analytical chemistry

Article

pubs.acs.org/ac

Tile-Based Fisher Ratio Analysis of Comprehensive Two-Dimensional Gas Chromatography Time-of-Flight Mass Spectrometry (GC × GC–TOFMS) Data Using a Null Distribution Approach

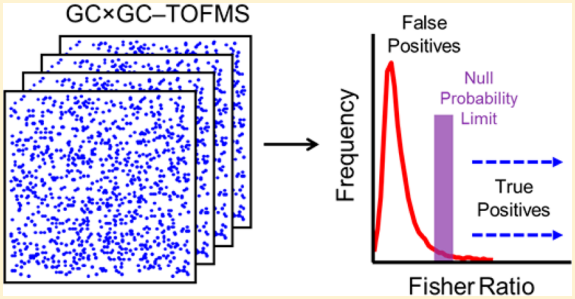
Brendon A. Parsons,[†] Luke C. Marney,[†] W. Christopher Siegler,[†,#] Jamin C. Hoggard,[†] Bob W. Wright,[‡] and Robert E. Synovec*[,†]

[†]Department of Chemistry, University of Washington, Box 351700, Seattle, Washington 98198, United States

[‡]Pacific Northwest National Laboratory, Battelle Boulevard, P.O. Box 999, Richland, Washington 99352, United States

Supporting Information

ABSTRACT: Comprehensive two-dimensional (2D) gas chromatography coupled with time-of-flight mass spectrometry (GC × GC–TOFMS) is a versatile instrumental platform capable of collecting highly informative, yet highly complex, chemical data for a variety of samples. Fisher-ratio (F-ratio) analysis applied to the supervised comparison of sample classes algorithmically reduces complex GC × GC–TOFMS data sets to find class distinguishing chemical features. F-ratio analysis, using a tile-based algorithm, significantly reduces the adverse effects of chromatographic misalignment and spurious covariance of the detected signal, enhancing the discovery of true positives while simultaneously reducing the likelihood of detecting false positives. Herein, we report a study using tile-based F-ratio analysis whereby four non-native analytes were spiked into diesel fuel at several concentrations ranging from 0 to 100 ppm. Spike level comparisons were performed in two regimes: comparing the spiked samples to the nonspiked fuel matrix and to each other at relative concentration factors of two. Redundant hits were algorithmically removed by refocusing the tiled results onto the original high resolution pixel level data. To objectively limit the tile-based F-ratio results to only features which are statistically likely to be true positives, we developed a combinatorial technique using null class comparisons, called null distribution analysis, by which we determined a statistically defensible F-ratio cutoff for the analysis of the hit list. After applying null distribution analysis, spiked analytes were reliably discovered at ~1 to ~10 ppm (~5 to ~50 pg using a 200:1 split), depending upon the degree of mass spectral selectivity and 2D chromatographic resolution, with minimal occurrence of false positives. To place the relevance of this work among other methods in this field, results are compared to those for pixel and peak table-based approaches.



Two-dimensional (2D) gas chromatography coupled with time-of-flight mass spectrometry (GC × GC–TOFMS) is a prominent instrumental platform for the study of complex samples for analytes either sufficiently volatile or amenable to analysis after derivatization.[1−8] Use of GC × GC–TOFMS and associated data analysis strategies aim to uncover meaningful chemical information from complex samples. However, meaningful chemical information is often buried in a background of less meaningful chemical signal and noise. Further complicating the effort, experiments often require analysis of replicate injections of different samples, increasing the data dimensionality. Feature selection software becomes increasingly important to extract the most meaningful chemical information. Ultimately, automatable software is needed to comprehensively analyze GC × GC–TOFMS data so important analytes and/or chemical fingerprints can be quickly and accurately ascertained during discovery-based experimentation.

Our focus is on nontargeted discovery-based GC × GC–TOFMS investigations, whereby chemical features of interest are not presumed known beforehand. Nontargeted approaches can be either supervised or unsupervised, where supervision refers to either external calibration or prior classification of samples/chromatograms as they relate to the experimental design.[9−11] Nontargeted analyses are typically performed by software comparing the chromatograms on a pixel, peak table, or peak region basis. These approaches have been reviewed in the context of nontargeted analysis of GC × GC data.[12] Specific approaches to GC × GC data processing include watershed algorithm[13] and template matching[14] methods, chromatographic region-based binning,[15] chromatographic and mass spectral inspection,[16] and scripting-based classification.[17] We recently reported an approach to GC × GC data processing, referred to as tile-based analysis, in which data is strategically binned, with the tile size based upon the 2D peak widths and the retention time shifting that may be present.[18] Recent

Received: December 1, 2014
Accepted: March 18, 2015
Published: March 18, 2015

DOI: 10.1021/ac504472s
Anal. Chem. 2015, 87, 3812−3819

studies have reported using a combination of multiple approaches (peak table, pixel, and peak region based).[19]

