

This document is the accepted manuscript version of the following article:
Chiaia-Hernández, A. C., Günthardt, B. F., Frey, M. P., & Hollender, J. (2017). Unravelling
contaminants in the anthropocene using statistical analysis of liquid chromatography–high-resolution
mass spectrometry nontarget screening data recorded in lake sediments. *Environmental Science and
Technology*. <http://doi.org/10.1021/acs.est.7b03357>

Unravelling contaminants in the Anthropocene using statistical analysis of LC-HRMS non-target screening data as recorded in lake sediments

Aurea C. Chiaia-Hernández^{1*}, Barbara F. Günthardt^{1,2}, Martin P. Frey³, and Juliane
Hollender^{1,2}

¹Eawag, Swiss Federal Institute of Aquatic Science and Technology (Eawag), Switzerland

²Institute of Biogeochemistry and Pollutant Dynamics, ETH Zurich, Switzerland

³Institute of Data Analysis and Process Design, ZHAW, Winterthur, Switzerland

*Corresponding author

Überlandstrasse 133, PO Box 611, 8600 Dübendorf, Switzerland; phone: +41 58 765 5309; fax:
+41 58 765 5028; E-Mail: aurea.chiaia@eawag.ch

Abstract

The significant increase in traces of human activity in the environment worldwide provides
evidence of the beginning of a new geological era, informally named the Anthropocene. The
rate and variability of these human modifications at the local and global scale remain largely
unknown, but new analytical methods such as high-resolution mass spectrometry (HRMS)
can help to characterize chemical contamination. We therefore applied HRMS to investigate
the contamination history of two lakes in Central Europe over the preceding 100 years. A
hierarchical clustering analysis (HCA) of the collected time series data revealed more than
13,000 profiles of anthropogenic origin in both lakes, defining the beginning of large-scale
human impacts during the 1950s. Our results show that the analysis of temporal patterns of
non-target contaminants is an effective method for characterizing the contamination pattern in
the Anthropocene and an important step in prioritizing the identification of organic
contaminants not yet successfully targeted by environmental regulation and pollution

reduction initiatives. As a proof-of-concept, the success of the method was demonstrated with the identification of the pesticide imazalil, which probably originated from imported fruits. This new approach applicable to palaeoarchives can effectively be used to document the time and rate of change in contamination over time and provide additional information on the onset of the Anthropocene.

Introduction

In the last century, multiple anthropogenic effects on the Earth System, such as perturbations of the carbon cycle, have led some researchers to call for a new geological epoch to be defined - the Anthropocene.¹⁻³ The onset of the Industrial Revolution is considered to be the beginning of large-scale human impact, however, finding global and preserved geological records to define a new geological epoch is challenging and the subject of debate. The beginning of the Anthropocene is considered to have occurred around the time of the onset of industrialization in the 1800s, and this period extended through the “Great Acceleration” of population growth and industrialization during the second half of the 20th century.^{1, 4} For example, deposits produced by mining, waste disposal, construction, and urbanization (e.g., aluminium, plastics, concrete and synthetic fibers) have been observed in ice and sedimentary archives.^{2, 5} The presence of these “technofossils”, which are materials unique to mankind, reflect a stratigraphic footprint that is noticeably different from that seen in Holocene records.¹⁻³

In recent decades, the demand of synthetic chemicals has reached ~300 million tons per year due to the global use in the domestic, agricultural, and industrial sectors.^{6, 7} The increase in production rates and diversification of pharmaceuticals and pesticides has exceeded that of most recognized agents of global change (e.g., CO₂, world population, agricultural land and biodiversity).⁸ Chemicals released to the environment via direct sources (e.g., waste water treatment plant (WWTP) effluents) or indirect input sources (e.g., runoff of pesticides from agricultural land after application, the drift of sprayed pesticides, accidental spills or atmospheric deposition) can enter natural waterways and ultimately be incorporated into aquatic sediments. Any of these chemicals that persist over time would become excellent archives of past environmental contamination, and can be used to reconstruct environmental changes over time within a catchment. While many studies have analyzed the sedimentary

record of anthropogenic pollution,⁹⁻¹³ gaining a comprehensive picture of many pollutants simultaneously in local palaeorecords remains challenging.

Recently, advances in analytical techniques such as high resolution mass spectrometry (HRMS) have led to new opportunities in environmental trace analysis to identify anthropogenic contaminants and transformation products (TPs) in different environmental matrices.¹⁴ HRMS has permitted the detection of a wider range of organic contaminants, owing to its high resolving power (10,000 – 1,000,000), mass accuracy (<2-5 ppm) and sensitivity, which is in the femtogram to picogram range in full scan mode.¹⁴ The coupling of HRMS with liquid chromatography (LC) through electrospray ionization and MS/MS capabilities enables reliable screening for molecular ions and their MS/MS fragments against complex matrix backgrounds (e.g., soils, sludge, and sediments).¹⁴ Until now, mostly target (quantitative analysis with reference standards) and to less extend suspect screening (qualitative analysis without reference standards) approaches have been used as the primary routine methods to monitor organic contaminants, such as widely used pesticides or pharmaceuticals in surface, drinking and waste waters.^{15, 16} The analysis of anthropogenic chemicals in complex matrix backgrounds is challenging due to the matrix interference from natural organic matter (NOM), which contains macromolecules and considerable amounts of decomposed biomolecules. Moreover, the large amount of data obtained from HRMS measurements requires the use of intelligent mining and statistical tools to fully explore the capabilities of HRMS. To date, only a few studies have explored the ability of HRMS to analyze sediment and soil matrices.^{9, 10, 13, 17, 18} These studies demonstrate the capabilities of HRMS to screen hundreds of compounds in a single run without losing sensitivity and selectivity. For example, the two lakes considered here, Greifensee (Switzerland) and Lake Lugano (Switzerland and Italy), were previously analyzed for several hundred of anthropogenic organic compounds providing evidence of the suitability of sediments not only to study highly lipophilic compounds but also polar emerging contaminants, including pharmaceuticals and all registered organic synthetic pesticides in Switzerland.^{9, 10, 19} The occurrence of less lipophilic compounds in sediments and the use of less generic ionization techniques (e.g., electrospray ionization (ESI) vs. electron impact (EI)) for their detection has been also demonstrated in other studies as in the case of nonylphenols, surfactants and their degradation intermediates.²⁰ However, no studies so far have attempted to provide a comprehensive overview of the contamination patterns over time and systematically document the Anthropocene as regards to organic contaminants, and their evolution since.

Therefore, the objective of this work was to identify the plethora of contaminants in sedimentary records deposited during the Anthropocene through statistical analysis of temporal trends from LC-HRMS non-target (unknown compounds) screening data. The statistical tools developed aim to provide a history of the human impacts in two lakes of Central Europe and offer a relatively rapid screening technique applicable to a geographically widespread palaeo-archive (i.e. lake sediments). Furthermore, the statistical approach, evaluated with a range of synthetic chemicals previously identified at these sites, was used to emphasize and further investigate trends in the contaminants that exhibit the highest concentrations at the present time to identify organic contaminants not yet successfully targeted by environmental regulation and pollution reduction initiatives. The lakes studied have similar point and non-point sources of contamination. However, Lake Lugano has different hydraulic retention times, redox conditions, climate conditions, and different water management when compared to Greifensee. To our knowledge, this is the first application of non-target LC-HRMS combined with statistical analysis to lake sediments to reconstruct a local history of persistent environmental contamination and to provide a comprehensive analysis of many polar non-volatile anthropogenic contaminants simultaneously.

Methods

Standards and reagents

Details on the sources, preparation and storage of target reference standards and reagents are described elsewhere¹⁰ and provided as Supporting Information.

