



Data-Mining von LC-HRMS-Daten zur Bewertung von Wasseraufbereitungsprozessen

Tobias Bader (bader.t@lw-online.de)

Zweckverband Landeswasserversorgung, Langenau

Zusammenfassung

Die Überwachung von Wasseraufbereitungsprozessen hinsichtlich des Verhaltens organischer Spurenstoffe wird bislang anhand relativ weniger bekannter Substanzen durchgeführt. Durch Kopplung von Flüssigkeitschromatographie mit hochauflösender Massenspektrometrie (LC-HRMS) lassen sich beim sogenannten Non-Target-Screening (NTS) auch unbekannte Spurenstoffe detektieren. Das NTS ist daher prädestiniert für eine weitgehende Überwachung von Aufbereitungsprozessen.

Die bei der automatisierten Datenauswertung verwendeten Algorithmen zeigen eine Anfälligkeit gegenüber falsch-positiv- bzw. falsch-negativ-Befunden. In dieser Arbeit wurden daher Strategien zur verlässlichen Auswertung von LC-HRMS-Daten entwickelt. Als Grundlage dienten Replikate, welche in einem kombinatorischen Validierungsansatz näher betrachtet wurden. Die rekursive Signalintegration führte zu einer verlässlicheren Signalklassifikation, die zur Bewertung von Prozessen herangezogen werden kann.

Einleitung

Die Emission von anthropogenen Spurenstoffen in die aquatische Umwelt stellt Wasserversorger fortwährend vor neue Herausforderungen [1-3]. Mit organischen Spurenstoffen belastete Rohwässer erfordern komplexe Aufbereitungsverfahren zur Sicherung der Trinkwasserqualität. Sowohl die routinemäßige Überwachung dieser Rohwässer als auch das Verhalten organischer Spurenstoffe während verschiedener Aufbereitungsprozesse ist von großem Interesse, um etwaige Gefahren frühzeitig zu erkennen [4]. Herkömmlich verwendete Analysetechniken erlauben die Überwachung bekannter Substanzen im Spurenbereich, versagen jedoch meist beim Erkennen unbekannter Verbindungen. Im Hinblick auf die Überwachung von Aufbereitungsprozessen stellt dies eine große Einschränkung dar, da Transformationsprodukte oftmals nicht oder nur unzureichend charakterisiert bzw. bekannt sind und demnach nur sporadisch erfasst werden können [5, 6]. Mit dem Non-Target-Screening (NTS) lassen sich tausende Verbindungen in einer Analyse erfassen und neben bekannten auch unbekannte Spurenstoffe detektieren [7, 8]. Im Hinblick auf die Überwachung von Roh- und Prozesswässern bei der Wasseraufbereitung stellt dies einen entscheidenden Vorteil gegenüber etablierten analytischen Techniken dar. Während die LC-HRMS-Messtechnik in den letzten Jahren stark weiterentwickelt wurde, zeigen die Auswertalgorithmen jedoch noch Schwachstellen. Die vorliegende Arbeit befasst sich daher mit

Strategien zur verlässlichen Auswertung von LC-HRMS-Daten [9].

Workflow

Aus den dreidimensionalen LC-HRMS-Daten (s. Abb. 1) werden mittels Peak-Finding-Algorithmen alle Signale (Features) extrahiert. Jedes Feature wird über das Masse-zu-Ladungsverhältnis (m/z), die Retentionszeit (RT) und die Intensität (Int) charakterisiert. Beim Peak-Alignment werden die Signale von mehreren Messungen (s. Abb. 1 Proben A, B und C) anhand von Masse- und Retentionszeit-Toleranzen zu einer Tabelle zusammengefasst. Hierdurch kann für jedes Feature der Intensitätsverlauf über mehrere Proben (z. B. Zeitreihen, Prozessketten etc.) betrachtet werden.

