Pacific Northwest
NATIONAL LABORATORY

PNNL-30902

Enhanced Ion Detection for Atomic Mass Spectrometry

October 2019

Anthony Carado
Greg Eiden
Jesse Ward
Ron Jasper
April Carman

U.S. DEPARTMENT OF ENERGY

Prepared for the U.S. Department of Energy
under Contract DE-AC05-76RL01830

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor Battelle Memorial Institute, nor any of their employees, **makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights**. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or Battelle Memorial Institute. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from the Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062
www.osti.gov
ph: (865) 576-8401
fax: (865) 576-5728
email: reports@osti.gov

Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
or (703) 605-6000
email: info@ntis.gov
Online ordering: http://www.ntis.gov

PNNL-30902

Enhanced Ion Detection for Atomic Mass Spectrometry

October 2019

Anthony Carado
Greg Eiden
Jesse Ward
Ron Jasper
April Carman

Prepared for
the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99354

Acknowledgments

The research described in this report was conducted under the Laboratory Directed Research and Development (LDRD) Program as a strategic investment at the Pacific Northwest National Laboratory (PNNL), a multiprogram national laboratory operated by Battelle for the U.S. Department of Energy. PNNL is a multi-program national laboratory operated for the U.S. Department of Energy (DOE) by Battelle Memorial Institute under Contract No. DE-AC05-76RL01830.

Contents

Acknowledgments ii
1.0 Introduction 1
2.0 Research Methods and Results 2
3.0 Summary 8

Figures

Figure 1. Plot of classification accuracy vs. training step for the m/z 15, m/z 16 (CH_3, O respectively) with different pulse averaging applied. The NN arrives at higher classification accuracy with increased pulse averaging. 3

Figure 2. A "clean" Kr isotope spectrum obtained from the noble gas instrument. 4

Figure 3. A Kr spectrum from the noble gas instrument showing significant levels of KrH^+ at m/z 85 and m/z 87. The results of the NN discrimination are shown as percent accuracy numbers for the ion pairs compared. All atomic ions showed nearly identical pulse signatures, while the polyatomic KrH^+ produced an easily distinguished pulse shape. 5

Tables

Table 1. Result of the NN discrimination of ion pulses produced for the Iodine NN discrimination data. Discrimination between two polyatomics or two atomic ions resulted in expected low discrimination accuracy at nominally 50%. Discrimination between ^{129}I and polyatomic 129 showed 95% accuracy. 6

Table 2. Ratio of uncertainty of the 1D-CNN method relative to conventional MS, as a function of polyatomic/atomic abundance ratio and classifier accuracy. 6

1.0 Introduction

In atomic mass spectrometry, polyatomic interferences that obscure ions of interest are a serious concern and in many cases hamper, or make impossible, the measurement of low signal analytes. Electron multipliers in pulse counting mode are used prolifically when low intensity ion beams are present during analysis as they provide on the order of 106-107 signal amplification, however simply counting the number of electron pulses ignores differences in the impact physics that are present depending on ion identity. The objective of this research was to investigate the efficacy of utilizing advanced neural network analysis and machine learning to recognize the differences in the analog electron multiplier pulse shapes for atomic vs. polyatomic ions of similar or identical mass as a means of distinguishing the two ion types. In this work, we've successfully discriminated polyatomic interferants from atomic analytes of interest for two different mass spectrometers, and three different ion pairs; Krypton isotope/hydride and methane/oxygen (noble gas mass spectrometer) and 129I atomic vs 129I polyatomic (Triton mass spectrometer). Discrimination accuracy for these experiments ranged from 76% to 95%.

2.0 Research Methods and Results

Pulse Shape Discrimination Approach

The primary neural network approach used to distinguish differences in the ion pair pulses was a 1D convolutional neural network (1D-CNN). Neural networks in general are groups of mathematical functions that, through training with known pulses, can self-optimize to produce improved output categorization from input data. 1D (dimensional) convolutional neural networks are particularly well-suited for analyzing sensor data with a temporal component, as is the case with electron multiplier pulses where the electron pulse if a function of intensity vs. time. The general approach to the analysis was as follows:

1. Identify atomic/polyatomic ion pair candidates where the identity of each ion was reasonably certain based on knowledge of the mass spectrometry experiment conducted. It was desirable to have these ions be as similar in mass as possible.
2. Collect approximately 50,000 pulses of each ion type with a fast, digital oscilloscope capable of a sampling rate of 10Gs/s. This acquisition rate equates to 500 data points for a 50 ns time window. Signals of interest have frequency components <1GHz so this sampling rate represents over-sampling of the relevant frequencies.
3. Train the neural network with 80% of the pulses
4. Test the effectiveness of the training with the remaining 20% of the pulses to arrive at a classification accuracy