Sampling collection and preservation

A sediment core of approximately 1 m long was collected from the southern basin of Lake Lugano (45° 57' 31.5" N, 8° 53' 38.3" E) and Greifensee (47° 21' 07" N, 8° 40' 41" E), which are located in the southern and northern parts of Switzerland, respectively, during the summer of 2012 and the fall of 2014. Sediment cores were extracted using a free-fall gravity corer and stored vertically in the dark at 4°C until analysis. The sediment cores were opened and sectioned into intervals between 3 and 5 cm in thickness representing 5 to 10 years' time-frame using stainless steel slicers. Additional "blank" samples were taken from the pre-industrial sediments of the core, when anthropogenic contaminants are assumed to be absent. The samples were transferred to airtight glass jars, frozen for 24 h and freeze-dried for 48 h

(LYO GT2 Basic, SRK-Systemtechnik GmbH, Riedstadt, Germany). Full site description of the lakes is provided in Table S1 in SI and elsewhere.^{21, 22}

Dating and geochemical analysis

Sediment cores from both lakes were dated using ²¹⁰Pb and ¹³⁷Cs signals from Chernobyl (1986) and the atomic bomb tests (1963) using a high-purity Germanium (HPGe) Well Detector (Gamma spectroscopy, GCW 3523) as reported in SI and elsewhere.²³ Total carbon (TC) content was measured using an elemental analyser (EURO EA 3000). Total inorganic carbon (TIC) was obtained from a titration Coulometer (CM5015). Total organic carbon (TOC) was calculated as TOC = TC – TIC. TOC values for both lakes are reported in Table S2 in SI. Based on the TOC values in Greifensee and Lake Lugano the change of TOC within both lakes do not influence the outcomes of this work.

LC-HRMS analyses

Sediment samples were freeze-dried and extracted by pressurized liquid extraction at 80 °C using a mixture of two polar aprotic solvents ethyl acetate and acetone at a ratio of 70:30 (% v/v) as reported in SI and elsewhere.^{10, 24} Detection of analytes was performed on an LC system connected to a QExactive™ Hybrid Quadrupole-Orbitrap Mass Spectrometer (Thermo Fisher Scientific, San Jose, U.S.A.) equipped with an ESI source, as described by Chiaia-Hernández et al.⁹ Data dependent and data independent acquisition (DIA) measurements were performed separately in the positive and negative ionization modes. Full range mass spectra were recorded over a mass range of 100 to 1100 *m/z* with a nominal resolving power of 140,000 referenced to *m/z* 400 and with a mass accuracy of ± 5 ppm. High-resolution product ion spectra were acquired in MS/MS experiments with a nominal resolving power of 17,500. For additional confirmation of prioritized compounds, samples were re-injected and measured at higher energy collision dissociation (HCD) values of 15, 25, 55, 75 and 90% for targeted MS/MS fragmentation. Specific instrumental settings, DIA mass isolation windows, quality control and quantification are reported in the SI. The complete list of organic contaminants detected in Greifensee and Lake Lugano is reported in Table S3 in SI.

Extraction of profiles

The web interface of the R package enviMass 3.0 was used to detect temporal trends in the LC-HRMS data in both lakes by sorting of identical exact masses into time series.²⁵ Originally created to detect temporal trends and chemical spills in water, the appropriateness of enviMass to processes sediment samples was investigated. LC-HRMS data from individual sediment layers, as well as sediment blanks, were uploaded as mzXML files in two separate projects for the positive and negative modes. Instrument raw files were converted to mzXML files using the open-source software ProteoWizard (version 3.0.9016).²⁶ Mass recalibration, global intensity correction and RT alignment were performed with the aid of 98 internal standards reported in Tables S4 and S5 in the SI. The main adducts, $[M+H]^+$ in the positive mode and $[M-H]^-$ in the negative mode, were included in the screening analysis. Furthermore, masses were restricted to an RT window from 3.8 min to 23.5 min to account for the elution of contaminants according to the settings chosen in the LC. At this stage, the profiles obtained were not further grouped with their isotopes and adducts to avoid loss of information. However, based on our experience, isotopes and adducts account for only ~20% of the total extracted masses. Detailed parameters for the workflow and settings are given in Table S6 in the SI and elsewhere.^{25, 27}

Spearman's rank correlation

Spearman's rank correlation was applied as a method for distinguishing, from thousands of time series, profiles with monotonically increasing intensities over time. The complete data set was extracted from enviMass and intensity vs. time was analyzed. The calculated coefficient (ρ) values ranged from -1 for decreasing monotonic trends to 1 for increasing monotonic trends. Calculations were performed with the open-source software R with the *cor* function using the "spearman" method.²⁸

Hierarchical cluster analysis (HCA)

Agglomerative HCA was chosen as an approach, due to its central role in classification and information retrieval.²⁹ Three main phases were optimized in Greifensee Lake sediments and later applied to sediments of Lake Lugano: data normalization, measurement of similarity and clustering. Normalization was tested by dividing by the maximum, mean and sum, where normalizing by the sum resulted in the best clusters. The similarity between the normalized description vectors was calculated using the widely used Euclidean distance, which gave better results than the Manhattan and Gower distances. Furthermore, Ward's minimum

variance method, which is based on variance analysis, was used for the clustering of data. The generation of clusters from the hierarchical tree was performed in different batches, based on visual criteria. The minimum number of clusters was found by capturing all the trends with increasing concentrations at one point in time and without losing information since this was the main goal of this work. All statistical tests were performed with the open-source software R using the functions *daisy* and *hclust* from the “*cluster*” package.^{28, 30} Heatmaps of dendograms of the hierarchical cluster analysis for both lakes are illustrated in Figure S3.

Structural elucidation of a non-target compound on Greifensee sediments

In silico fragmentation was performed for the identification of non-targets using the full scan and MS/MS data at a collision energy (HCD) of 90% due to the higher retrieval of information when compared to other collision energies. This step was executed using the R version of MetFrag 2.3 with the databases ChemSpider, PubMed and Mass Bank North America (MoNA).³¹ Possible structures were further evaluated using similar methods described in previous work.⁹ These methods included detailed inspection of peak shapes, isotopic patterns, correlation of RT with log K_{ow} values, the number of extracted references, and matching to spectral libraries using mzCloud and MassBank.^{10, 32, 33} The identification of the compound was ultimately confirmed with the purchase of reference standards. Estimated concentrations of the identified non-target were determined by comparing the ratios between standard and internal standard from a 9 point calibration curve prepared in methanol with the ratios obtained in the sediment extracts in different runs.

Results and Discussion

Greifensee Lake sediments

Construction of organic chemical time-series using LC-HRMS

The software enviMass was originally created to detect temporal chemical profiles and chemical spills in water. Its appropriateness for detecting temporal chemical profiles in a more complex matrix was explored for the first time in this work by detecting more than >71,000 compound profiles in positive ionization mode and >148,000 in negative ionization mode. The results were validated, to ensure the accuracy of the algorithms, by comparing the results from 16 target substances previously reported in Greifensee sediments which include personal care products, pesticides, biocides, additives, corrosion inhibitors, musk fragrances,

and industrial chemicals entering the lake via direct and diffuse point sources (Table S3).^{9, 10} Profiles of two pesticides, irgarol and dichlorophen, obtained in this study are superimposed on previously published data in Figure 1.^{9, 10} enviMass was successful in retrieving all the 16 substances by reflecting good performance and overlap with reported concentration patterns. To differentiate between anthropogenic origin and non-anthropogenic compounds (i.e. NOM), and to eliminate instrumental noise (i.e., from the source, the input transducer, signal processing elements and output transducer), the compound time-series were filtered using Spearman's rank correlation coefficient (ρ) to select compounds whose concentration increases over time—an indication that they are anthropogenically sourced through (i) the increased use of chemicals, (ii) their formation from other chemicals with increasing concentrations or (ii) due to increasing degradation of chemicals with time and accumulation in surface sediments as have been shown in other studies.^{9-11, 20, 34, 35} The distribution of Spearman's ρ values is illustrated in Figure 2. Only 11% and 12% of the profiles correspond to time series with increasing concentrations over time ($\rho \geq 0.5$) in the positive and negative ionization modes, respectively. This suggests that a significant percentage of profiles (approx. 88%) can be attributed to instrumental noise, or non-anthropogenic compounds (NOM or compounds that decrease in intensity over time). Therefore, this data filtering is an important step in isolating, from thousands of possible profiles, those that are anthropogenically sourced, and therefore most relevant for the analysis.