Pixel-based analysis compares every GC × GC−TOFMS data point in the three-dimensional space (3D), i.e., pixel, in each chromatogram with relatively limited preprocessing (beyond baseline correction and normalization steps). Since chromatographic misalignment may lead to false positives and the diminished ability to discover true positives, 2D alignment[20−25] is often required to reduce misalignment but may distort peak shapes and areas in the pixel level data. Further, the typical GC × GC parameters applied (a low number of modulations per first dimension peak width, e.g., a modulation ratio of $M_R \sim 2$ to 4)[26] limit the data density defining first dimension peaks due to the need to balance the peak capacities in both GC × GC dimensions; this reduces the information available for first dimension alignment.[27] On the other hand, peak-based approaches utilize baseline correction, peak deconvolution and identification, integration, and arrangement of every peak in a single chromatogram into a table (i.e., peak table), and tables are then compared across samples. Tile-based analysis[18] addresses the challenges due to misalignment of multiple GC × GC−TOFMS chromatograms without explicitly aligning the data, thus avoiding chromatographic and/or mass spectral warping or other distortions of the raw data. Tile-based analysis combines advantages of peak table-based approaches (including mitigation of misalignment and improvement in signal-to-noise) and the advantages of pixel-based approaches (remaining unbiased to chromatographic peak shape or mass spectral signal characteristics).

After the data have been prepared by either pixel, tile, or peak table-based approaches, supervised statistical analysis can be performed to discover chemical features that distinguish the sample classes, producing lists of significant features referred to as hit lists. Two popular supervised methods include Fisher-ratio (F-ratio) analysis and partial least-squares discriminant analysis (PLS-DA). F-ratio analysis has been applied to GC × GC−TOFMS data previously at the pixel level[28−33] and at the peak table level using LECO ChromaTOF Statistical Compare.[34−37] PLS-DA has also been applied to peak table-based GC × GC−TOFMS data.[38−41] For nontargeted analysis, false positives should be minimized while maximizing the sensitivity with which true positives are detected. False positives are the rejection of the null hypothesis when it is actually true, often referred to as Type I errors. For F-ratio analysis, a false positive is the discovery of a feature that does not chemically distinguish the sample classes. While a true positive is a feature that is chemically class distinguishing (e.g., a peak signal inferring a statistically significant difference in concentration between two or more sample classes), a false positive may be due to either spurious covariance of detector fluctuations or chromatographic misalignment. Tile-based F-ratio analysis significantly reduces the number of false positives found with pixel-based analysis of GC × GC−TOFMS data that were associated with the aforementioned causes.[18] While tile-based F-ratio analysis reduces the frequency and rank of false positives in the hit list, false positives are still present in the analysis simply due to the large number of statistical hypotheses being simultaneously tested in the entire tiled GC × GC−TOFMS chromatogram. This is encountered in multiple hypothesis testing and is commonly described by the false discovery rate (FDR).[42]

For efficient and informative nontargeted F-ratio analysis, the relative distributions of true and false positives must be understood so the analysis can focus on true positives. This is typically performed in a subjective manner by selecting an F-ratio cutoff in the hit list at the point where the selected features become unreliable in differentiating the sample classes.[29,36] However, manual selection of a cutoff comes with two complications. First, the cutoff must be determined by working through the ranked hit list and validating the results to determine the point at which the analysis becomes unreliable. This is problematic as it requires time-consuming analysis of features with little value and because the cutoff must be determined in a subjective manner, as true positives are often interspersed with false positives as one works down the hit list. The second concern is that this exercise must be repeated for every experiment, since sampling and instrumental variance may differ from one experimental campaign to the next.

Ideally, an objective metric should be implemented that expresses the confidence by which a given feature can be deemed a true positive. Herein, we critically study utilization of false positive distribution estimations, called "null distributions", which ultimately allows for a statistically-based determination of whether a discovered class distinguishing feature is likely to be either a true positive or a false positive. We utilize tile-based F-ratio analysis in concert with the standard addition method by spiking non-native chemicals into a diesel fuel matrix at low concentrations. While the previous work studied the concentration range of 100−1000 ppm,[18] the current study focuses on the 0−100 ppm analyte spike range. The current study demonstrates the sensitivity and selectivity of tile-based F-ratio analysis for discovery of true positives in the nontargeted analysis of a chemically complex and analytically challenging sample matrix. Additionally, we demonstrate the utility of null distribution analysis to defensibly limit the resulting hit list to only the features most likely to be class distinguishing and, ultimately, useful to the analyst. By exploring the low concentration spike levels, we gain a better understanding of the "discovery" limit of detection (LOD) of tile-based F-ratio analysis and compare its merits to pixel-based and peak table-based approaches.

EXPERIMENTAL SECTION

Four chemical compounds were spiked at the following nominal concentrations into an ultralow sulfur diesel (ULSD) fuel: 100, 50, 25, 12.5, 6.2, 3.2, 1.6, and 0 parts-per-million by mass (ppm). The 0 ppm spike level served as the matrix blank. The four spiked compounds were bromobenzene, 1-chlorohexane, 5-decyne, and 3-octanone, none of which was naturally present in the fuel, as verified by the standard addition method. An internal standard, 1-bromoheptane, was also spiked into each sample at 1 part-per-thousand (ppt). All samples were prepared gravimetrically using a 5-place analytical balance, and the actual (not nominal) concentration for each analyte at each spike level is provided in Table S-1 (Supporting Information, Section S1.1). Additionally, the calculated mass of each analyte injected on the column is provided in Table S-2, Supporting Information. The actual concentrations of the spiked analytes are used for software evaluation. However, for clarity, we use the nominal spike concentrations when referring to a particular concentration comparison with multiple analytes. The samples were analyzed by GC × GC−TOFMS, as in the previous report.[18] Under the selected instrumental conditions, the modulation ratio was 3 to 4, depending on the first dimension peak width.[26] Full instrument parameters and a representative

DOI: 10.1021/ac504472s
Anal. Chem. 2015, 87, 3812−3819

chromatogram are provided in the Supporting Information (Section S1.2, Figure S-1).