As an addition to the 1D CNN approach for pulse shape discrimination, Random Forest and GradientBoosting models were built for both the methane/oxygen pulses as well as the Kr pulses. The Random Forest algorithm is a machine learning approach that uses decision trees built on bootstrap sampled training data. The tree maximum depth was 50 with 100 tree instances. The GradientBoosting algorithm builds a weighted ensemble of weak-learner classifiers incrementally by focusing on a residual. The maximum depth for boosting was five with 100 mini-tree instances using a learning rate of 0.05.

The general process for pulse evaluation/discrimination was similar to that used for 1D CNN approach, most typically 80% of the pulses collected were used for training, while the other 20% were used for evaluating the training success. During our experiments, we saw no significant difference between the accuracy of Random Forest, GradientBoosting, and 1D CNN approaches across a variety of experiments after conducting hyperparameter optimization.

Discrimination results

Two different mass spectrometers were used to produce the SEM pulses tested. The first instrument used is a home built noble gas magnetic sector mass spectrometer with an electron cyclotron resonance ion source and second is the Triton mass spectrometer by Thermo Scientific, which is a magnetic sector thermal ionization (TIMS) instrument. The results of three different ion groups will be shown below. The m/z 15-16 ion pair and the Krypton isotope comparisons were produced by the noble gas instrument while the Iodine pulses were produced by the Triton instrument. The classification accuracy of atomic vs. polyatomic pulses for these three examples ranged from 76% in the case of Krypton to 95% accuracy for ^{129}I atomic vs ^{129}I polyatomic. Results for methane (m/z 15) vs. oxygen (m/z 16) fell in between these

two at 84% accuracy. In figure 1 we demonstrated the substantial positive effect of averaging (2 or 4 ion pulses were averaged to produce training and challenge data sets) on classification accuracy, however since real-time application of this technique won't allow averaging, all subsequent efforts used single pulses.

Methane and Oxygen

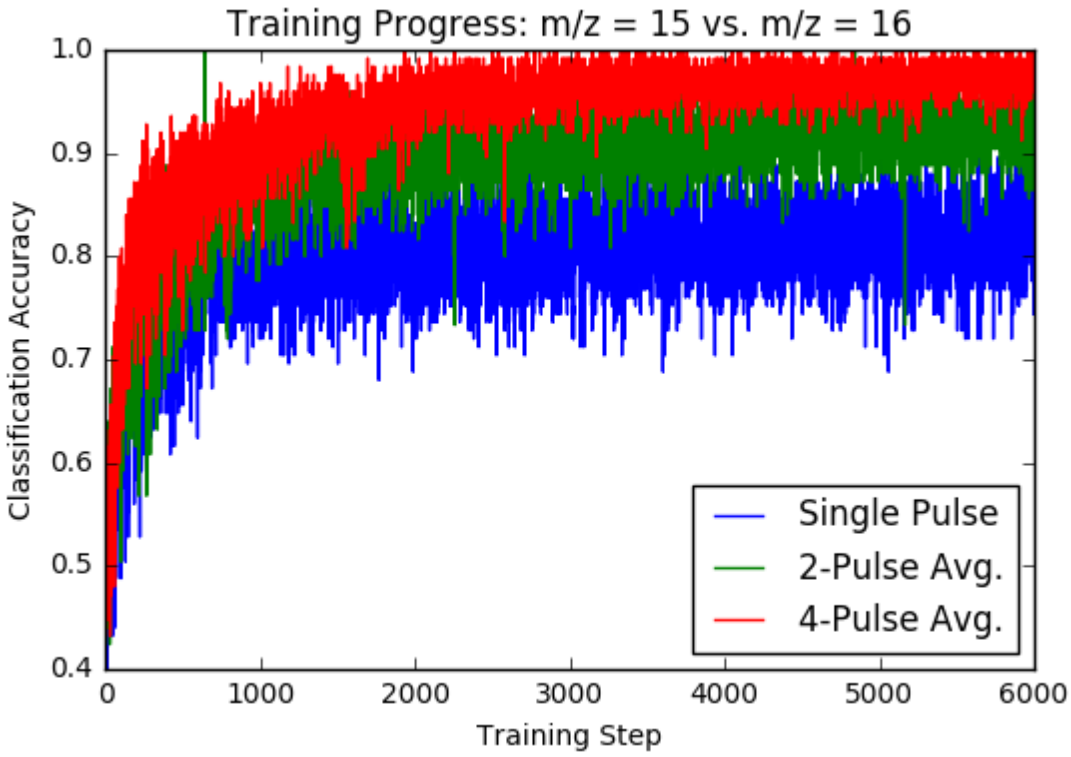

Figure 1. Plot of classification accuracy vs. training step for the m/z 15, m/z 16 (CH_3, O respectively) with different pulse averaging applied. The NN arrives at higher classification accuracy with increased pulse averaging.