Profiles with steadily decreasing intensities over time can be attributed to compounds that experience higher ion suppression in younger sediment layers due to a stronger matrix effect. As the main focus of this study was to determine signals of anthropogenic contaminants that can be of potential concern due to their constant increase in concentration, profiles with constant decreasing intensities for the past 100 years, although potentially important components for the characterization of the Anthropocene, were excluded from further analysis. All of the 16 previously published anthropogenic compounds with increasing concentrations over time observed in Greifensee had a Spearman's rank correlation $\rho \geq 0.5$. Therefore, this was chosen as the cut-off value to pre-filter the data. Moreover, the selection of the Spearman's rank correlation cutoff was evaluated by randomly selecting 200 profiles in R²⁸ and manually assign if they should belong to the Anthropocene or not, due to the increase of the compound at one point in time. Afterwards the labels were compared with the spearman cutoff. A threshold of $\rho > 0.5$ leads to a sensitivity of 95% and specificity of 97% as derived from the confusion matrix as illustrated in Table S7 in SI. The selected threshold ($\rho >$

0.5) worked well for both lakes, however, this cutoff should be evaluated for other studies in different matrices.

Grouping time-series by hierarchical cluster analysis (HCA)

HCA analysis was used on pre-filtered data ($\rho \geq 0.5$) to further cluster profiles into groups with similar trends within a sediment core. Alternatively, HCA was also applied to un-filtered data (no Spearman's rank correlation) for comparative purpose as discussed later.

For the pre-filtered data in positive ionization mode, 29 clusters were chosen as the minimal number of clusters for a clear separation of trends. Seven of these clusters reflect increasing concentrations over time in recent years. 9 of these clusters peak at one point in time before concentrations decline towards 2014, and 13 of these clusters were attributed to NOM or instrumental noise due to the inconsistencies of the profiles (Figure S2). In negative ionization mode, 19 clusters were selected. Three of these clusters reflect increasing concentrations over time, 9 of these clusters have a maximum at one point in time, and 7 of these clusters are attributed to NOM (Figure 3 and S2). Previously published compounds from Greifensee, including dichlorophen, irgarol, and the antibacterial agent triclocarban, fit into these clusters (displayed in blue within their respective clusters in Figure 3, together with other anthropogenic profiles in each cluster). Based on HCA, 46% (4,000 profiles) and 43% (6,000 profiles) of all detected temporal trends with $\rho \geq 0.5$ in both the positive and negative ionization modes, respectively, were eliminated from further analysis, as they display profiles attributed to NOM and instrumental noise. Profiles with a defined maximum over time (e.g., the compound triclocarban and dichlorophen in Figure 3a and 3c) account for 30% (2,500 profiles) in the positive ionization and 36% (5,300 profiles) in the negative ionization mode, while 24% (2,000 profiles) and 21% (3,000 profiles) of the total trends observed with $\rho \geq 0.5$ correspond to profiles with increasing concentrations up to the present day in the positive and negative ionization modes, respectively. Therefore, between 54% and 57% of all the filtered profiles were recognized as compounds having anthropogenic origins or to NOM with constant increasing intensity values. However, it is unlikely that NOM-sourced compounds display increasing intensities over time. According to Peng et al.,¹⁷ only organo-bromine compounds (NSOBCs) with increasing abundances over time provide evidences of anthropogenic input and about half of the NSOBCs detected in sediment cores of Lake

Michigan (~800 compounds) were from natural sources, which include NSOBCs with decreasing abundances over time.

The cluster analysis was also tested on unfiltered data for purposes of comparison, i.e. using data not filtered according to a Spearman's rank correlation coefficient. Only compounds with a signal intensity $> 10^6$ were used due to the computational limitations of what would otherwise be a very large dataset (e.g., up to a gigabyte per environmental sample). The data show similar clusters as those identified in the pre-filtered data, but with additional clusters having constant or a single peak in the profiles and with decreasing concentrations over time. Compounds whose signal intensities vary little throughout the core, or consist of a single peak unrelated to chemical spills or events, are unlikely to be of anthropogenic origin. While this approach can identify clusters of organic contaminant time-series, it underrepresents or neglects equally relevant compounds that have either small ionization efficiencies, experience ion suppression due to a stronger matrix effect, are present at low concentrations or come from missing spill events. As a result, cluster analysis on pre-filtered data provides a more comprehensive screening of non-target compounds from anthropogenic origin as well as systematically reduce data to overcome computational limitations.

Organic contaminants in the Anthropocene

Based on the statistical analysis, Greifensee contained few or no organic contaminants before the ~1950s (Figure 3). The results are consistent with the sudden acceleration of human enterprise after the end of the Second World War, i.e., the rise in petroleum consumption by a factor of 3.5 since 1960, which marks the increase in human activity globally,^{1, 36} the 14% increase in pesticide use since the mid-1950s,³⁶ the rise in the production of synthetic chemicals,^{6, 7} and the increase in gross domestic product (GDP) on a global scale.³⁷ In Switzerland, imports raised from 273 CHF million in 1950 to an average of 6178 CHF million from 1950 until 2017 driven mainly by chemicals and pharmaceuticals, vehicles and energy products.³⁸ The cluster analysis highlights the dissipation of many profiles around the 1970s and 1980s (Figure 3a and 3c). Clusters with maxima occurring in the 1970s and 1980s account for 34% of all profiles of anthropogenic origin as reported for both ionization modes combined, encompassing more than 7,770 profiles. The decline in concentration of these profiles since the 1980s is consistent with the construction of 7 WWTPs between 1956 and 1971, and their subsequent upgrade in the early 1980s, which reduced the amount of compounds in the effluents reaching the lake. Implementation of environmental regulations,

policies, and management practices in Switzerland since the 1980s also coincide with the decline in anthropogenic compounds in Greifensee (Figure 3 a and c). All these measures to improve water quality were extremely effective in removing potential organic contaminants reaching the lake. The decrease in concentrations and similar patterns observed in Figure 3a have been also reported for highly hydrophobic compounds like polychlorinated dibenzo-p-dioxins and furans (PCDD/F) and PCBs in Greifensee with an increase in concentrations starting in the earlier 1940s and a decreased in concentration starting in the 1960s as reported by Zennegg et al.¹⁹ However, increasing concentrations observed in some of the clusters, such as irgarol (Figure 3b), that replaced antifouling paints containing organotin compounds after their ban in Switzerland in the 1990s, show that anthropogenic contaminants are still reaching the lake. These compounds might be released from non-point sources (e.g., paint leaching from boats for irgarol), WWTPs effluents or transform through different processes in the water phase or in the sediment. In Figure 3d a cluster is depicted with increasing intensities over time that contains more than 400 unknown profiles.

The overall increase in the number of profiles detected (5,000) in both ionization modes is consistent with the steady increase in organic contaminants reported elsewhere in the world; for example, in sediment profiles from a sewage-impacted estuary in Jamaica Bay, NY, where the sediment profiles of pharmaceuticals correlate with the increase in prescription drug usage in the United States.¹¹ Furthermore, sediments from Lake Erie and Lake Ontario document the increased production of synthetic musk fragrances in the United States.³⁹ However, these trends have not yet been demonstrated for the full class of polar organic anthropogenic compounds.

The cluster analysis presented here provides an effective method for characterizing thousands of synthetic chemicals emitted throughout the Anthropocene and provides a more holistic overview of contamination than known target compound analysis. In addition, it is an important step in prioritizing the identification of organic contaminants with increasing concentrations over time reaching the lake through different input sources as well as delivers further understanding in the management and protection of aquatic environments.