GC × GC−TOFMS data from all 36 runs (7 spike levels × 4 replicates, and 8 replicates of the 0 ppm) were imported from the LECO ChromaTOF software v 3.32 (LECO, St. Joseph, MI) to Matlab v 8.0.0.783 (MathWorks, Inc., Natick, MA) via an in-house developed data converter.[43] The imported data were analyzed with the in-house developed tile-based F-ratio software, performed on a midlevel personal computer, having an Intel Core i7-4770 processor (3.4 GHz), a 256 GB Samsung 840 solid state hard drive, and 16 GB dual-channel DDR3 RAM. An in-house developed PARAFAC GUI was used to measure the quantitative signal volume for the internal standard, 1-bromoheptane, for each sample injection, which was used to normalize the data to account for injection variation.[9] The F-ratio was calculated as the class-to-class variation of the detected signal divided by the sum of the within-class variations of the signal.[28,44,45] Details of the F-ratio calculation and a summary of the preprocessing steps are included in the Supporting Information (Sections S2.1 and S2.2).

The GC × GC pixel level data was summed using a novel method of binning, the tile method.[18] Details on the tile method and its advantages are available in the Supporting Information (Section S2.3 and Figure S-2). Redundant hits were removed using a novel "pin and cluster" algorithm, detailed in the Supporting Information (Section 2.4 and Figure S-3 and Tables S-3 and S-4). The complete tile-based F-ratio software is summarized by a flowchart in the Supporting Information, Figure S-4. The specific parameters for the tile method and redundant hit removal are included in the Supporting Information (Section S1.3). The 2D chromatographic features discovered by tile-based analysis are arranged in a ranked table sorted by decreasing average F-ratio, i.e., a "hit list", an example of which is provided in Table 1. To compare

Table 1. Hit List for the 12.5 ppm versus 6.2 ppm Comparison after Redundant Hit Removal[a]

F-ratio hit number	average F-ratio	^1t_R (s)	^2t_R (s)	null classification	compound
1	214.4	360	0.13	hit	bromobenzene
2	114.5	255	0.93	hit	1-chlorohexane
3	60.1	534	0.73	hit	5-decyne
4	14.8	945	0.52	potential hit	
5	11.6	1202	0.61	potential hit	
6	11.2	441	0.91	potential hit	3-octanone
7	9.9	973	0.51	nonhit	
8−5033				nonhits	

[a]There are a total of 5033 entries in the list. The four spiked analytes are found within the first six entries. There are two potential hits interspersed with the features for the spiked analytes. The corresponding hit list prior to redundant hit removal may be found in the Supporting Information (Table S-5).

the performance of the tile-based F-ratio analysis with other relevant methods, the data were also processed by pixel-based and peak table-based methods, using in-house software and LECO ChromaTOF v 3.32, respectively. Details of the pixel and peak table analyses are provided in the Supporting Information (Section S4).

Following the generation of hit lists by the tile-based F-ratio analysis, null distribution analysis was implemented to control the false discovery rate by determining the probability that a discovered feature is a false positive. Due to the large number of statistical hypotheses being simultaneously tested in the tiled data, a portion of the statistical tests will falsely reject the null hypothesis when it is true, i.e., false positives (Type I errors).[46−53] The large number of simultaneous hypotheses tested in the tiled data were leveraged to determine the distribution of potential false positives. These distributions were utilized to limit the analysis of the hit list to only the range that is likely to have few or no false positives, using an objective metric as we demonstrate herein. False positive distributions can be determined by repeated testing of null hypotheses (i.e., that there is no difference between sample classes), created either by comparing replicates of a single class or by evenly mixing the sample classes to eliminate between-class variation. In this study, such comparisons of GC × GC−TOFMS data of the diesel samples were subjected to tile-based analysis to generate null comparison hit lists, from which the F-ratio values are used to make a histogram. Because the comparisons are testing the null hypothesis, we refer to the resulting curves as "null distributions".

RESULTS AND DISCUSSION

Null comparisons were performed on the basis of a pairwise rearrangement (switching of two samples at a time from each class, keeping balanced classes) of the eight injection replicates of the 0 ppm diesel fuel that served as the matrix blank. Six null class comparisons were deemed sufficient to capture the variation in the null comparisons, as different combinations of the injection replicates ultimately resulted in only slightly different distributions of null F-ratios. This approach is a form of combinatorial analysis.[54] Each of these null comparisons was subjected to tile-based F-ratio analysis in order to generate a total of six null hit lists. Null distributions (see Figure 1A) were then prepared from the hit lists. The F-ratio values for each feature from the null hit lists were then combined into a histogram with a bin size of 0.2. The average null distribution (black trace) has an average of 4860 ± 77 features. The range of null distributions encountered is also provided in Figure 1A by the blue and red traces. The inset plot in Figure 1A is a zoomed section, emphasizing the range of the null distribution tails at higher F-ratios. This region is sparse due to the infrequent occurrence of values greater than ∼10 in the null distributions, and it follows that in true class comparisons (discussed shortly) there may be occasional false positives interspersed with the true positives at values that have some overlap with the null distribution. It is this region of the null distribution that relates to the discovery LOD for the tile-based F-ratio analysis that we explored.

