Application Note

Consumer Products

Agilent

Trusted Answers

Agilent AI Peak Integration for GC/MS Analysis of Phthalates

Authors

Ruoji Luo,
Tamas King,
Winnie Chau,
Thomas Bispham
Agilent Technologies, Inc.

Abstract

Gas chromatography/mass spectrometry (GC/MS) data analysis is often manual, time consuming, and human reliant. Even though automated integrators are available, a chemist may still need to manually correct integration baselines. This application note demonstrates a workflow using artificial intelligence (AI) during the data analysis process for phthalates. A machine learning (ML) model is custom-trained during a user's normal data analysis workflow, observing manual integration actions. It then replaces manual peak integration with adaptable AI-assisted peak detection and integration with continuous learning. Agilent AI Peak Integration for MassHunter Quantitative Analysis Software produced reproducible and accurate results. It also dramatically reduced the data analysis time.

Introduction

One of the most time-consuming processes in test laboratories is manual integration during data analysis, which includes baseline correction, peak combination or splitting, and eliminating false positive/negative peaks. There are many reasons manual integration might be needed, such as contamination, matrix effect, insufficient separation, non-Gaussian peak shape (fronting/tailing), and the nature of analytes.

An example of a particularly difficult analyte is phthalates. Phthalates, or phthalic acid esters, are used frequently as plasticizers. They can be found in a large variety of products, such as toys, perfume, electrical and electronic equipment. Due to their potential health risks, the concentration in consumer products/food is regulated worldwide.[1] In the case of phthalate analysis, isomeric compounds, such as diisononyl phthalate (DINP) and diisodecyl phthalate (DIDP), result in broad, irregularly shaped peaks usually requiring additional manual integration.[2]

In this application note, we demonstrate a new workflow of phthalate analysis using GC/MS, where the manual integration is replaced by Agilent AI Peak Integration, a plug-in for Agilent MassHunter Quantitative Analysis software.

Experimental

Equipment and software

Agilent GC/MSD instruments were used to acquire data. Data were acquired in synchronous SIM/Scan mode. Agilent MassHunter Quantitative Analysis software (Version 10.2) was used for data analysis.

The acquired data were first processed using a built-in parameter-less integrator from MassHunter Quantitative Analysis. The integration results obtained from this built-in integrator were used as the basis for the machine learning (ML) model. Then, these results are uploaded to the cloud-based AI Peak Integration database through a plug-in to MassHunter Quantitative Analysis.

Machine learning model

To use Agilent AI Peak Integration, an ML model for a specific application must be trained. Proper training requires high quality, real-world data. Additionally, the way the model is trained by the user has a tremendous impact on the model's accuracy. A technical overview for AI Peak Integration covers the following aspects:[3]

- Key features of AI Peak Integration
- Development and training of AI Peak Integration
- Quality and accuracy control of AI Peak Integration

For this application note an ML model was developed for the analysis of phthalates in consumer products. Data from over 1000 samples were used to train this model in phthalate detection and quantification.

Data analysis workflow

Figure 1 illustrates a general workflow in MassHunter Quantitative Analysis software with AI Peak Integration. After successful user authentication, the user prepares the batch of samples as they usually would without AI Peak Integration in their standard workflow. This includes adding samples and loading the quantification method. Once the batch of samples is prepared, which means that the batch is analyzed and the peaks integrated by the built-in integrator, the user can then initiate the prediction by clicking the "Process" button in the UI. A trained model in the cloud will predict the correct integrations and update the results in MassHunter Quantitative Analysis software. In an ideal case, the user can directly save this batch with the predicted results.

One of the key features of AI Peak Integration is continuous learning. The optional manual adjustment step (Figure 1) is designed to capture any changes the user made to the peaks. This allows the model to continuously improve itself and quickly adapt to the new dataset.

Figure 1. Agilent AI Peak Integration assisted workflow for data analysis in MassHunter Quantitative Analysis software.

Results and Discussion

Model reproducibility

Reproducibility is another key feature of AI Peak Integration. Figure 2 shows the integration results for the quantifier ion (*m/z* = 307) of DIDP in the same sample, repeatedly analyzed by AI Peak Integration with a fixed model version. All 5 predictions produced a well-integrated peak with a peak area of 1050 for DIDP at 0.4 ppm.

This feature of AI Peak Integration is extremely important. It is even more important for the analysis of phthalates containing isomeric compounds, such DINP and DIDP, in complex matrices.[2] With this feature users can review the previous results retrospectively or validate the results, which are predicted by a model with a known model version.