Non-target identification

The developed method was used to identify the most relevant contaminants present in Greifensee today that are absent in monitoring studies. Data-dependent MS/MS acquisition was performed on 460 prioritized masses obtained from clusters with steadily increasing

concentrations over time in the positive ionization mode and having intensities $>10^6$ or with very characteristic isotopic pattern (e.g., Cl and Br). The data were evaluated using two approaches: the environmental and food safety compounds high-resolution, accurate-mass MS/MS spectral library (EFS, Thermo Scientific, >1700 organic contaminants) was used to match spectra, and *in silico* fragmentation using MetFrag 2.3 was performed.⁴⁰

Both approaches resulted for one predominant profile in the tentative identification of the fungicide imazalil, based on retention time (RT) in reversed phase chromatography which is correlated to the hydrophobicity value ($\log K_{ow}$), and the number of explained MS/MS peaks (> 5) at a collision energy (HCD) of $\geq 35\%$ (according to spectra matched using the EFS). In addition, an overall highest ranking score and a total of 12 explained peaks with a HCD of 90% in MetFrag was reached. The measured spectra that were suspected to represent imazalil were further compared with an available mass spectrum in the library m/z cloud (No. 3110)³³ and a reference standard as illustrated in Figure 4. Estimated concentrations of imazalil (C_{ised}) are reported to be between 200 to 400 pg/g_{dw} (Figure 4c). These values are in the range of other organic contaminants (e.g., personal care products and pesticides) previously found in Greifensee sediments.^{9, 10}

There are a number of possible sources of this contaminant in Greifensee. Imazalil is a registered fungicide in Switzerland used in the treatment of potato seeds and the spraying of vegetables in greenhouses. In Switzerland, imazalil is used at low quantities of < 50 kg/year in agriculture, making this route unlikely. Imazalil is also registered as a veterinary product (it is administered topically) under the name enilconazole.⁴¹ Despite its limited use, Kahle et al.⁴¹ reported imazalil concentrations of up to 26 ng/L and 10 ng/L in Swiss WWTP influents and effluents, respectively, suggesting its veterinary use as a possible pathway for imazalil to enter surface waters. Imazalil has been also reported to be present in wastewater effluents and sediments in other parts in Europe such as Spain.⁴² The fact that imazalil is also used as a post-harvest fungicide for preventing moisture loss during the storage, shipment and marketing of citrus fruits was overlooked until now.⁴³ Ortelli et al.⁴³ reported imazalil as the top post-harvest fungicide (70% detection rate) of 240 citrus fruits (lemons, oranges, mandarin oranges, grapefruit, limes, pomelos and kumquats) from the Swiss market in Geneva with concentrations of up to 2.5 mg/kg. Imazalil residues on citrus fruits may enter surface waters through the washing of fruits or hands after contact with the fruit. This contamination pathway offers the most probable source of imazalil in Greifensee via imported citrus fruits which had, so far, largely been neglected.

The clustering analysis presented here, combined with non-target screening, is a new tool for identifying such contaminants with unexpected sources and input pathways to surface waters.

Lake Lugano sediments

HCA of Lake Lugano sediments

In order to validate the method developed for Greifensee as well as generalise the obtained results with regards to synthetic chemicals in the Anthropocene, the same approach was applied on data obtained from the southern basin of Lake Lugano, on the border between Switzerland and Italy.

The cluster analysis indicates little or no organic contamination before the 1950s. Lake Lugano sediments record similar patterns to Greifensee, but with some clusters delayed by > 10 years. The occurrence of contaminants known to be present in Lake Lugano and reported elsewhere⁹ were again used to evaluate the efficiency and reproducibility of the developed method. For example, triclocarban and its transformation product dichlorocarbanilide (DCC), as shown in Figure 5a, show a peak in the early 1990s, 10 years later than in Greifensee (Figure 3c). 23% (9,000 profiles) of the detected anthropogenic compounds in Lake Lugano peak at one point in time before declining in recent years, while 52% (20,000 profiles for both ionization modes) of all detected anthropogenic compounds increase over time towards the present day. Therefore, 29,000 of all these profiles are recognized as profiles having possible anthropogenic origins – more than twice as many as in Greifensee (~13,000 anthropogenic profiles). These include pharmaceuticals such as the analgesic flufenamic acid (Figure 5b). Pharmaceuticals were not detected in Greifensee and have only been reported in literature in sediments impacted strongly by sewage treatment plant discharges.^{11, 44} In Lake Lugano, the concentration of the antibacterial agent triclosan began to decrease in the late 1990s (Figure 5b), whereas triclosan has been eliminated from Greifensee already since the 1970s.¹⁰ This is in accordance with the construction of WWTPs in the Swiss territory of the lake since 1976 and further upgrade treatment in 1995.

Two clusters with increasing concentrations encompassed more than 1000 profiles, including irgarol and the isomeric pesticides terbutryn and prometryn (Figure 5c and 5d). As in Greifensee, concentrations of irgarol in Lake Lugano continue to increase towards the present day. These results are consistent with the high degree of contamination of Lake Lugano in the

1960s and 1970s, when swimming in the lake was officially forbidden. Despite the introduction of wastewater treatment since the 1970s, the high population density and the strong development of tourism in this relatively small catchment have led to substantial contamination of the lake despite the considerable efforts in water management. For example, today, only ~20% of the population in the Italian part of the catchment is connected to a WWTP with phosphorus removal. These observations are consistent with the large number of anthropogenic signals observed in the lake sediments in this study.

Insights and Future Research

HRMS combined with statistical analysis proved to be extremely valuable to gain a comprehensive picture of the occurrence of organic contaminants and their trends over time in lake sediments. Environmental changes recorded in sediments of Greifensee and Lake Lugano give evidence of past and present water quality, as well as a record of management and mitigation measures. Sediments from the two lakes record signatures that clearly define the beginning of large-scale human impact around the 1950s in accordance with the Great Acceleration. The results are also in agreement with other geological deposits (e.g. nutrient pollution and radioisotope fallout signals) and clearly ratify a long-lived phenomenon as was shown for thousands of profiles with steady increasing trends until today and not yet successfully targeted by environmental regulation and pollution reduction initiative. This includes chemicals with unrecognized input pathways, as shown here for the fungicide imazalil as a first example of application of this new method. Moreover, the results highlight the need of databases and mass libraries to identify contaminants not recognized until now, as have been shown for thousands of profiles, as well as to explore other ionization techniques to capture other persistent organic contaminants.

The HRMS method in combination with the statistical tools presented in this work offers a rapid screening technique applicable to other palaeoarchives and environmental matrices to document the time and rate of changes in contamination within a catchment and provides additional signals of a global environmental problem. In addition, more research is needed to confirm the ubiquitous presence of human impact in the environment by providing additional signals on the Anthropocene that are global and preserved in the geological record as has been presented here for two lakes in Central Europe.

Acknowledgements

The authors thank Adrian Müller and Andreas Maccagnan for their work in the laboratory; Heinz Singer for his help with non-target analysis; Emma L. Schymanski for her help with MetFrag; Martin Loos for his help with enviMass; the Sedimentology Group at Eawag for their help in the collection of cores and dating of the sediments, Nathalie Dubois and Pablo Lara-Martin for their feedback and helpful discussions, and Cresten Mansfeldt for comments and suggestions

Supporting Information

Analytical methods, site description, chemical analysis, and profile extraction methods are provided as additional information.