Transforming a null distribution into a "null probability" curve facilitates limiting a hit list to features that are likely to be either a false positive or a true positive. The null probability provides the quantitative likelihood that a given feature is indeed a false positive, based on its average F-ratio. This calculation is performed by dividing the summed number of null features above a given value by the total number of null features in the entire null distribution, multiplied by 100 to yield a percentage. The null probability is equivalent to the percentage area of the null distribution that is above a given value, which represents the proportion of the null features that would be included in the analysis with a given F-ratio cutoff. This calculation is performed for the range of values in the null distribution to create a curve which is the null probability versus average F-ratio. Figure 1B is the null probability curve for the

DOI: 10.1021/ac504472s
Anal. Chem. 2015, 87, 3812−3819

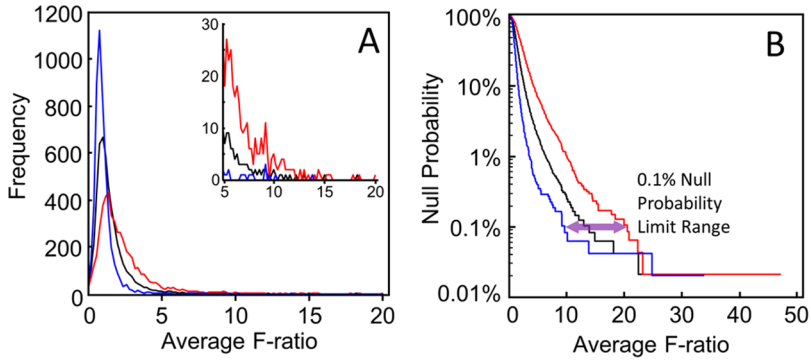


Figure 1. (A) Average and the two extreme null distributions from the null comparisons of the 0 ppm versus 0 ppm diesel fuel sample (i.e., the unspiked matrix blank) are shown. The frequency plotted is the number of features within a bin interval of 0.2 per the average F-ratio. Six rearrangements produced six different null distributions, and others were omitted for clarity. The range of the null distributions is represented by the red and blue null distributions. The zoomed section shows the range of the rightmost tail of the null distributions that were encountered. (B) The average null probability curve is shown on the basis of the average F-ratio null distribution (the black trace) in (A), along with the null probability curves resulting from the two extreme null distributions presented in (A). The null probability limit range of ∼0.1% is indicated by the purple double arrow (∼equivalent to 5 spurious features in a hit list), suggesting the use of an average F-ratio cutoff of ∼10 to 20 for the "discovery" LOD.

average null distribution (the black trace), as well as the range of individual null distributions shown in Figure 1A, the blue trace and red trace. The horizontal range between the blue and red traces gives rise to a null probability range, which expresses the variation in the average F-ratio which would be selected for various null probability limits. If one were to limit a comparison to features with less than 0.1% probability that the feature is a false positive (this would be equivalent to accepting ∼5 spurious features in the analysis, based on the 4860 features in the average null distribution for this data set), this would correspond to a threshold of ∼14, such that only features with average F-ratios above 14 would be inspected. However, since there is variation in the distribution of false positives in the null comparisons, as shown in Figure 1A, it is necessary to designate a range for an F-ratio limit at a particular null probability. Accordingly, as illustrated in Figure 1B, the range for the 0.1% null probability limit is ∼10 to 20. On the basis of the 0.1% null probability limit, an average F-ratio at or above 20 is referred to as a "hit" while a value ranging from 10 to 20 is referred to as a "potential hit". A value below 10 is referred to as a "nonhit". Therefore, the analyst should work from the top of the hit list, since "hits" have the highest likelihood of being true positives. As defined in Figure 1, the hit lists in this report include a null classification, whereby the features were assigned as hit, potential hit, or nonhit based on their F-ratio relative to the 0.1% null probability limit.

Tile-based F-ratio analyses for true class comparisons were performed in two regimes: spike versus blank comparisons (e.g., 6.2 ppm versus 0 ppm) and concentration ratio of 2 comparisons (e.g., 25 ppm versus 12.5 ppm). The spike versus blank comparisons test tile-based F-ratio analysis in cases where a low abundance analyte is present in one sample class and absent from the other. The concentration ratio of two comparisons tests cases where a low abundance analyte is different in concentration by a factor of 2 between the sample classes. Each of these cases is important for the performance of a nontargeted analysis platform. For these true class comparisons, and analogous to the null distributions in Figure 1A, *false positive* distributions were prepared and summarized in Figure 2. Note that the features corresponding to the four true positives (i.e., the spiked analytes) in each comparison were removed in the preparation of Figure 2. The resulting