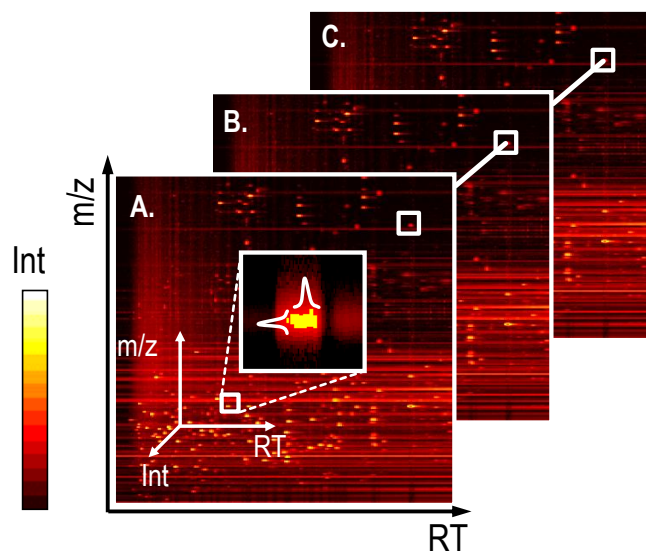


Abb. 1. Schematische Darstellung der dreidimensionalen LC-HRMS-Datenstruktur, Vergleich mehrerer Proben A, B und C nach Peak-Alignment.

Falsch-positiv- und falsch-negativ-Problematik

Basierend auf Wiederholmessungen (technische Replikate) konnte die Problematik von falsch-positiv- (Feature detektiert obwohl nicht vorhanden) sowie falsch-negativ-Befunden (Feature nicht detektiert obwohl vorhanden) beleuchtet werden. Hier wurde die gleiche Probe 21-fach analysiert. Im Idealfall würde jedes Feature in allen 21 Replikaten erfasst werden (Erkennungsrate 100 %). In Abb. 2 wurden verschiedene Thresholds (untere Signalgrenzen) herangezogen, um die gleichen Daten zu prozessieren. Während bei einem Threshold von 100 cps die Erkennungsrate unter 50 % liegt, steigt diese auf über 95 % bei einem Threshold von 1000 cps. Gleichwohl

nimmt jedoch die Anzahl der erfassten Features mit steigendem Threshold deutlich ab.

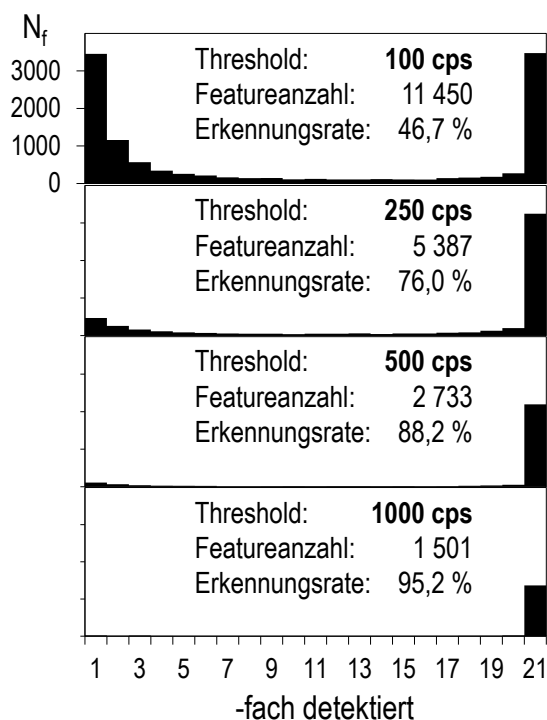


Abb. 2. Verteilung der Erkennungsrate über 21 Replikatmessungen bei verschiedenen Thresholds.

Die Wahrscheinlichkeit falsch-positiv-Befunde zu erzeugen nimmt mit steigendem Threshold ab, die Wahrscheinlichkeit falsch-negativ-Befunde zu erzeugen hingegen zu. Während sich falsch-positiv-Befunde im Nachgang durch weitere Filter wieder korrigieren lassen, sind falsch-negativ-Befunde irreversibel aus den Daten entfernt. Daher wurden die Kriterien beim Peak-Finding großzügig gewählt (hier: 100 cps) und wissentlich viele falsch-positiv-Befunde generiert, welche im Nachgang jedoch gefiltert werden.