Model processing speed

As mentioned in the introduction, for some applications (such as phthalate analysis) analysts spend most of their time reviewing and adjusting the baseline integration that was automatically calculated by the data analysis software. The manual tasks during data analysis are associated with baseline correction, peak combination or splitting, and eliminating false positive/negative peaks. The data analysis review time per each sample is dependent on the type of application, the number of analytes and the complexity of the sample matrix.

The results of the processing speed for a batch with 2, 5, 10, 25, 50 and 100 samples are summarized in Figure 3. The average processing time per sample was about 12 seconds.

For this phthalate analysis under ideal conditions with an experienced analyst, reviewing data and performing manual integration requires on average 60 to120 seconds per sample on average (illustrated in Figure 3, orange area). AI peak integration provides significant time savings and productivity, especially as sample size grows.

Model accuracy

The ability to deliver reliable, accurate predictions is the decisive quality of a trained ML model. With this caliber of quality, the ML model can ultimately reduce overall data analysis time and save money for a company, while increasing the lab's throughput. The model accuracy is calculated using the ML-predicted and manually integrated peak area. Several metrics, including mean error, median error, maximum error, and error standard deviation, are used to evaluate model accuracy.[3]

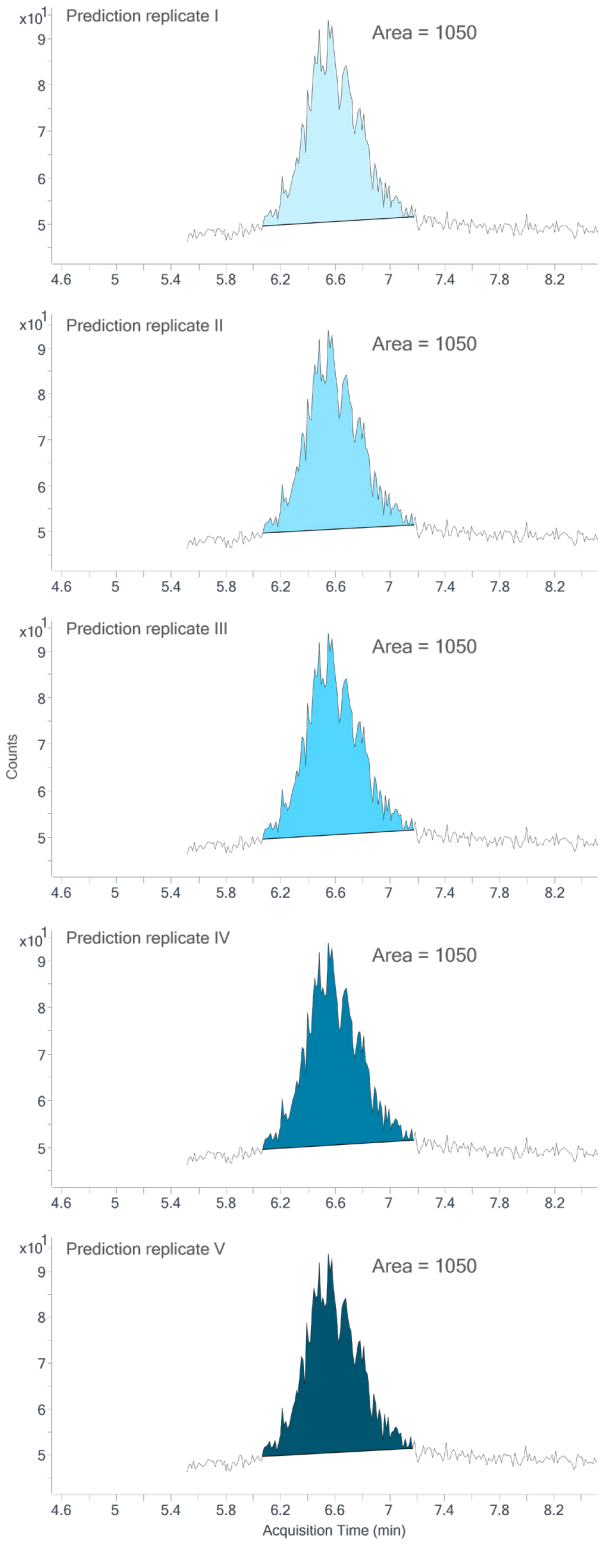

Figure 2. SIM chromatograms for diisodecyl phthalate (DIDP, *m/z* = 307), the same sample from 5 different batches. Results predicted by Agilent AI Peak Integration using the same model.