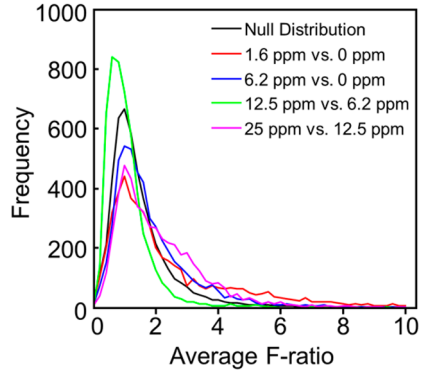


Figure 2. False positive distributions from four representative F-ratio true class comparisons are illustrated in the context of the average null distribution from the six null comparisons derived from the comparisons of the 0 ppm diesel fuel sample (black curve, see Figure 1A). The frequency plotted is the number of features within a bin interval of 0.2 per the average F-ratio.

distributions have an average of 4944 ± 141 false positives. Figure 2 shows false positive distributions from four comparisons representing the range of false positive F-ratio values encountered. The average null distribution from Figure 1A is also plotted in Figure 2 for comparison. As in Figure 1A, in Figure 2, the F-ratios for the false positives from each comparison were made into histograms using a bin size of 0.2. The range of false positive distributions indicates there is some modest variation in the shape of the distributions, though nearly all of the false positives have F-ratios below 10, which strongly supports the implementation of null distribution data to determine the relationship between false positives and their corresponding F-ratio, as delineated in the null probability curve (Figure 1B). Indeed, the null distributions in Figure 1A are of similar range to the false positive distributions in Figure 2, and the average null distribution is essentially the average of the false positive distributions. In this study, we have taken advantage of knowing the true positives (the spiked analytes) in order to critically compare the underlying false positive distributions in the true class comparisons with the null distributions prepared from replicates of the nonspiked diesel fuel (0 ppm matrix blank). For analysis of comparisons where the true positives are not known, false positive distributions

DOI: 10.1021/ac504472s
Anal. Chem. 2015, 87, 3812−3819

cannot be determined; instead, null distribution analysis is needed to estimate the underlying distribution of false positives and to illustrate the discovery LOD for tile-based F-ratio analysis, as is described next.

The true class comparisons, in the context of the 0.1% probability limit presented in Figure 1B, are presented in Figure 3A,B for the spike versus blank comparisons and the concentration ratio of two comparisons, respectively. The plots show the average F-ratios for each of the four spiked analytes versus their actual concentration (in ppm) in the comparisons. In the concentration ratio of two plots, the concentration plotted is that of the lower concentration spike in the comparison. The 0.1% null probability limit range (illustrated in Figure 1B) is included in the figure to show whether analytes would be found if the 0.1% null probability limit was applied. Features above the 0.1% null probability limit (hits) would be found in the course of the analysis of the hit list, while those below the limit (nonhits) would not, unless the analyst is willing to look through a larger number of false positives. Analytes within the 0.1% probability limit (potential hits) often may be readily found, depending on the distribution of false positives in the particular comparison, and the extent to which the hit list is examined.

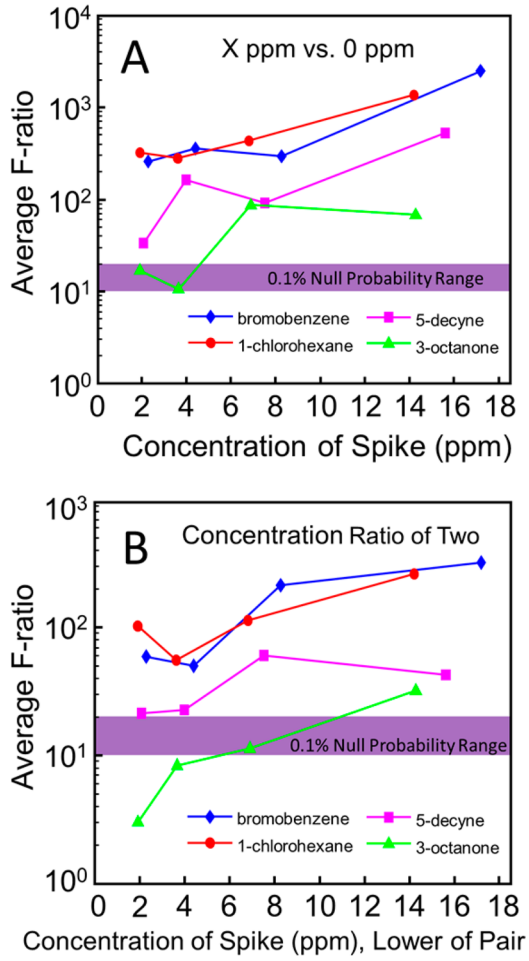


Figure 3. (A) Average F-ratios for the four spiked analytes are shown for the spike versus matrix blank comparisons (e.g., 12.5 ppm versus 0 ppm) in the context of the 0.1% null probability limit (Figure 1B). The actual concentration in ppm of the spiked sample in each pair for each spiked analyte is plotted (see Table 1). (B) Average F-ratios for the four spiked analytes are shown for the nominal concentration ratio of two comparisons (e.g., 25 ppm versus 12.5 ppm) in the context of the 0.1% null probability limit (Figure 1B). The actual concentration in ppm (see Table 1) of the lower spike level of each pair for each spiked analyte is plotted.