M/z^+ 15 and 16 are always present in the noble gas instrument and are tentatively identified as methane and oxygen, respectively. This ion pair is a good candidate for NN discrimination testing as they are close in mass, likely to represent an atomic/polyatomic pair, and are low in mass. This latter attribute is important as all other experiments were performed with substantially higher mass. We've done many trials with these two ions and learned a lot about what parameters are important in these experiments such as detector warm up, signal intensity during pulse collection with the oscilloscope, the impact of pulse averaging, and the effect (or lack thereof) that detector gain, and oscilloscope trigger level has on the results. Given the ultimate importance of being able to successfully discriminate single pulses if this technique were to be employed to a real time pulse stream, analysis of averaged pulses was eliminated in all subsequent experiments. The final, best result of the methane, oxygen pulse analysis was 84% classification accuracy.

Krypton

The Argon gas that is used to generate the plasma for the noble gas instrument ion source contains small amounts of Krypton which has a very beneficial isotope fingerprint for these studies. A typical Kr spectrum with its major isotopes is shown below.

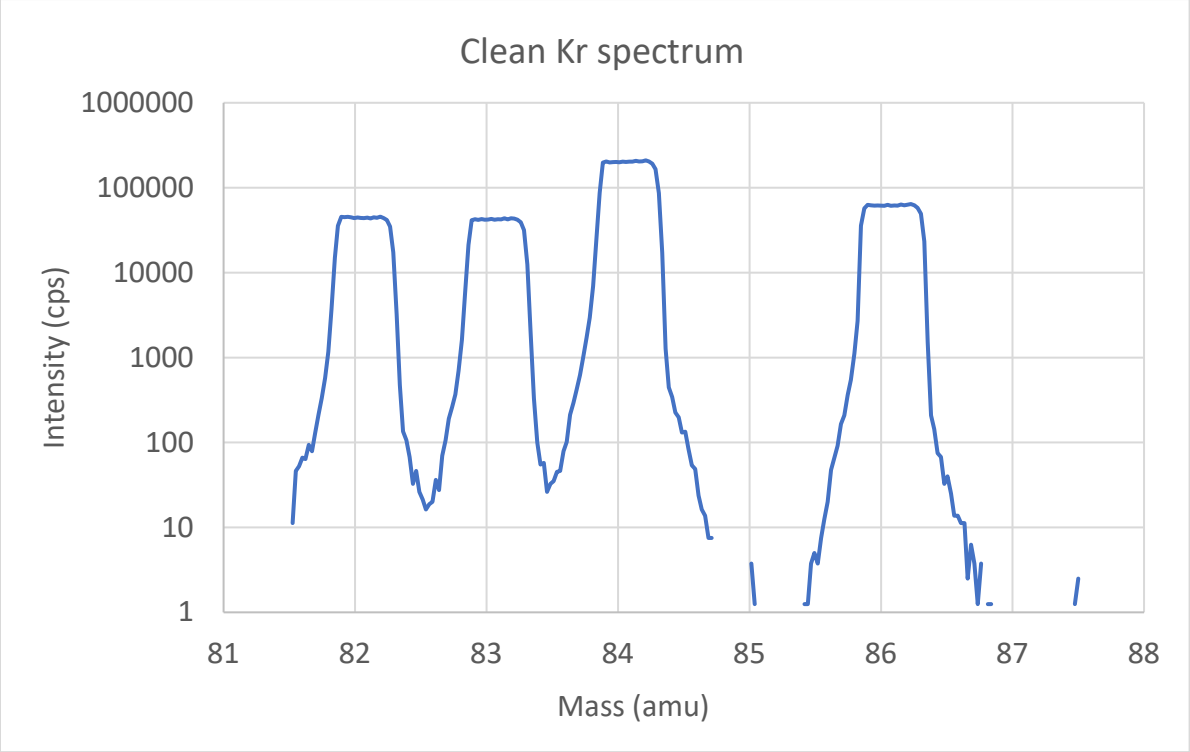

Figure 2. A "clean" Kr isotope spectrum obtained from the noble gas instrument.

