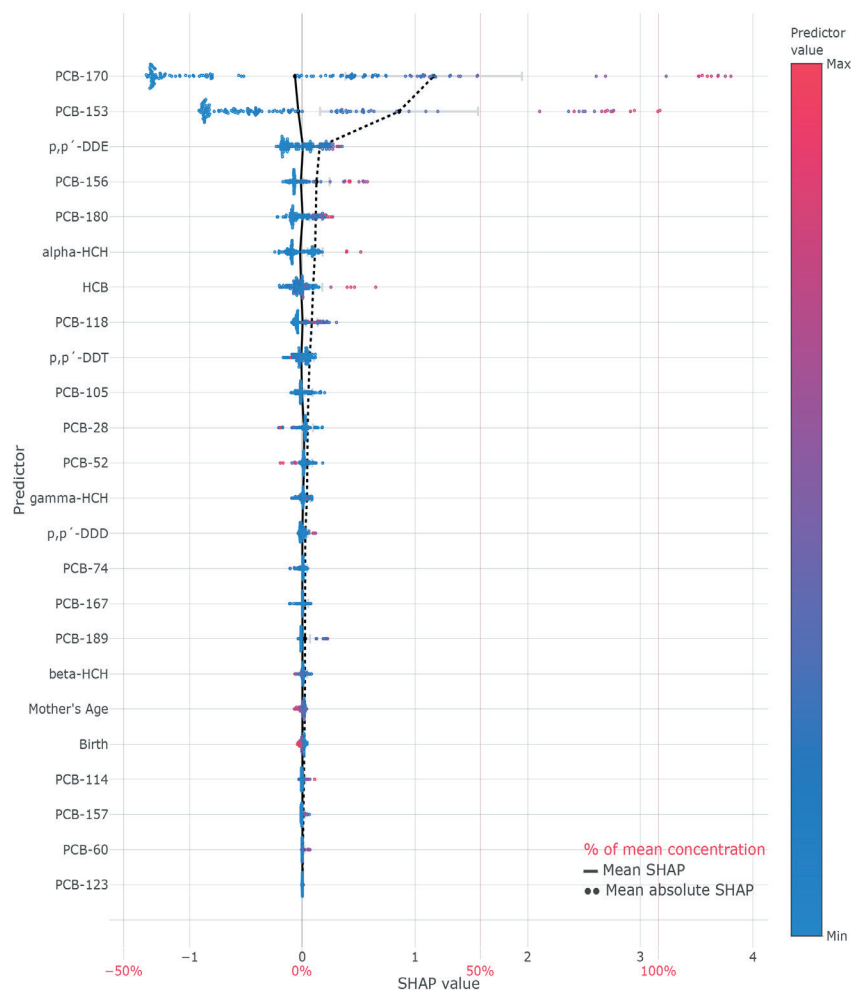


Fig. 4. PCB-138 SHAP summary.

ACKNOWLEDGMENT

The authors acknowledge funding provided by the Science Fund of the Republic of Serbia #GRANT No. 6524105, AI – ATLAS.

REFERENCES

- [1] J. Fång, E. Nyberg, U. Winnberg, A. Bignert, and Å. Bergman, "Spatial and temporal trends of the Stockholm Convention POPs in mothers' milk — a global review." *Environ. Sci. Pollut. Res.*, 2015, pp. 8989–9041.
- [2] United Nations Environment Programme, 2013. Results of the global survey on concentrations in human milk of persistent organic pollutants by the United Nations Environment Programme and the World Health Organization, Conference of the Parties to the Stockholm Convention on Persistent Organic Pollutants, Sixth meeting, Geneva, 28 April–10 May
- [3] S. Lignell, A. Winkvist, F. Bertz, K.M. Rasmussen, A. Glynn, M. Aune, and H.K. Brekke "Environmental organic pollutants in human milk before and after weight loss." *Chemosphere*, 2016, pp. 96–102.
- [4] A. Polder, J.U. Skaare, E. Skjerve, K.B. Løken, and M. Eggesbø "Levels of chlorinated pesticides and polychlorinated biphenyls in Norwegian breast milk (2002–2006), and factors that may predict the level of contamination." *Sci. Total Environ.*, 2009, pp. 4584–4590.
- [5] D. Klinčić, S. Herceg Romanić, M. Matek Sarić, J. Grzunov, and B. Dukić, "Polychlorinated biphenyls and organochlorine pesticides in human milk samples from two regions in Croatia." *Environ. Toxicol. Pharmacol*, 2014, pp. 543–552.
- [6] D. Klinčić, S. Herceg Romanić, I. Brčić Karačonji, M. Matek Sarić, J. Grzunov Letinić, and N. Brajenović "Organochlorine pesticides and PCBs (including dl-PCBs) in human milk samples collected from multiparae from Croatia and comparison with primiparae." *Environ. Toxicol. Pharmacol*, 2016, pp. 74–79.



- [7] G. Jovanović, S. Herceg Romanić., A. Stojić, D. Klinčić, M. Matek Sarić, J. Grzunov Letinić, and A. Popović, "Introducing of modeling techniques in the research of POPs in breast milk – A pilot study." *Ecotoxicol. Environ. Saf.*, 2019, pp. 341-347.
- [8] A. Stojić, N. Stanić, G. Vuković, S. Stanišić, M. Perišić, A. Šoštarić, and L. Lazić, "Explainable extreme gradient boosting tree-based prediction of toluene, ethylbenzene and xylene wet deposition." *Sci. Total Environ.*, 2019, pp. 140-147.
- [9] G. Vasios, A. Kosmidi, O.-I. Kalantzi, A. Tsantili-Kakolidou, N. Kavantzas, S. Theocharis, and C. Gigaginis "Simple physicochemical properties related with lipophilicity, polarity, molecular size and ionization status exert significant impact on the transfer of drugs and chemicals into human breast milk." *Expert Opin. Drug Metab. Toxicol.*, 2016, pp. 1273–1278.
- [10] R. Aerts, I. Van Overmeire, A. Colles, M. Andjelković, G. Malarvannan, G. Poma, E. Den Hond, E. Van de Mieroop, M.-C. Dewolf, F. Charlet, A. Van Nieuwenhuysse, J. Van Loco, A. Covacci, "Determinants of persistent organic pollutant (POP) concentrations in human breast milk of a cross-sectional sample of primiparous mothers in Belgium", *Environ. Int.*, 2019, pp. 104979.