Most of the spiked analytes were consistently "found" above the 0.1% probability limit for the range of concentration comparisons studied, even for the more analytically challenging concentration ratio of two comparisons. Nonetheless, there were sufficient cases that facilitated the study of the discovery LOD, which is taken as an average F-ratio that falls below the 0.1% probability limit. Bromobenzene and 1-chlorohexane have consistently high average F-ratios for all concentration comparisons, due to having selective m/z and limited chromatographic interference. 5-Decyne and 3-octanone have fewer selective m/z and more chromatographic interference, which results in lower average F-ratios for these analytes. In the spike versus blank comparisons, each of the four spiked analytes has a high average F-ratio at nearly all concentration levels. For the comparisons at or above 6.2 ppm versus 0 ppm, all four spiked analytes have average F-ratios greater than the 0.1% null probability limit, indicating that these hits have little to no overlap with the null distribution. For the 3.2 ppm versus 0 ppm and the 1.6 ppm versus 0 ppm comparisons, 3-octanone has an average F-ratio which falls within the 0.1% null probability limit, indicating that these features for 3-octanone are likely to be interspersed with a small number of false positives.

The concentration ratio of two comparisons is more challenging, has a lower overall F-ratio, and thus tends to overlap more with the null distribution. For the 25 ppm versus 12.5 ppm comparison, all four spiked analytes have average F-ratios greater than the 0.1% null probability limit. At lower absolute concentrations, 3-octanone has a diminished average F-ratio. For the 12.5 ppm versus 6.2 ppm comparison, 3-octanone is in the 0.1% null probability limit, indicating the 3-octanone feature is likely to be interspersed with a small number of false positives. Table 1 is the abbreviated (first 7 entries) hit list for the 12.5 ppm versus 6.2 ppm comparison, in which 3-octanone is within the 0.1% null probability limit, along with two other potential hits that were not spiked analytes. As the absolute concentration further decreases, the average F-ratio for 3-octanone likewise falls, increasing the number of false positives that will be encountered while working down the hit list toward the 3-octanone feature. This is demonstrated in Table 2, which is the abbreviated hit list for the 3.2 ppm versus 1.6 ppm comparison. At this lower absolute

Table 2. Hit List for the 3.2 ppm versus 1.6 ppm Comparison after Redundant Hit Removal[a]

F-ratio hit number	average F-ratio	^1t_R (s)	^2t_R (s)	null classification	compound
1	102.2	255	0.93	hit	1-chlorohexane
2	59.3	360	0.13	hit	bromobenzene
3	25.8	2150	0.42	hit	false positive
4	21.4	534	0.73	hit	5-decyne
5–220				potential hits and nonhits	
221	3.0	441	0.92	nonhit	3-octanone
222–5099				nonhits	

[a]There are a total of 5099 entries in the list.

DOI: 10.1021/ac504472s
Anal. Chem. 2015, 87, 3812–3819

concentration, the F-ratios for the true positives have all dropped. There is now a false positive hit interspersed between the hits for bromobenzene and 5-decyne and 216 potential hits and nonhits between the features for 5-decyne and 3-octanone. Due to lower mass spectral selectivity and chromatographic interference from the diesel fuel matrix, 3-octanone is found below the 0.1% null probability limit at the concentration ratio of two comparisons below the 12.5 ppm versus 6.2 ppm comparison and is thus assigned as a nonhit. The remaining three spiked analytes have F-ratios greater than the 0.1% null probability limit for the concentration ratio of two comparisons, though 5-decyne approaches the 0.1% null probability limit for comparisons below 12.5 ppm versus 6.2 ppm. Additional abbreviated hit lists showing the relative ranks for the spiked analytes and false positives, and a discussion thereof, are included in the Supporting Information (Section S3, Tables S-6 and S-7). To summarize, spiked analytes were reliably discovered at ∼1 to ∼10 ppm (∼5 to ∼50 pg on column using a 200:1 split), depending on the degree of mass spectral selectivity and 2D chromatographic resolution from the diesel fuel matrix.

Figure 4 illustrates the challenge successfully addressed by tile-based F-ratio analysis by examining the limited selectivity at

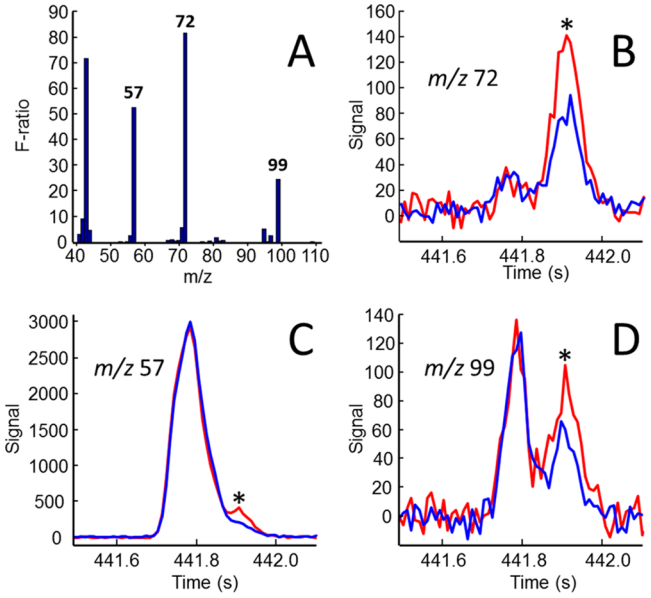


Figure 4. 3-Octanone in the 12.5 ppm spike (red) versus 6.2 ppm spike (blue) comparison. (A) The F-ratio "mass spectrum". (B–D) The three most selective m/z values per F-ratio analysis are plotted from one sample injection from each class. The plots are extracted ion chromatograms (XIC) from the column 2 separation for the most intense modulation of the 3-octanone peak. The rightmost peak is 3-octanone (denoted by ∗), while the leftmost peak is comprised of several interfering compounds from the diesel fuel matrix.