Filterkriterien: Auswahl, Parametrierung und Validierung

Zur Reduktion von falsch-positiv-Befunden wurde ein Filterkonzept entwickelt. Im ersten Schritt wurden knapp 3000 Features in einem Probenreplikat manuell betrachtet und in die Gruppen „Peaks“ bzw. „No Peaks“ eingeteilt (→ Trainingsdaten). Ausgehend von dieser Einteilung wurden für verschiedene Kenngrößen eines Peaks die Verteilungen in den Gruppen „Peaks“ und „No Peaks“ betrachtet, um darüber die Filter zu parametrieren. Aus den Verteilungen der Kenngrößen wurden die Grenzen (Min, Max und Standardabweichung) abgeleitet. Beispielhaft sind in Abb. 3 die Verteilungen für die Peakbreiten dargestellt. Des Weiteren wurde die Schwankung über das Triplikat in Form der Standardabweichung einbezogen. Erwartungsgemäß zeigt sich, dass die Schwankung bei echten Peaks sehr viel geringer ausfällt als beim Rauschen („No Peaks“) und dieses Kriterium daher einen effizienten Filter darstellt. In gleicher Weise wurden mehr als 20 Kenngrößen parametriert.

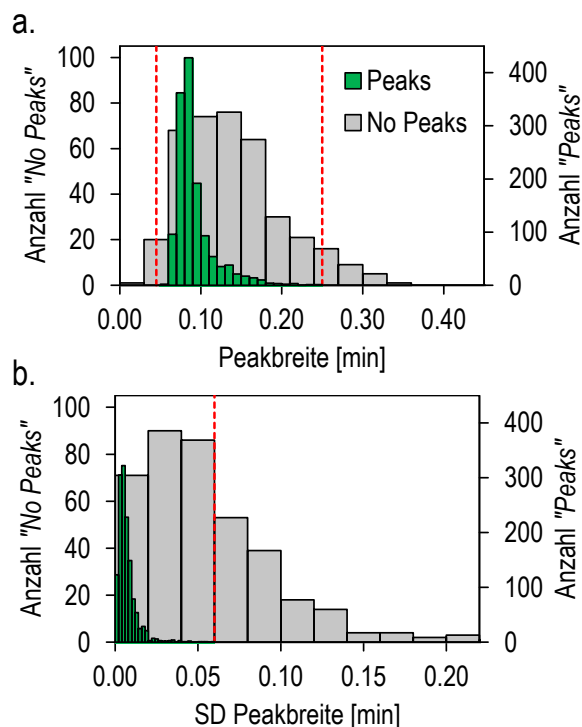


Abb. 3: Verteilung der a. Peakbreiten (bei 50 % der Peakhöhe) und b. Standardabweichungen (SD) der Peakbreiten für die Gruppen „Peaks“ und „No Peaks“. Nachdruck mit Genehmigung von [9]. Copyright 2017 American Chemical Society.

Die Validierung erfolgte in mehreren Stufen. Zum einen wurde ein Testdatensatz mit mehr als 38 000 Peaks erneut manuell klassifiziert und mit den Ergebnissen der Filtermethode verglichen. Die Kennzahlen Precision (> 98 %) und Recall (> 96%) sowie die falsch-positiv-Rate (< 5 %) unterstreichen die Güte des Modells.

Im zweiten Schritt wurden 411 Targets bei einer Konzentration von 0,1 µg/L betrachtet. Von den 411 Targets waren nach Durchlaufen des gesamten Workflows noch 406 vorhanden. Aufgrund dieser Daten kann die falsch-negativ-Rate als gering eingeschätzt werden.

Signalklassifizierung mittels rekursiver Integration

Zur Beschreibung von Prozessen werden die Features im Prozesszu- und ablauf verfolgt. Basierend auf der relativen Intensitätsänderung (fold change, fc) zwischen Zu- und Ablauf werden die Signale nach Tab. 1 eingeteilt. Der Einfluss von Matrixeffekten wird über dotierte isotope-markierte Standards (IS) abgeschätzt. Liegen alle IS nach dem Prozessvergleich in der Kategorie Konstanz, spielen Matrixeffekte eine untergeordnete Rolle.