When the instrument is dirty, such as after a vent to atmosphere, the diatomic hydride cations are formed in addition to atomic ions. In general, this instrument is too clean to show hydrides, but a recent servicing of the ion source resulted in an elevation of hydrides for a few weeks afterward. During this time period, EM pulses were collected from m/z 84 ($^{84}Kr^+$), 85 ($^{84}KrH^+$), and 86 ($^{86}Kr^+$). Previously, with the instrument operating at its typically clean, low hydride state, we had collected m/z 82, and 83 pulses. These two ions are very similar in that they are both atomic isotopes of Kr, one amu apart. Even though the m/z 82 and 83 discrimination was a separate experiment, the results are shown on a spectrum collected later while KrH^+ was present. The classification results for the various Kr isotopes and hydride are shown in Figure 3.

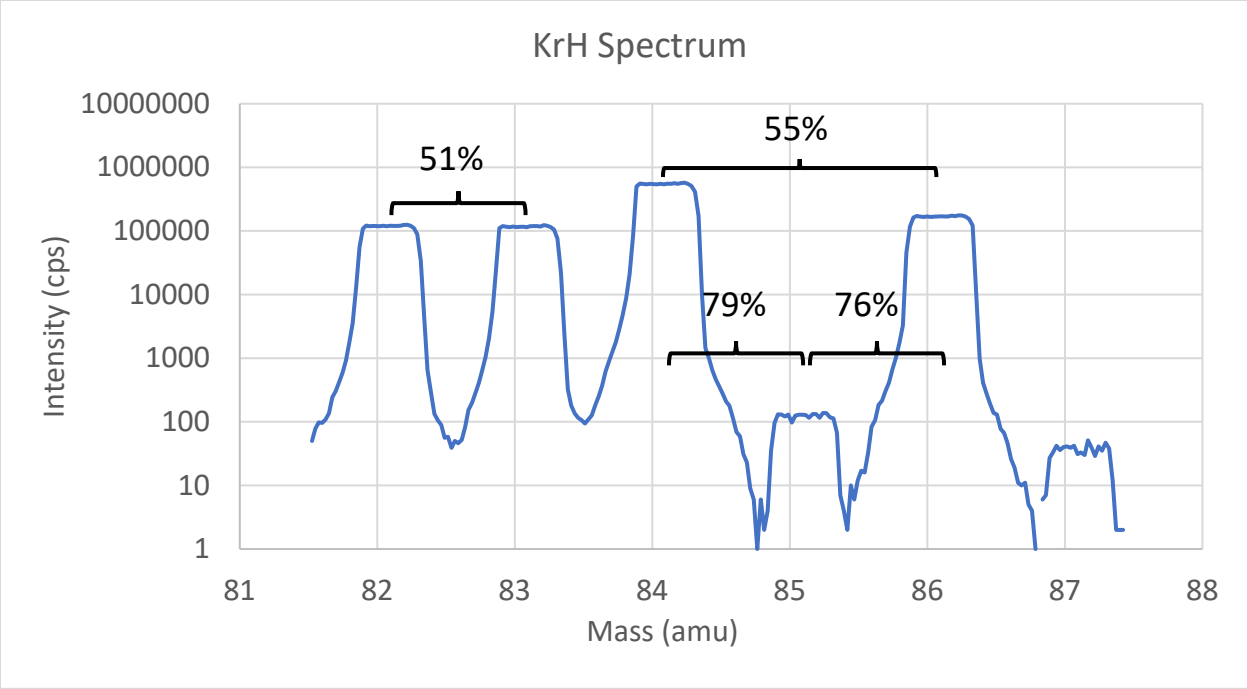

Figure 3. A Kr spectrum from the noble gas instrument showing significant levels of KrH^+ at m/z 85 and m/z 87. The results of the NN discrimination are shown as percent accuracy numbers for the ion pairs compared. All atomic ions showed nearly identical pulse signatures, while the polyatomic KrH^+ produced an easily distinguished pulse shape.

The interpretation of this data set is as follows:

M/z 82, 83, 84, and 86 are all atomic isotopes of Kr. Mass 85 is essentially pure $^{84}KrH^+$. There's an infinitesimally small amount of ^{85}Kr isotope but we can consider it absent, as shown in the first spectrum, as well as ~1ppm m/z 84 tail. All atomic isotopes show nearly 50% accuracy, meaning the NN result essentially did no better than a guess. This is as we would expect since they're all atomic cations of the same element. The only classification that indicated a significantly different pulse signature was the diatomic $^{84}KrH^+$.

