



## **AG Chemometrie und multivariate Datenauswertung Jahresbericht 2018**

**Obmann: Prof. Dr. Philipp Weller, Mannheim**

Die Arbeitsgruppe Chemometrie und multivariate Datenanalyse hatte 2018 einen Mitgliederstand von 10 aktiven und 6 korrespondierenden Mitgliedern.

Im Berichtsjahr 2018 und den Sitzungen der Arbeitsgruppe, die gemäß eines gemeinsamen Beschlusses von 2017 zwei Mal im Jahr am 8.5.18 und am 12.11.18 stattfanden, wurden folgende Schwerpunktthemen diskutiert und bearbeitet:

### *Weiterführung des Aufbaus der AG:*

Die AG Chemometrie und multivariate Datenanalyse befindet sich im zweiten Jahr nach ihrer Gründung und damit immer noch im Aufbau. Es hat sich gezeigt, dass die Vorstellungen der AG-Mitglieder zum Leitthema der Arbeitsgruppe eine größere Bandbreite aufweisen, als zu Beginn angenommen. Es ist daher sinnvoll, im kommenden Jahr die Ausrichtung zu präzisieren und die Aufgaben der AG zu schärfen. Schwerpunkte für die AG werden vor allem sein:

- Standardisierung von Prozessen und Strategien zur Datenanalyse und geeigneter Validierung
- Erarbeitung von Toolboxen für die Datenanalyse aus typischen Analysesystemen
- Multivariate Analyse in der Lehre: Einbindung in die Ausbildung von Studenten

### *Zusammenarbeit mit anderen Arbeitsgruppen innerhalb der LChG:*

Ein wichtiger Aspekt ist die Anwendung der Datenanalyse seitens der Anwender aus den anderen Fachgruppen, z.B. Isotopenanalytik, Aromen, etc., um die Anwendbarkeit der Werkzeuge in der Realität zu evaluieren und anzupassen. Die AG Chemometrie und MVA hat es sich daher zum Ziel gesetzt, die Interaktion mit anderen Fachgruppen zu intensivieren.

### *Zusammenarbeit mit anderen Chemometrie-Gruppen:*

Über die Zusammenarbeit innerhalb der LChG hinaus ist angedacht, auch mit internationalen Experten im Bereich der Chemometrie zusammen zu arbeiten, so wurde 2017 der Experte Prof. Dr. Leardi aus Genua zur Sitzung eingeladen, für das kommende Jahr sind ähnliche Gastvorträge angedacht.

### *Besuch auf dem Forschungs-Campus der Robert Bosch GmbH:*

Das Wissen und das Verständnis der Anwendung der Chemometrie und Datenanalyse in der Praxis soll durch regelmäßige Interaktionen und Besuche bei Firmen und Behörden verbessert werden. In diesem Rahmen wurde der zweite Sitzungstermin der AG auf dem Gelände des Forschungs-Campus der Robert Bosch GmbH abgehalten. Hierbei ergab sich die Gelegenheit zur regen Diskussion über die Anwendung von chemometrischen Tools und Ansätzen in der Analytik.

### *Evaluierung von Software-Paketen für die Datenanalyse:*

Ein zentraler Aspekt der multivariaten Datenanalyse ist es, geeignete Software-Tools dafür zu identifizieren. Im kommerziellen Umfeld sind im Wesentlichen MATLAB, Tableau oder RapidMiner in den vorderen Positionen zu finden und weisen ein sehr großes Leistungsspektrum auf. Problematisch hierbei wird von der Arbeitsgruppe vor allem das jeweilige Lizenzierungsmodell gesehen, das zum Teil Gebühren von mehreren Zehntausend Euro pro Jahr vorsieht und dementsprechend für viele kommerzielle Anwender kaum nutzbar sein kann.

Aus diesem Grund sieht die AG großes Potential in open-source Lösungen, die kostenfrei nutzbar sind und über eine sehr große, aktive Community verfügen, die beständig Pakete und Toolboxen weiterentwickelt.

Als wichtige Systeme wurden hier von der AG diskutiert und vorgestellt:

- R (Grundpaket, sowie angepasste Distributionen, wie bspw. R-GUI von R. Leardi)
- Orange3
- KNIME
- Python
- OpenChrom mit der Option, multivariate Analyse zukünftig auch für Spektroskopie durchführen zu können

Diese Plattformen können einzeln oder gemeinsam entsprechend auf die Bedürfnisse der Anwender angepasst werden.

*Standardisierung von Datenanalyse-Strategien:*

Als eines der Kernprobleme der Chemometrie/MVA sieht die AG die fehlende Standardisierung in der Anwendung: wenn zwei Anwender von „ihrer“ PCA sprechen, was wird gemacht? Im Hinblick auf das Ziel eines flächendeckenden Einsatzes der multivariaten Analyse sieht hier die AG großen Bedarf in der Erarbeitung von Vorschlägen für die Anwendung typischer Datenanalyse-Tools, wie PCA, LDA, PLS-DA etc. Ein Aspekt hierbei ist auch die Präzisierung der Terminologie.

Der Vorschlag der Arbeitsgruppe ist die Durchführung von Workshops, bei der Anwender einen in der AG erarbeiteten Workflow, z.B. mit IR- oder NMR-Daten selbst anwenden und nachvollziehen können.

*Angebot von Workshops zur Datenanalyse:*

Als direkte Konsequenz aus den zuvor genannten Diskussionen zum Thema Standardisierung wurde in der AG beschlossen, 2019 einen oder mehrere Workshops anzubieten. Ein Vorschlag ist hierbei ein Workshop auf dem Lebensmittelchemikertag 2019 in Dresden, weitere Möglichkeiten wären beispielsweise auch die Regionaltagungen. Ziel ist hierbei vor allem, auch den Nachwuchs zeitig an die Anwendung solcher Werkzeuge zu bringen und das Bewusstsein für die Potentiale zu schärfen.