Tab. 1: Signalklassifikation basieren auf relativer Intensitätsänderung

Gruppe	Kriterium
(E) Elimination	$0,0 \leq fc < 0,2$
(T) Teilelimination	$0,2 \leq fc < 0,5$
(K) Konstanz	$0,5 \leq fc \leq 2,0$
(Z) Zunahme	$2,0 < fc \leq 5,0$
(N) Neubildung	$5,0 < fc \leq \infty$

$$\text{fold change: } fc = \frac{\text{Signalintensität}_{\text{Ablauf}}}{\text{Signalintensität}_{\text{Zulauf}}}$$

Um eine verlässlichere Klassifizierung zu erreichen, wurde ein zweiter Integrationsalgorithmus (*Summation*) etabliert. Die Besonderheit besteht darin, dass immer eine Integration erzwungen wird. Auch für nicht erfasste Features lässt sich eine Signalhöhe rekursiv aus dem Rauschen extrahieren (s. Abb. 4 b).

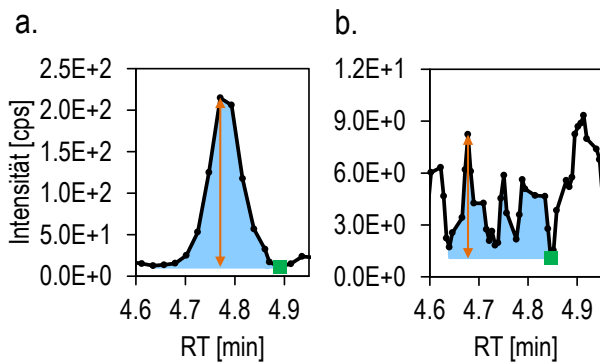


Abb. 4: Extrahiertes Ionenchromatogramm aus a.) Prozesszu- und b.) -ablauf, integriert mittels Summation-Algorithmus. Die Peakhöhe wird zwischen dem Datenpunkt mit minimaler (grünes Quadrat) und maximaler Intensität berechnet. Nachdruck mit Genehmigung von [9]. Copyright 2017 American Chemical Society.

Dies führt zu einer deutlich verlässlicheren Eingruppierung der Signale, vor allem in der Nähe des Thresholds, beispielhaft in Abb. 4 dargestellt.

- a. Zulauf: 204 cps
- b. Ablauf: kein Peak erfasst
- Threshold: 200 cps

Die fehlende Signalhöhe im Ablauf könnte mit „0“ (optimistische Betrachtung) oder dem Threshold (pessimistische Betrachtung) ersetzt werden. Basierend auf dieser Festlegung würde das Signal entweder in die Gruppe Elimination (0 cps / 204 cps) oder in die Gruppe Konstanz (200 cps / 204 cps) eingeordnet werden. Durch die rekursive Methode wird die Rauschhöhe im gleichen Retentionszeitfenster (s. Abb. 4 b) herangezogen um die „richtige“ Klassifizierung - in diesem Falle: Elimination (7 cps / 204 cps) - vorzunehmen.

Kombinatorischer Ansatz zur Beurteilung der Wiederholbarkeit

Um die Wiederholbarkeit des Verfahrens für die Prozessbeschreibung zu validieren, wurde eine kombinatorische Herangehensweise gewählt. Als Datengrundlage diente jeweils eine Neunfachmessung des Zu- und Ablaufs einer Ozonungsanlage (Wasserwerk Langenau, Donauwasseraufbereitung).

Zur Plausibilitätsprüfung wurden theoretische Prozesse innerhalb der Replikate von den gleichen Proben (s. Abb. 5 a und b) betrachtet. Insgesamt lassen sich 840 Prozessvergleiche (theoretische Prozesse) innerhalb einer Probe durchführen. Im Idealfall würden alle Features stets der Kategorie Konstanz zugeordnet werden. Beim realen Prozess der Ozonung wurden ebenfalls alle paarweisen Triplikatsvergleiche durchgeführt (Abb. 5 c). Für diesen Fall lassen sich 7056 mögliche Prozessvergleiche (realer Prozess) bilden.