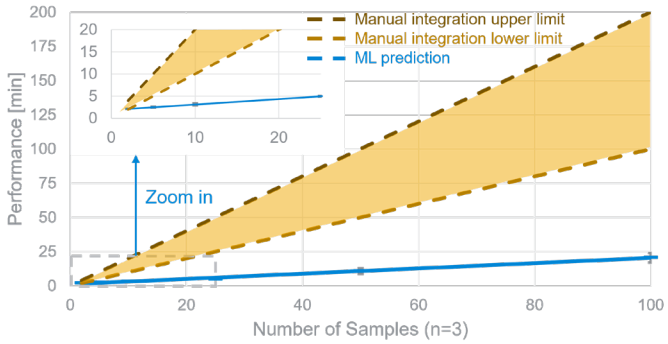

Figure 3. Agilent AI Peak Integration processing speed vs. number of samples per batch (n = 3).

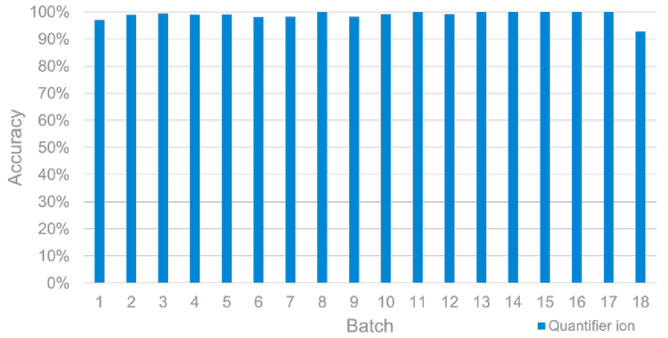

Figure 4. The accuracy for the quantifier ion of each analyte in 576 samples from 18 batches.

Figure 4 demonstrates an excellent model accuracy for the quantifier ion of each phthalate. Over 550 samples from 18 batches were processed by the trained ML model. An average model accuracy of over 98% was achieved.

A trained ML model can not only correctly integrate the quantifier of each phthalate, but also the corresponding qualifier ion(s). This is sometimes challenging for some specific phthalates, such as DIDP and DINP. They partially coelute even with optimized acquisition methods, and they exist as mixtures of positional isomers.[4,5] The results on peak integration by the built-in integrator and AI Peak Integration are compared in Figure 5. Both the quantifier (*m/z* = 293) and qualifier (*m/z* = 149) ion of DINP at 0.3 ppm were perfectly integrated by the ML model.

A well-trained ML model should not only be able correctly to predict the baseline of a peak and therefore its area at a given concentration, but the peaks across the calibration range or sometimes even below the limit of quantification (LOQ). Figure 6 displays the AI Peak Integration results for the quantifier ion (*m/z* = 293) of DINP at 0.1, 0.2, 0.3, 0.4, 0.5, 1, 2, 5, 10, and 20 ppm. Figure 7 shows the calibration curve (0.5 to 20 ppm) of DINP with the peak area of each calibration point predicted by the ML model. Across the concentration range investigated and especially at the lower concentration (0.1 to 0.5 ppm), the trained ML model did an outstanding job in predicting the correct baseline. This significantly reduces the amount of manual adjustment, and therefore the data analysis time.