References

1. Steffen, W.; Grinevald, J.; Crutzen, P.; McNeill, J., The Anthropocene: conceptual and historical perspectives. *Philos Trans A Math Phys Eng Sci* **2011**, *369*, (1938), 842-867.
2. Waters, C. N.; Zalasiewicz, J.; Summerhayes, C.; Barnosky, A. D.; Poirier, C.; Gałuszka, A.; Cearreta, A.; Edgeworth, M.; Ellis, E. C.; Ellis, M.; Jeandel, C.; Leinfelder, R.; McNeill, J. R.; Richter, D. d.; Steffen, W.; Syvitski, J.; Vidas, D.; Waple, M.; Williams, M.; Zhisheng, A.; Grinevald, J.; Odada, E.; Oreskes, N.; Wolfe, A. P., The Anthropocene is functionally and stratigraphically distinct from the Holocene. *Science* **2016**, *351*, (6269).
3. Zalasiewicz*, J.; Williams, M.; Steffen, W.; Crutzen, P., The New World of the Anthropocene. *Environ. Sci. Technol.* **2010**, *44*, (7), 2228-2231.
4. Steffen, W.; Crutzen, P. J.; McNeill, J. R., The Anthropocene: Are Humans Now Overwhelming the Great Forces of Nature. *Ambio* **2007**, *36*, (8), 614-621.
5. Ford, J. R.; Price, S. J.; Cooper, A. H.; Waters, C. N., An assessment of lithostratigraphy for anthropogenic deposits. *Geological Society, London, Special Publications* **2014**, *395*, (1), 55-89.
6. Schwarzenbach, R. P.; Escher, B. I.; Fenner, K.; Hofstetter, T. B.; Johnson, C. A.; von Gunten, U.; Wehrli, B., The Challenge of Micropollutants in Aquatic Systems. *Science* **2006**, *313*, (5790), 1072-1077.
7. FAO, Statistical Database, <http://faostat.fao.org/> (Food and Agriculture Organization of the United Nations, Rome 2006).
8. Bernhardt, E. S.; Rosi, E. J.; Gessner, M. O., Synthetic chemicals as agents of global change. *Front. Ecol. Evol.* **2017**, *15*, (2), 84-90.
9. Chiaia-Hernandez, A.; Schymanski, E.; Kumar, P.; Singer, H.; Hollender, J., Suspect and nontarget screening approaches to identify organic contaminant records in lake sediments. *Anal. Bioanal. Chem.* **2014**, *406*, (28), 7323-7335.
10. Chiaia-Hernandez, A. C.; Krauss, M.; Hollender, J., Screening of Lake Sediments for Emerging Contaminants by Liquid Chromatography Atmospheric Pressure Photoionization and Electrospray Ionization Coupled to High Resolution Mass Spectrometry. *Environ. Sci. Technol.* **2013**, *47*, (2), 976-986.

497 11. Lara-Martín, P. A.; Renfro, A. A.; Cochran, J. K.; Brownawell, B. J., Geochronologies of
 498 Pharmaceuticals in a Sewage-Impacted Estuarine Urban Setting (Jamaica Bay, New York). *Environ. Sci.*
 499 *Technol.* **2015**, *49*, (10), 5948-5955.

500 12. Li, Z.; Maier, M. P.; Radke, M., Screening for pharmaceutical transformation products formed
 501 in river sediment by combining ultrahigh performance liquid chromatography/high resolution mass
 502 spectrometry with a rapid data-processing method. *Anal. Chim. Acta* **2014**, *810*, 61-70.

503 13. Peng, H.; Chen, C.; Saunders, D. M. V.; Sun, J.; Tang, S.; Codling, G.; Hecker, M.; Wiseman, S.;
 504 Jones, P. D.; Li, A.; Rockne, K. J.; Giesy, J. P., Untargeted Identification of Organo-Bromine
 505 Compounds in Lake Sediments by Ultrahigh-Resolution Mass Spectrometry with the Data-
 506 Independent Precursor Isolation and Characteristic Fragment Method. *Anal. Chem.* **2015**, *87*, (20),
 507 10237-10246.

508 14. Krauss, M.; Singer, H.; Hollender, J., LC–high resolution MS in environmental analysis: from
 509 target screening to the identification of unknowns. *Anal. Bioanal. Chem.* **2010**, *397*, (3), 943-951.

510 15. Moschet, C.; Wittmer, I.; Simovic, J.; Junghans, M.; Piazzoli, A.; Singer, H.; Stamm, C.; Leu, C.;
 511 Hollender, J., How a Complete Pesticide Screening Changes the Assessment of Surface Water Quality.
 512 *Environ. Sci. Technol.* **2014**, *48*, (10), 5423-5432.

513 16. Schymanski, E.; Singer, H.; Slobodnik, J.; Ipolyi, I.; Oswald, P.; Krauss, M.; Schulze, T.; Haglund,
 514 P.; Letzel, T.; Grosse, S.; Thomaidis, N.; Bletsou, A.; Zwiener, C.; Ibáñez, M.; Portolés, T.; de Boer, R.;
 515 Reid, M.; Onghena, M.; Kunkel, U.; Schulz, W.; Guillon, A.; Noyon, N.; Leroy, G.; Bados, P.; Bogialli, S.;
 516 Stipanicev, D.; Rostkowski, P.; Hollender, J., Non-target screening with high-resolution mass
 517 spectrometry: critical review using a collaborative trial on water analysis. *Anal. Bioanal. Chem.* **2015**,
 518 *407*, (21), 6237-6255.

519 17. Peng, H.; Chen, C.; Cantin, J.; Saunders, D. M. V.; Sun, J.; Tang, S.; Codling, G.; Hecker, M.;
 520 Wiseman, S.; Jones, P. D.; Li, A.; Rockne, K. J.; Sturchio, N. C.; Giesy, J. P., Untargeted Screening and
 521 Distribution of Organo-Bromine Compounds in Sediments of Lake Michigan. *Environ. Sci. Technol.*
 522 **2016**, *50*, (1), 321-330.

523 18. Chiaia-Hernandez, A. C.; Keller, A.; Wächter, D.; Steinlin, C.; Camenzuli, L.; Hollender, J.;
 524 Krauss, M., Long-Term Persistence of Pesticides and TPs in Archived Agricultural Soil Samples and
 525 Comparison with Pesticide Application. *Environ. Sci. Technol.* **2017**, *51*, (18), 10642-10651.

526 19. Zennegg, M.; Kohler, M.; Hartmann, P. C.; Sturm, M.; Gujer, E.; Schmid, P.; Gerecke, A. C.;
 527 Heeb, N. V.; Kohler, H.-P. E.; Giger, W., The historical record of PCB and PCDD/F deposition at
 528 Greifensee, a lake of the Swiss plateau, between 1848 and 1999. *Chemosphere* **2007**, *67*, (9), 1754-
 529 1761.

530 20. Lara-Martín, P. A.; Petrovic, M.; Gómez-Parra, A.; Barceló, D.; González-Mazo, E., Presence of
 531 surfactants and their degradation intermediates in sediment cores and grabs from the Cadiz Bay
 532 area. *Environ. Pollut.* **2006**, *144*, (2), 483-491.

533 21. Barbieri, A.; Simona, M., Trophic evolution of Lake Lugano related to external load reduction:
 534 Changes in phosphorus and nitrogen as well as oxygen balance and biological parameters. *Lakes*
 535 *Reserv. Res. Manag.* **2001**, *6*, (1), 37-47.

536 22. Singer, H.; Müller, S.; Tixier, C.; Pillonel, L., Triclosan: Occurrence and Fate of a Widely Used
 537 Biocide in the Aquatic Environment: Field Measurements in Wastewater Treatment Plants, Surface
 538 Waters, and Lake Sediments. *Environ. Sci. Technol.* **2002**, *36*, (23), 4998-5004.

539 23. Pennington, W.; Tutin, T. G.; Cambray, R. S.; Fisher, E. M., Observations on Lake Sediments
 540 using Fallout ¹³⁷Cs as a Tracer. *Nature* **1973**, *242*, (5396), 324-326.

541 24. Pintado-Herrera, M. G.; González-Mazo, E.; Lara-Martín, P. A., In-cell clean-up pressurized
 542 liquid extraction and gas chromatography–tandem mass spectrometry determination of hydrophobic
 543 persistent and emerging organic pollutants in coastal sediments. *J. Chromatogr. A* **2016**, *1429*, 107-
 544 118.