which 3-octanone is discovered in the 12.5 ppm versus 6.2 ppm comparison. Figure 4A is the F-ratio spectrum (the F-ratio for each m/z that was present in the tile above the S/N threshold) for the feature corresponding to 3-octanone. The m/z plotted were averaged to yield the average F-ratio, by which the hit list is currently ranked. The m/z 57, 72, and 99 are the most selective versus the significant matrix peaks, which overlap the 3-octanone peak at low chromatographic resolution in both separation dimensions. There is potential to improve ranking by selecting a subset of m/z for the calculation of average F-ratios, requiring further study. Figure 4B–D includes segments of the column 2 separation from the most abundant modulation of 3-octanone, for the extracted ion chromatograms (XIC) at m/z 57, 72, and 99. These chromatograms illustrate both the low absolute abundance of 3-octanone at this concentration level and the limited chromatographic resolution and mass spectral selectivity versus the interfering matrix peak. Due to low S/N and low mass spectral selectivity, 3-octanone has a moderate average F-ratio at the 12.5 ppm versus 6.2 ppm comparison level and is found within the 0.1% null probability limit, in which two other potential hits are interspersed. The discovered feature for 3-octanone in the 6.2 ppm versus 0 ppm comparison is detailed in a similar fashion in Figure S-5 (Supporting Information, Section S3.1).

We also studied the performance of the tile-based F-ratio method in the context of the pixel-based and peak table-based analyses (see Supporting Information, Section S4). The pixel-based method (see Section S4.1, Supporting Information, for details) performs similarly to the tile-based method, particularly at the higher absolute concentration comparisons (at or above the 50 ppm nominal concentration). However, at lower concentration, the pixel-based method encounters more false positives interspersed between the spiked analytes. Likewise, the peak table comparison is provided in Section S4.2, Supporting Information. Because the peak table processing is highly influenced by the selected parameters,[55] we carefully optimized the peak table method to discover the low concentration analytes, while minimizing false positives. The peak table method performs well at the higher absolute concentrations. However, at lower concentrations, some of the four spiked analytes are not found or are disqualified due to low deconvolution quality and/or misidentified peaks.

CONCLUSIONS

Tile-based F-ratio analysis of GC × GC–TOFMS data facilitates confident analyte detection for small concentration changes between sample classes for discovery-based experimentation, even for analytes that are heavily interfered with little mass spectral selectivity. The software is computationally fast, requiring less than 10 min of computation time per comparison (including null distribution analysis with six null comparisons) on a midlevel personal computer. After applying null distribution analysis, the results of the study showed that spiked analytes could be reliably discovered at ∼1 to ∼10 ppm (∼5 to ∼50 pg using a 200:1 split) in a complex diesel matrix with minimal occurrence of false positives above the 0.1% null probability range. The determination of the null probability limits using the combinatorial technique described herein is rapid and provides a defensible F-ratio cutoff that is determined algorithmically for each experiment based on the variation present in the data. This approach is superior to rule-of-thumb thresholds or manually determining an appropriate threshold for each experiment based on inspecting the hit lists.

This analytical platform is broadly applicable for the analysis of samples where class distinguishing analytes may be present at a range of signals and in which minor chromatographic misalignment may be present, without the need for explicit 2D chromatographic alignment. The tile-based approach is demonstrated to reduce the occurrence of false positives compared to pixel-based methods and permits lower discovery limits compared to peak table methods. A future study could include an expanded selection of spiked analytes (substantially

DOI: 10.1021/ac504472s
Anal. Chem. 2015, 87, 3812–3819

more true positives), allowing for characterization of a "true positive distribution" and a more thorough comparison to other nontargeted analysis methods.

■ ASSOCIATED CONTENT

Supporting Information

Additional information as noted in text. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Tel.:+1-206-685-2328. Fax: +1-206-685-8665. E-mail: synovec@chem.washington.edu.

Present Address

#W.C.S.: Dow Chemical Corporation, 2301 North Brazosport Blvd., Freeport, TX 77541, U.S.A.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the Internal Revenue Service (IRS) under an Interagency Agreement with the US Department of Energy (DOE) under Contract DE-AC05-76RLO 1830 with the Pacific Northwest National Laboratory.