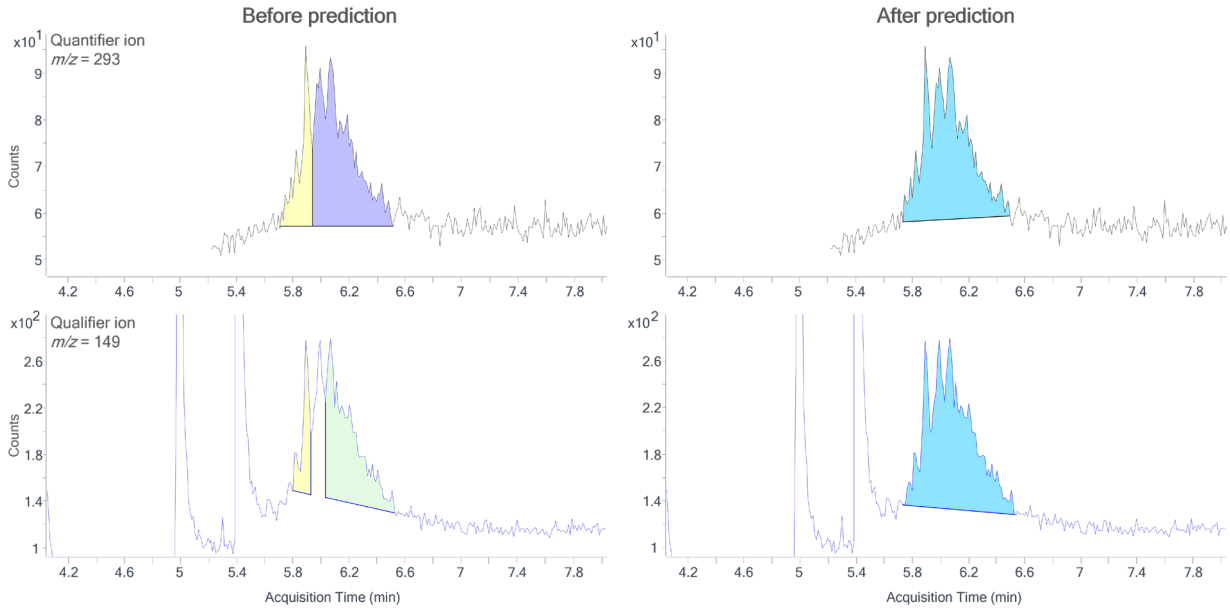

Figure 5. SIM chromatograms for diisononyl phthalate (DINP, *m/z* = 293 (quantifier ion, top row), *m/z* = 149 (qualifier ion, bottom row)), before and after Agilent AI Peak Integration.

Another capability of a well-trained model is the correction of a false positive or negative integration made by the built-in integrator. The technical overview explains the peak screening correctness metrics in detail.[3] Briefly, a false positive peak means that the peak is incorrectly identified as positive, whereas a false negative peak means that the peak is incorrectly identified as negative (zeroed out or integrated incorrectly).