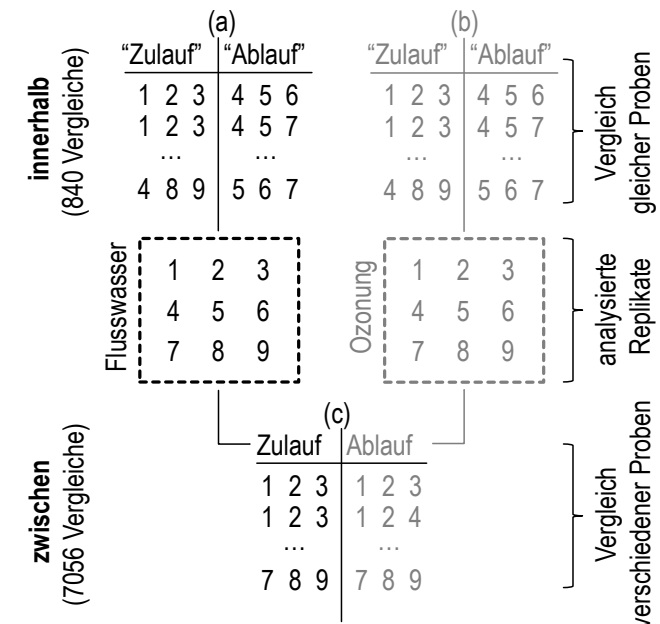


Abb. 5: Paarweiser Triplikatsvergleich von 9 Replikaten innerhalb einer Probe (a und b, theoretische Prozesse) sowie zwischen den Proben (c, realer Prozess). Nachdruck mit Genehmigung von [9]. Copyright 2017 American Chemical Society.

Für jeden Prozessvergleich wurde die Featureanzahl in den jeweiligen Gruppen betrachtet. In Abb. 6 sind die Daten als Boxplots dargestellt. Die Vergleiche innerhalb der Proben (Abb. 6 a und b) zeigen, dass fast alle Features über alle Prozessvergleiche stets erwartungsgemäß der Kategorie Konstanz zugeordnet wurden. Lediglich einige wenige Ausreißer (rote Pluszeichen) fallen teilweise in die verbleibenden Gruppen. Der Großteil dieser Ausreißer lässt sich auf reale - wenn auch nicht plausible - Unterschiede in den Replikaten zurückführen. Die Breite der Boxen zeigt außerdem, dass die Featureanzahl in der Gruppe Konstanz wiederholbar erreicht wird. Die höhere Boxbreite bei (b) lässt sich vermutlich auf geringere Signal-

intensitäten und damit höhere Unsicherheiten in der Ablaufprobe zurückführen. Bei der Betrachtung des realen Prozesses der Ozonung (Abb. 6 c) zeigt sich ein anderes Bild. Der Großteil der Features wird den Gruppen Elimination und Teilelimination zugeordnet, was die Effizienz des Prozesses verdeutlicht. Erneut zeigen die Boxbreiten eine ausreichende Wiederholbarkeit über alle 7056 Prozessvergleiche.

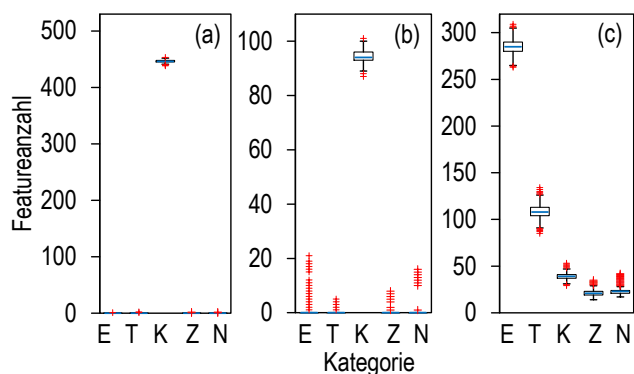


Abb. 6: Boxplots der Featureanzahlen, welche den jeweiligen Kategorien (E, T, K, Z, N) zugeordnet wurden; (a) und (b) jeweils 840 Vergleiche, (c) 7056 Prozessvergleiche. Nachdruck mit Genehmigung von [9]. Copyright 2017 American Chemical Society.