Iodine Results

Thermal ionization mass spectrometry (TIMS) is a common tool used for iodine isotope analysis for the purpose of environmental monitoring. Two versions of Thermo's Triton mass spectrometer are in use at PNNL; the Triton and the Triton plus. For these experiments, the Triton was used to produce pulses of both Iodine (127 and 129) as well as a known polyatomic interferent at m/z^- 129. In addition, a suspected polyatomic ion at m/z^- 128 that has been observed to track with the polyatomic m/z^- 129 was collected. These are all negative ions (as opposed to all other ions studied in this E-ion work). Isotope ratios relative to ^{127}I confirmed the identity of the ^{129}I ion as well as that of the interferent at m/z^- 129. The primary result of interest was the comparison of ^{129}I to m/z^- 129 polyatomic, however discrimination between other combinations of the four ions collected for these experiments was also performed with the results shown below in Table 1.

Table 1. Result of the NN discrimination of ion pulses produced for the Iodine NN discrimination data. Discrimination between two polyatomics or two atomic ions resulted in expected low discrimination accuracy at nominally 50%. Discrimination between ^{129}I and polyatomic 129 showed 95% accuracy.

Comparison	Result (% discrimination accuracy)
^{127}I vs. ^{129}I	55%
Polyatomic 128 vs. Polyatomic 129	54%
^{129}I vs. Polyatomic 129	95%

The most important comparison for the Iodine study, ^{129}I vs. Polyatomic 129, showed the most impressive result to date for the NN discrimination at 95%. Neither visual examination of the pulses nor any other experimental details revealed an obvious reason for the exceedingly good result found. It is possible that the lower energy spread of ions produced by the TIMS (Triton) ion source relative to the noble gas instrument played a role. It should be noted that the Iodine experiment on the Triton instrument was only performed once, and more trials would be needed before the very high classification result of 95% could be reported with confidence.

Statistical treatment of E-ion Results

An effort to quantify the uncertainty in the estimated total number of analyte ions based on E-ion classification accuracy led to the development of a probabilistic model. This model provided a formula for a best estimate of the quantity of analyte and a basis upon which to compare the uncertainty in this estimate to that obtained in two alternatives: "perfect classification," a purely hypothetical case in which one knows with certainty which ions are atomic and which are polyatomic, and an alternative using a common correction based on a calibration sample. One of the useful results was to show how much greater the uncertainty in the E-ion estimate would be relative to hypothetical "perfect classification." In particular, the table below indicates how that uncertainty changes as a function of classifier accuracy and polyatomic/atomic ratio.

Magnitude of Uncertainty in the Number of Analyte Ions Relative to "Perfect"

Table 2. Ratio of uncertainty of the 1D-CNN method relative to conventional MS, as a function of polyatomic/atomic abundance ratio and classifier accuracy.

Polyatomic / Atomic Ratio	Classifier Accuracy		
	60%	75%	90%
1	3.6	1.6	1.1
10	8.2	3.0	1.6
100	24.6	8.8	3.9
1000	77.5	27.4	11.9

The trend shows, as expected, that the uncertainty of the estimated total number of analyte ions increases with interferent/atomic ion ratio and lower classifier accuracy.

The other very important result of this work was a theoretical result indicating that, for polyatomic/atomic ratios of about 10:1 or greater, only ~**73%** classification accuracy is needed to exceed the precision of the common mass spectrometry correction method. Further investigation of this theoretical result is needed in consideration of typical results in the application of this type of correction. There is some heavy math going on with these calculations, and we'll need to make sure this is an accurate assessment, but it's very promising and would change the early classification accuracy target of ~95% first suggested.

3.0 Summary

In summary, we have demonstrated that with two different examples of neural network pulse shape discrimination, 1D convolutional neural network, Random Forest and GradientBoosting algorithms, we can distinguish between atomic and polyatomic ion pulses of either the same mass (Iodine), or with very similar masses (Kr isotopes and KrH; Methane/Oxygen). The degree to which we can distinguish these pulses, reported as discrimination accuracy, varied between 76% and 95% for the ion pairs presented here. These results are extremely encouraging with respect to the minimum threshold required to exceed the ion identity uncertainty, as determined by a statistical treatment of E-ion discrimination compared to classic correction methods, of 73%. Further exploration of the threshold for various analytical scenarios (ratios of atomic : polyatomic; method for classically correcting the interference) is needed, but it is encouraging that even the lowest classification accuracy obtained (76%) compares well with our initial estimate of such threshold accuracy.

Pacific Northwest National Laboratory

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99354

1-888-375-PNNL (7665)

www.pnnl.gov