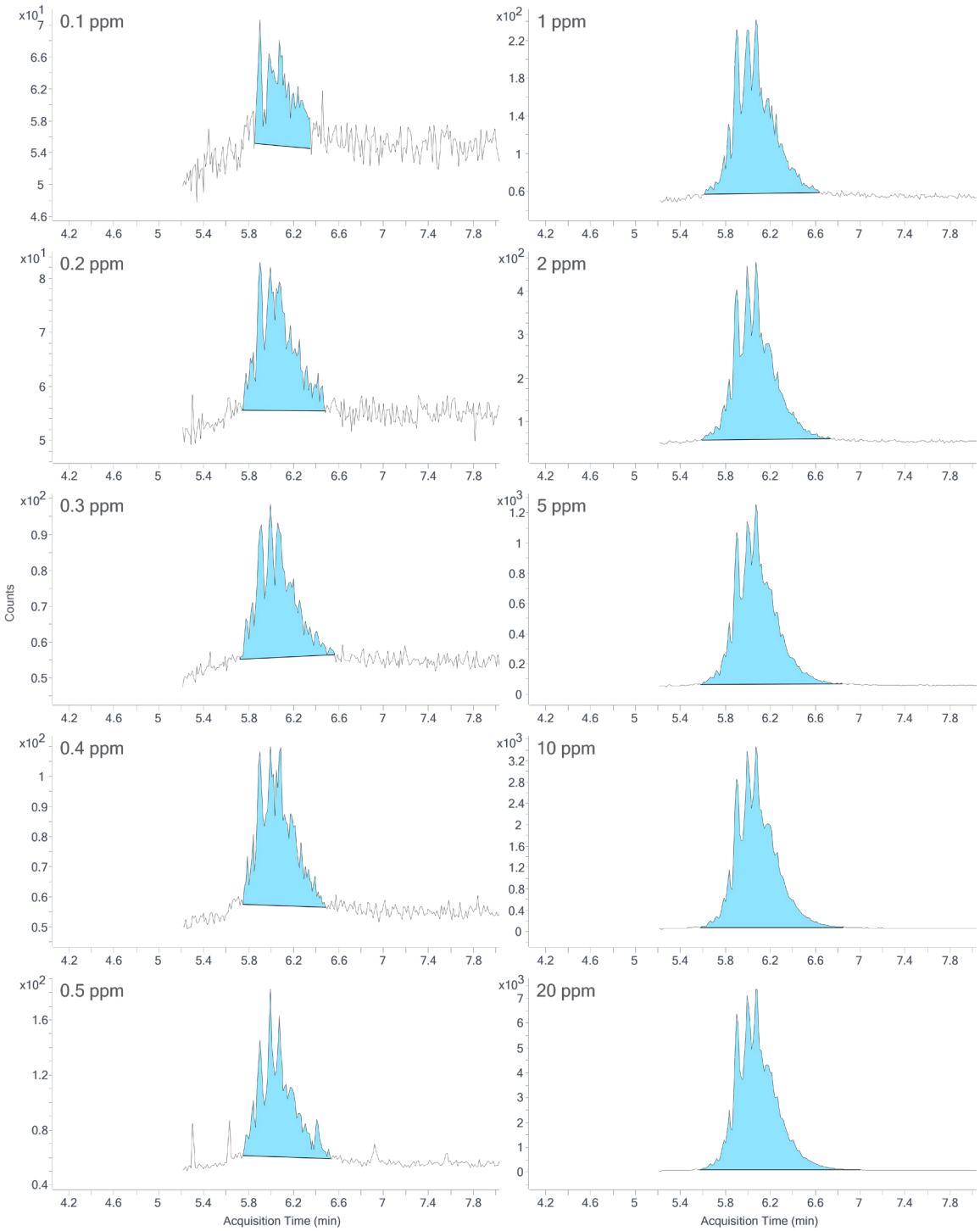

Figure 6. SIM chromatograms for diisononyl phthalate (DINP, *m/z* = 293) at various concentration. Results predicted by AI Peak Integration.

One example for dealing with false negative integrations is already shown in Figure 5. Figure 8 gives an example for dealing with false positive integrations. Here, the false positive integration for the quantifier ion of bis(2-ethylhexyl) phthalate (highlighted in the red dashed rectangular) was deleted by the ML model. That particular ion (*m/z* = 149) belongs to dicyclohexyl phthalate (Figure 8, right column), which elutes almost at the same time as bis(2-ethylhexyl) phthalate in this method.

A trained ML model cannot solve all chromatographic issues (such as unresolved peaks), but Figure 8 demonstrates at least one case, where a trained ML model partially compensated for the challenge of using a faster method resulting in lower chromatographic resolution.

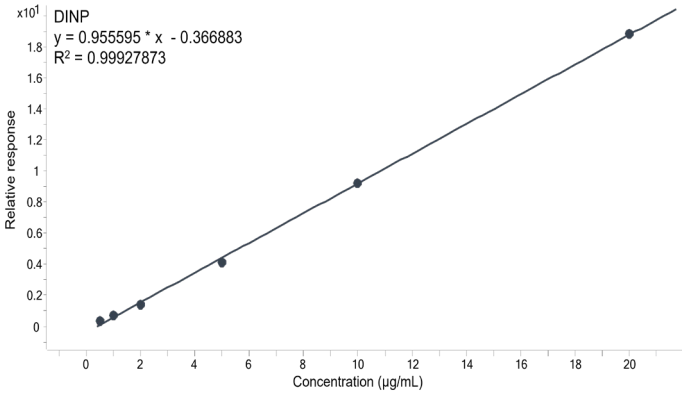

Figure 7. Calibration curve (0.5-20 ppm) of diisononyl phthalate (DINP), results predicted by AI Peak Integration.