545 25. Loos, M.; Ruff, M.; Singer, H. enviMass version 3.0 target sceening software.
 546 <https://github.com/blosloos/enviMass> (01.10.2015),

547 26. Chambers, M. C.; Maclean, B.; Burke, R.; Amodei, D.; Ruderman, D. L.; Neumann, S.; Gatto, L.;
548 Fischer, B.; Pratt, B.; Egertson, J.; Hoff, K.; Kessner, D.; Tasman, N.; Shulman, N.; Frewen, B.; Baker, T.
549 A.; Brusniak, M.-Y.; Paulse, C.; Creasy, D.; Flashner, L.; Kani, K.; Moulding, C.; Seymour, S. L.;
550 Nuwaysir, L. M.; Lefebvre, B.; Kuhlmann, F.; Roark, J.; Rainer, P.; Detlev, S.; Hemenway, T.; Huhmer,
551 A.; Langridge, J.; Connolly, B.; Chadick, T.; Holly, K.; Eckels, J.; Deutsch, E. W.; Moritz, R. L.; Katz, J. E.;
552 Agus, D. B.; MacCoss, M.; Tabb, D. L.; Mallick, P., A cross-platform toolkit for mass spectrometry and
553 proteomics. *Nat Biotech* **2012**, *30*, (10), 918-920.

554 27. Ruff, M.; Mueller, M. S.; Loos, M.; Singer, H. P., Quantitative target and systematic non-target
555 analysis of polar organic micro-pollutants along the river Rhine using high-resolution mass-
556 spectrometry – Identification of unknown sources and compounds. *Water Res.* **2015**, *87*, 145-154.

557 28. *R Development Core Team (2008). R: A language and environment for statistical computing. R*
558 *Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL [http://www.R-](http://www.R-project.org)*
559 *[project.org](http://www.R-project.org).*

560 29. Murtagh, F.; Contreras, P., Algorithms for hierarchical clustering: an overview. *Wiley*
561 *Interdisciplinary Reviews: Data Mining and Knowledge Discovery* **2012**, *2*, (1), 86-97.

562 30. Maechler, M.; Rousseeuw, P.; Struyf, A.; Hubert, M.; Hornik, K. *cluster: Cluster Analysis Basics*
563 *and Extensions. R package version 2.0.5. 2016.*

564 31. Ruttkies, C.; Schymanski, E. L.; Wolf, S.; Hollender, J.; Neumann, S., MetFrag relaunched:
565 incorporating strategies beyond in silico fragmentation. *Journal of Cheminformatics* **2016**, *8*, (1), 3.

566 32. Horai, H.; Arita, M.; Kanaya, S.; Nihei, Y.; Ikeda, T.; Suwa, K.; Ojima, Y.; Tanaka, K.; Tanaka, S.;
567 Aoshima, K.; Oda, Y.; Kakazu, Y.; Kusano, M.; Tohge, T.; Matsuda, F.; Sawada, Y.; Hirai, M. Y.;
568 Nakanishi, H.; Ikeda, K.; Akimoto, N.; Maoka, T.; Takahashi, H.; Ara, T.; Sakurai, N.; Suzuki, H.; Shibata,
569 D.; Neumann, S.; Iida, T.; Funatsu, K.; Matsuura, F.; Soga, T.; Taguchi, R.; Saito, K.; Nishioka, T.,
570 MassBank: A public repository for sharing mass spectral data for life sciences. *J. Mass Spectrom.*
571 **2010**, *45*, (7), 703-714.

572 33. HighChem, (2016), mzCloud <https://mzcloud.org>, HighChem Ltd., Bratislava, Slovakia.
573 Accessed 17.08.2016.

574 34. Ferguson, P. L.; Bopp, R. F.; Chillrud, S. N.; Aller, R. C.; Brownawell, B. J., Biogeochemistry of
575 Nonylphenol Ethoxylates in Urban Estuarine Sediments. *Environ. Sci. Technol.* **2003**, *37*, (16), 3499-
576 3506.

577 35. León, V. M.; González-Mazo, E.; Pajares, J. M. F.; Gómez-Parra, A., Vertical distribution
578 profiles of linear alkylbenzene sulfonates and their long-chain intermediate degradation products in
579 coastal marine sediments. *Environ. Toxicol. Chem.* **2001**, *20*, (10), 2171-2178.

580 36. Agrios, G. N., chapter one - INTRODUCTION. In *Plant Pathology (Fifth Edition)*, Academic
581 Press: San Diego, 2005; pp 3-75.

582 37. Steffen, W.; Broadgate, W.; Deutsch, L.; Gaffney, O.; Ludwig, C., The trajectory of the
583 Anthropocene: The Great Acceleration. *The Anthropocene Review* **2015**, *2*, (1), 81-98.

584 38. Trading Economics. Switzerland imports.
585 <http://www.tradingeconomics.com/switzerland/imports>. Accessed on 31.03.2017.

586 39. Peck, A. M.; Linebaugh, E. K.; Hornbuckle, K. C., Synthetic Musk Fragrances in Lake Erie and
587 Lake Ontario Sediment Cores. *Environ. Sci. Technol.* **2006**, *40*, (18), 5629-5635.

588 40. Ruttkies, C.; Schymanski, E. L.; Wolf, S.; Hollender, J.; Neumann, S., MetFrag relaunched:
589 incorporating strategies beyond in silico fragmentation. *J. Cheminform.* **2016**, *in press*.

590 41. Kahle, M.; Buerge, I. J.; Hauser, A.; Müller, M. D.; Poiger, T., Azole Fungicides: Occurrence and
591 Fate in Wastewater and Surface Waters. *Environ. Sci. Technol.* **2008**, *42*, (19), 7193-7200.

592 42. Masiá, A.; Campo, J.; Vázquez-Roig, P.; Blasco, C.; Picó, Y., Screening of currently used
593 pesticides in water, sediments and biota of the Guadalquivir River Basin (Spain). *J. Hazard. Mater.*
594 **2013**, *263*, 95-104.

595 43. Ortelli, D.; Edder, P.; Corvi, C., Pesticide residues survey in citrus fruits. *Food Addit. Contam.*
596 **2005**, *22*, (5), 423-428.

44. Lahti, M.; Oikari, A., Vertical distribution of pharmaceuticals in lake sediments—citalopram as potential chemomarker. *Environ. Toxicol. Chem.* **2012**, *31*, (8), 1738-1744.

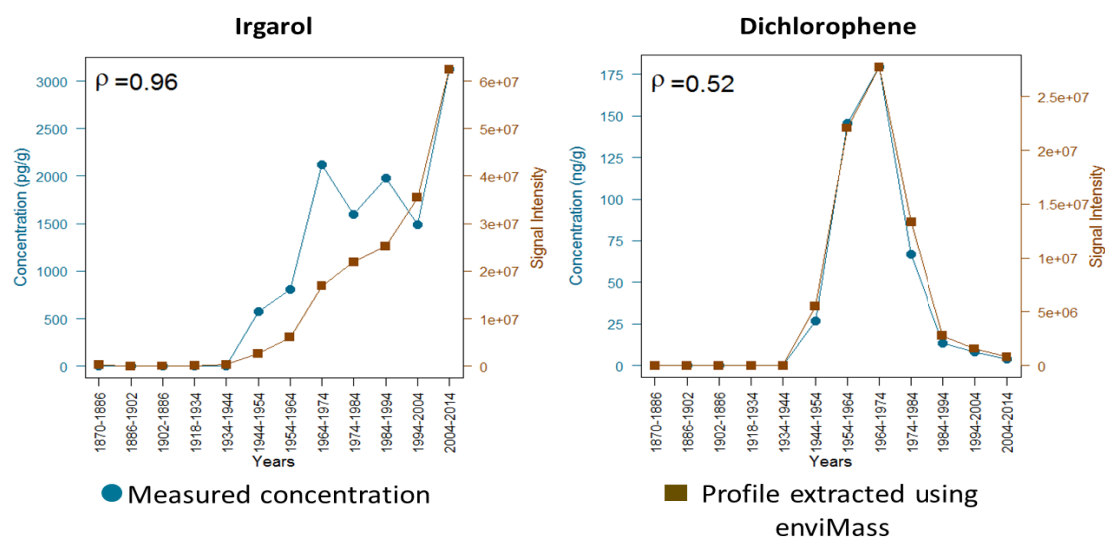


Figure 1. Interpolated profiles based on intensities from enviMass output data (this study, brown) and concentration profiles reported in previous studies (blue) for the pesticides irgarol (left) and dichlorophen (right) in Greifensee.^{9, 10} Spearman's rank correlation coefficient (ρ) values are shown on the top left of each panel.