Für den Prozess der Ozonung wurden tendenziell mehr Transformationsprodukte (Kategorie N) erwartet. Oftmals sind die polaren Ozonungs-Transformationsprodukte schlechter oder nicht mittels Umkehrphasenchromatographie zu erfassen. In zukünftigen Arbeiten müssen daher Screening-Methoden für (hoch)polare Spurenstoffe, beispielsweise mit der *Hydrophilic interaction liquid chromatography* (HILIC) oder mittels *Mixed-mode chromatography* (MMC) entwickelt werden.

Fazit und Ausblick

In dieser Arbeit wurden ein Konzept und eine LC-HRMS-Datenprozessierungsstrategie zur Bewertung von Wasseraufbereitungsprozessen entwickelt. Dabei wurde eine Methode zur Filterung von falsch-positiv-Befunden entwickelt und validiert. Ein wichtiger Meilenstein war die Triplikat-Analyse und die damit verbundene Möglichkeit, auch die Signalschwankungen über Replikatmessungen als effiziente Filtermethode einzusetzen. Durch die entwickelten Herangehensweisen konnte die Datenqualität stark verbessert werden. Auf dieser Grundlage wurde ein Konzept zum verlässlicheren Probenvergleich erarbeitet. Ein Schlüsselpunkt war hierbei die rekursive Integration zur verlässlicheren Berechnung des *fold change*. Beim kombinatorischen Validierungskonzept konnte eine ausreichende Wiederholbarkeit festgestellt werden.

Das Non-Target-Screening eröffnet eine neue Dimension zur Überwachung von organischen Spurenstoffen bei der Wasseraufbereitung. Beispielsweise lassen sich Betriebsbedingungen auf einer deutlich umfassenderen Datengrundlage optimieren. Dieses umfassendere Bild wird auch dabei helfen, Wissens-

lücken zu reduzieren. Die Anwendung der entwickelten Methoden ist jedoch nicht nur auf Aufbereitungsprozesse beschränkt, sondern kann auf andere Felder übertragen werden.

Mining of LC-HRMS data for the assessment of water treatment processes, Dissertation von Tobias Bader, Betreuer: Prof. Dr. Klaus Kümmerer, Leuphana Universität Lüneburg, Fakultät Nachhaltigkeit, Dr. Wolfgang Schulz, Zweckverband Landeswasserversorgung, kostenloser Download unter <http://opus.uni-lueneburg.de/opus/volltexte/2018/14493/>

Literatur

- [1] Schwarzenbach, R. P. et al., *Science* 2006, 313, 1072-1077.
- [2] Richardson, S. D., Ternes, T. A., *Anal. Chem.* 2014, 86, 2813-2848.
- [3] Howard, P. H., Muir, D. C. G., *Environ. Sci. Technol.* 2010, 44, 2277-2285
- [4] Jekel, M. et al., *Chemosphere* 2015, 125, 155-167.
- [5] Petrie, B. et al., *Water Res.* 2015, 72, 3-27.
- [6] Menz, J. et al., *Environ. Int.* 2017, 98, 171-180.
- [7] Moschet, C. et al., *Anal. Chem.* 2013, 85, 10312-10320.
- [8] Schymanski, E. L. et al., *Anal. Bioanal. Chem.* 2015, 407, 6237-6255.
- [9] Bader, T. et al., *Anal. Chem.* 2017, 89, 13219-13226.

Korrespondenzadresse

Dr. Tobias Bader
Zweckverband Landeswasserversorgung
Betriebs- und Forschungslabor
Am Spitzigen Berg 1
89129 Langenau
E-Mail: bader.t@lw-online.de
Tel. +49 (0) 7345 9638 2865