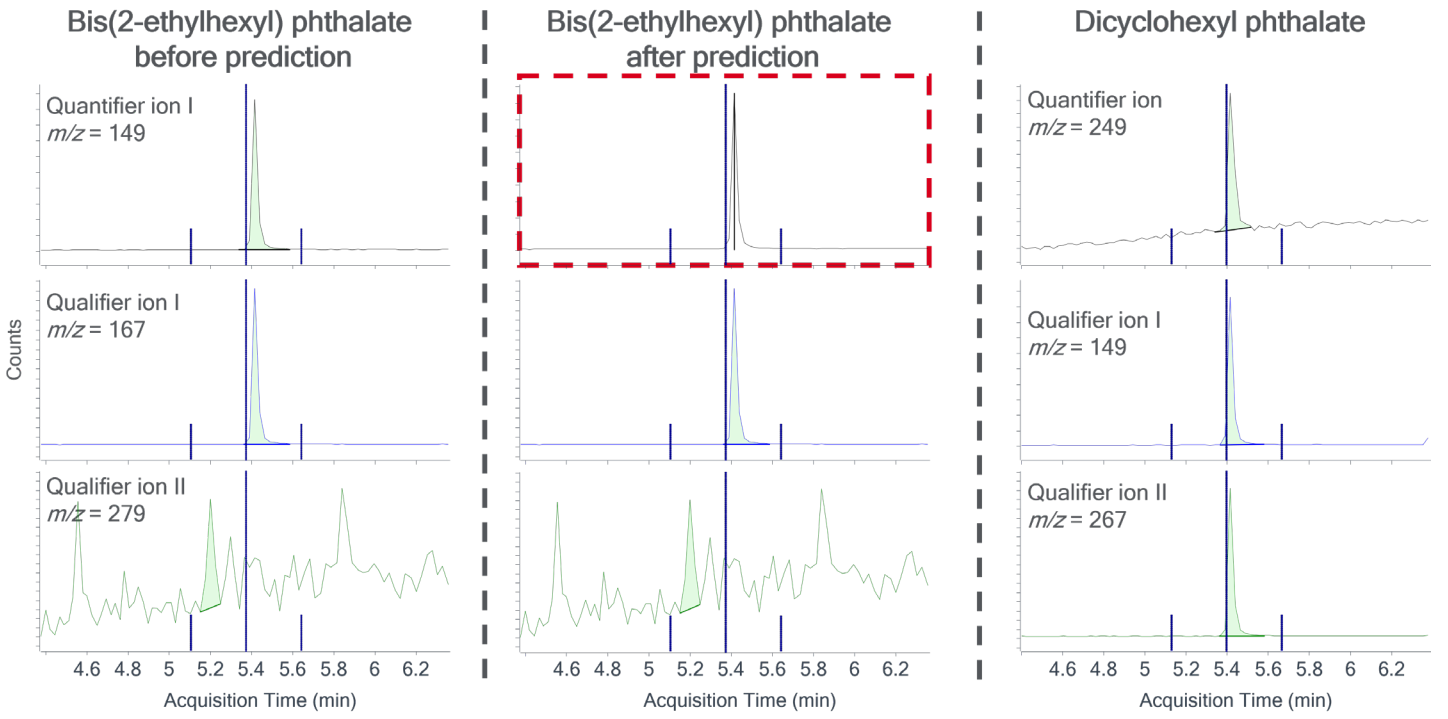

Figure 8. SIM chromatograms (quantifier and qualifier ions) for bis(2-ethylhexyl) phthalate (left: before AI Peak Integration, middle: after Agilent AI Peak Integration) and dicyclohexyl phthalate (right).

Conclusion

In this application note we present Agilent's AI Peak Integration solution for handling traditional manual integration in GC/MS phthalate analysis. The trained machine learning model generated reproducible results when the same sample is processed. It took only 12 seconds on average to process one sample, while for an experienced analyst, this will take between 60 and 120 seconds. The trained model delivered reliable and accurate predictions for the quantifier as well as qualifier ions of targeted phthalates at different concentrations.

References

1. Net, S. *et al.* Occurrence, fate, behavior and ecotoxicological state of phthalates in different environmental matrices, *Environmental Science and Technology*, **2015**, 49 (7), 4019-4035. DOI: 10.1021/es505233b.
2. Dong, C. *et al.* Detecting phthalate esters in sludge particulates from wastewater treatment plants, *Journal of Environmental Science and Health*, Part A **2020**, 55:10, 1233-1240, DOI: 10.1080/10934529.2020.1780850
3. Sosnovshchenko, O.; Luo, R.; King, T. Agilent AI Peak Integration for MassHunter. *Agilent Technologies technical overview*, publication number 5994-6728EN, **2023**
4. Bushey, J. Phthalate Analysis Using an Agilent 8890 GC and an Agilent 5977A GC/MSD. *Agilent Technologies application note*, publication number 5994-0483EN, **2018**.
5. Andrianova, A.; Quimby, B. Rapid Screening Workflow for Phthalates in Plastics by GC/MSD in Under Six Minutes. *Agilent Technologies application note*, publication number 5994-2727EN, **2020**.

To learn more about AI Peak Integration for MassHunter, visit:

www.agilent.com/mass-spec/ai-peak-integration

DE87223787

Printed in the USA, February 23, 2024
5994-7181EN

Agilent
Trusted Answers