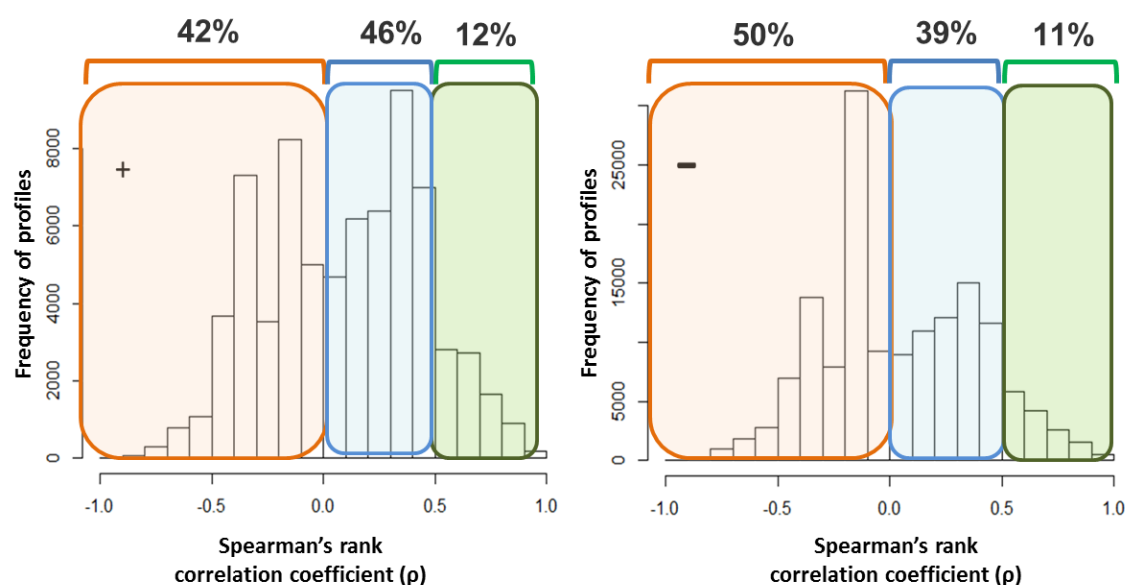


Figure 2. Distribution of Spearman's rank correlation coefficient (ρ) values for all Greifensee sediment compound time-series in the positive (left) and negative (right) ionization modes,

respectively. The profiles were retrieved from envMass. $p < 0.0$ shown in orange indicates decreasing or constant concentrations; $p \geq 0.5$ shown in green indicates increasing concentrations; values in between shown in blue indicate no clear trend.

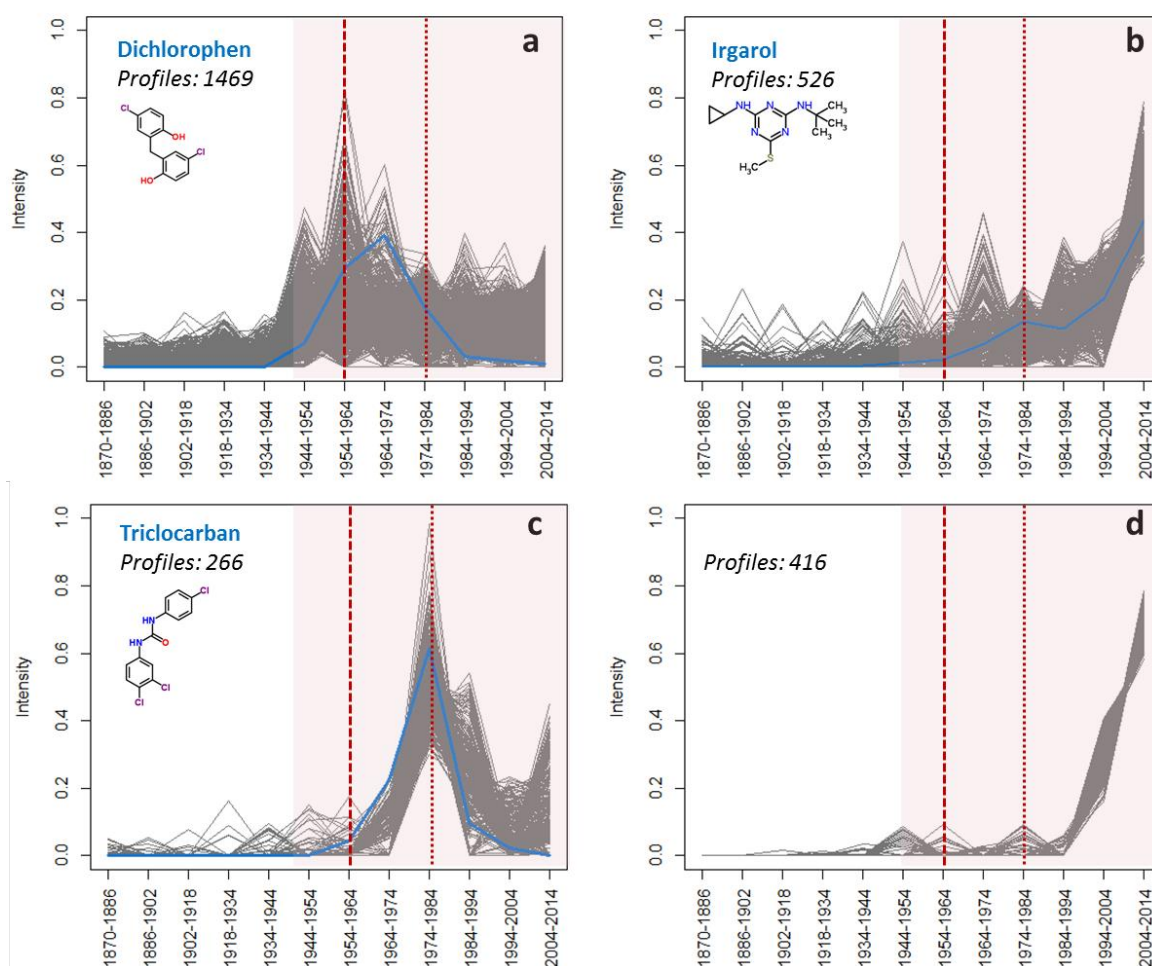
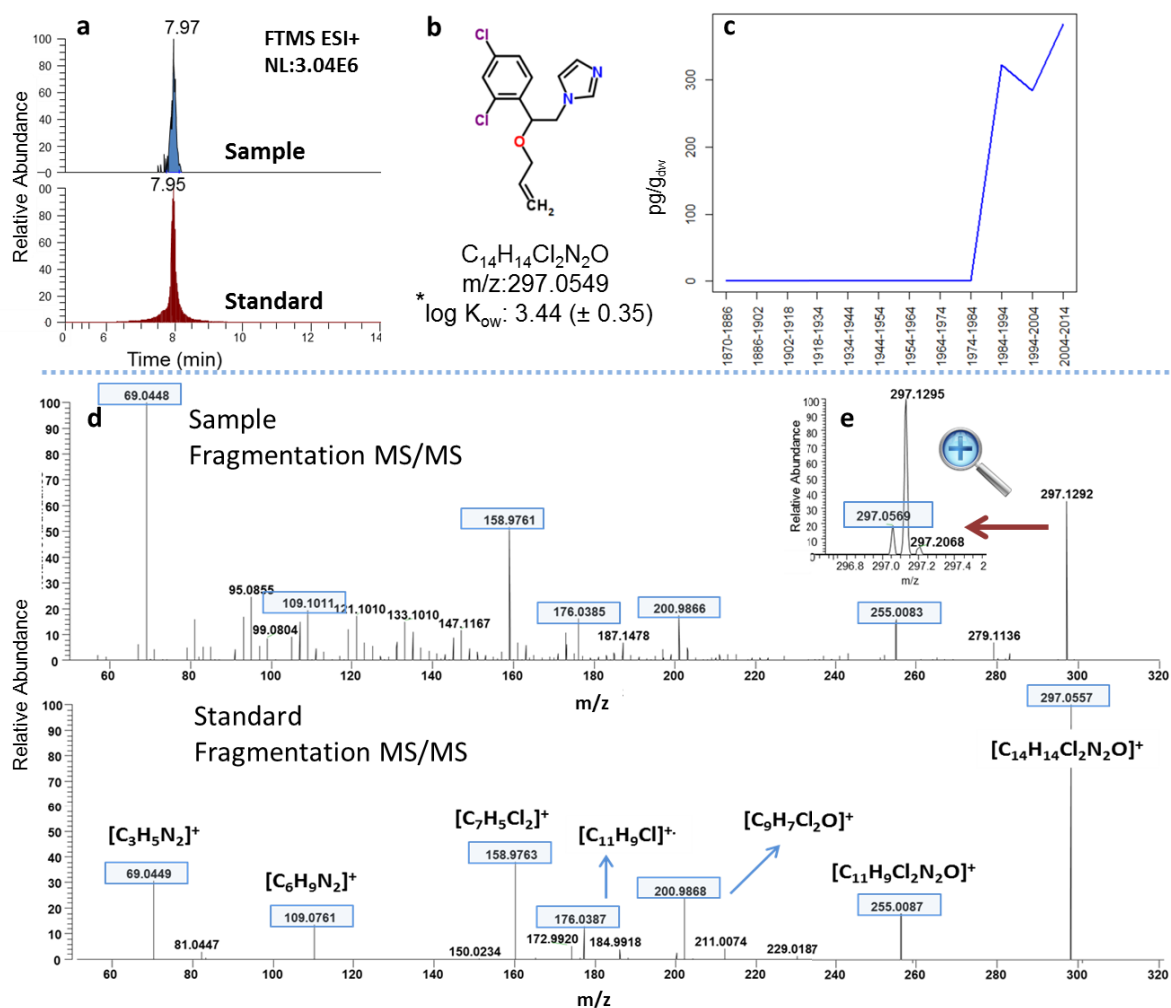


Figure 3. Examples of four clusters from Greifensee: two clusters peak around the 1970s and 1980s (a and c), and two clusters where the compounds increase in intensity over time (b and d). Grey lines are unidentified (unknown) compounds. Blue lines indicate previously published compounds (dichlorophen, irgarol, and triclocarban).^{9, 10} No compounds were identified in the cluster shown in panel d. The number of profiles in each cluster is also listed. The Great Acceleration is illustrated in shadow red with two lines illustrating the construction of the first WWTP in the catchment of Greifensee in 1956 (red dashed line) and later the upgrade of the WWTP and the implementation of environmental regulations, policies, and management practices in Switzerland in the 1980s (red dotted line).



¹Molecular formulas were assigned using Mass Frontier Spectral Interpretation Software (ThermoScientific, San Jose, U.S.A.).

*log K_{ow} values were predicted using vclab software.²⁷

Figure 4. Identification of the fungicide imazalil. Chromatograms of sample and reference standard (a), imazalil structure and log K_{ow} index value (b), estimated sediment concentration pattern over time (c) and MS/MS spectra used for the identification of the fungicide imazalil (d). For clarification purposes, a zoomed-in view of the molecular ion in the sample spectrum is illustrated with a red arrow (e). The most probable molecular formulas have been assigned for the main fragments.

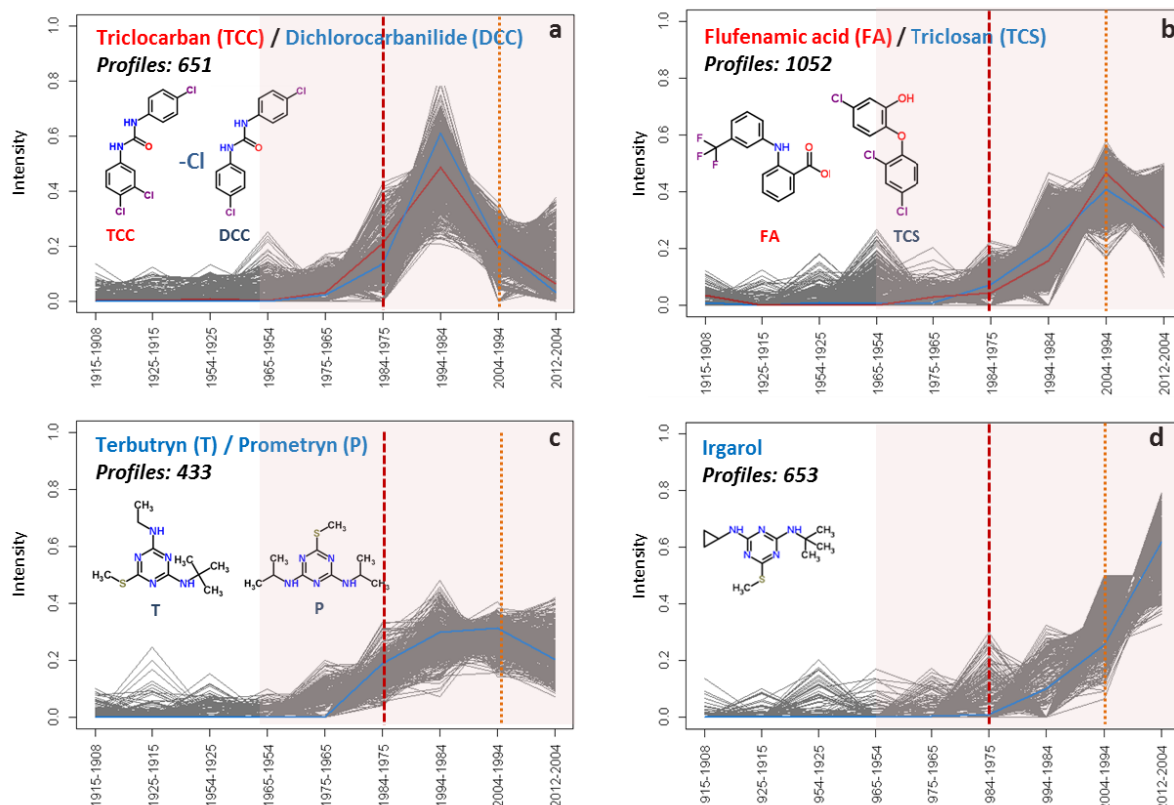


Figure 5. Characterization of Lake Lugano sediments showing four example clusters. Two clusters peak in intensities in the 1980s (a) and 1990s (b). Two clusters demonstrate increasing intensities towards the present day (c and d). Grey lines are unidentified (unknown) compounds. Blue and red lines correspond to the antibacterial agent triclocarban and its transformation product dichlorocarbanilide (a), the pharmaceutical flufenamic acid and the antibacterial triclosan (b), the isomeric pesticides terbutryn and prometryn (c) and irgarol (d) reported to be present in sediments of Lake Lugano.⁹ The number of profiles in each cluster is also compiled. The Great Acceleration is illustrated in shadow red with two lines illustrating the beginning of the first WWTP construction in the catchment of Lake Lugano (red dashed line) in 1976 and later the upgrade of some WWTPs in 1995 (orange dotted line).