■ REFERENCES

(1) Liu, Z.; Phillips, J. B. *J. Chromatogr. Sci.* **1991**, *29*, 227−231.
(2) Adahchour, M.; van Stee, L. L.; Beens, J.; Vreuls, R. J.; Batenburg, M. A.; Brinkman, U. A. Th. *J. Chromatogr. A* **2003**, *1019*, 157−172.
(3) Beens, J.; Adahchour, M.; Vreuls, R. J. J.; van Altena, K.; Brinkman, U. A. Th. *J. Chromatogr. A* **2001**, *919*, 127−132.
(4) Bruckner, C. A.; Prazen, B. J.; Synovec, R. E. *Anal. Chem.* **1998**, *70*, 2796−2804.
(5) Dallüge, J.; van Rijn, M.; Beens, J.; Vreuls, R. J.; Brinkman, U. A. Th. *J. Chromatogr. A* **2002**, *965*, 207−217.
(6) Kinghorn, R. M.; Marriott, P. J. *J. High Resolut. Chromatogr.* **1998**, *21*, 620−622.
(7) Seeley, J. V.; Kramp, F.; Hicks, C. J. *Anal. Chem.* **2000**, *72*, 4346−4352.
(8) Shellie, R.; Mondello, L.; Marriott, P.; Dugo, G. *J. Chromatogr. A* **2002**, *970*, 225−234.
(9) Pierce, K. M.; Kehimkar, B.; Marney, L. C.; Hoggard, J. C.; Synovec, R. E. *J. Chromatogr. A* **2012**, *1255*, 3−11.
(10) Pierce, K. M.; Nadeau, J. S.; Synovec, R. E. In *Gas Chromatography*; Poole, C. F., Ed.; Elsevier: Amsterdam, 2012; pp 415−434.
(11) Yang, S.; Hoggard, J. C.; Lidstrom, M. E.; Synovec, R. E. In *Metabolomics in Practice*; Lämmerhofer, M., Weckwerth, W., Eds.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, 2013; pp 69−92.
(12) Reichenbach, S. E.; Tian, X.; Cordero, C.; Tao, Q. *J. Chromatogr. A* **2012**, *1226*, 140−148.
(13) Reichenbach, S. E.; Ni, M.; Kottapalli, V.; Visvanathan, A. *Chemom. Intell. Lab. Syst.* **2004**, *71*, 107−120.
(14) Reichenbach, S. E.; Tian, X.; Boateng, A. A.; Mullen, C. A.; Cordero, C.; Tao, Q. *Anal. Chem.* **2013**, *85*, 4974−4981.
(15) Harvey, P. M.; Shellie, R. A. *Anal. Chem.* **2012**, *84*, 6501−6507.
(16) Organtini, K. L.; Myers, A. L.; Jobst, K. J.; Cochran, J.; Ross, B.; McCarry, B.; Reiner, E. J.; Dorman, F. L. *J. Chromatogr. A* **2014**, *1369*, 138−146.
(17) Vogt, L.; Gröger, T.; Zimmermann, R. *J. Chromatogr. A* **2007**, *1150*, 2−12.
(18) Marney, L. C.; Siegler, W. C.; Parsons, B. A.; Hoggard, J. C.; Wright, B. W.; Synovec, R. E. *Talanta* **2013**, *115*, 887−895.
(19) Brokl, M.; Bishop, L.; Wright, C. G.; Liu, C.; McAdam, K.; Focant, J.-F. *J. Chromatogr. A* **2014**, *1370*, 216−229.
(20) Fraga, C. G.; Prazen, B. J.; Synovec, R. E. *Anal. Chem.* **2000**, *72*, 4154−4162.
(21) Fraga, C. G.; Prazen, B. J.; Synovec, R. E. *Anal. Chem.* **2001**, *73*, 5833−5840.
(22) Pierce, K. M.; Wood, L. F.; Wright, B. W.; Synovec, R. E. *Anal. Chem.* **2005**, *77*, 7735−7743.
(23) Kempa, S.; Hummel, J.; Schwemmer, T.; Pietzke, M.; Strehmel, N.; Wienkoop, S.; Kopka, J.; Weckwerth, W. *J. Basic Microbiol.* **2009**, *49*, 82−91.
(24) Wang, B.; Fang, A.; Heim, J.; Bogdanov, B.; Pugh, S.; Libardoni, M.; Zhang, X. *Anal. Chem.* **2010**, *82*, 5069−5081.
(25) Gros, J.; Nabi, D.; Dimitriou-Christidis, P.; Rutler, R.; Arey, J. S. *Anal. Chem.* **2012**, *84*, 9033−9040.
(26) Khummueng, W.; Harynuk, J.; Marriott, P. J. *Anal. Chem.* **2006**, *78*, 4578−4587.
(27) Nadeau, J. S.; Wilson, R. B.; Hoggard, J. C.; Wright, B. W.; Synovec, R. E. *J. Chromatogr. A* **2011**, *1218*, 9091−9101.
(28) Johnson, K. J.; Synovec, R. E. *Chemom. Intell. Lab. Syst.* **2002**, *60*, 225−237.
(29) Mohler, R. E.; Dombek, K. M.; Hoggard, J. C.; Pierce, K. M.; Young, E. T.; Synovec, R. E. *Analyst* **2007**, *132*, 756−767.
(30) Humston, E. M.; Dombek, K. M.; Tu, B. P.; Young, E. T.; Synovec, R. E. *Anal. Bioanal. Chem.* **2011**, *401*, 2387−2402.
(31) Beckstrom, A. C.; Humston, E. M.; Snyder, L. R.; Synovec, R. E.; Juul, S. E. *J. Chromatogr. A* **2011**, *1218*, 1899−1906.
(32) Hantao, L. W.; Toledo, B. R.; de Lima Ribeiro, F. A.; Pizetta, M.; Pierozzi, C. G.; Furtado, E. L.; Augusto, F. *Talanta* **2013**, *116*, 1079−1084.
(33) Gröger, T.; Schäffer, M.; Pütz, M.; Ahrens, B.; Drew, K.; Eschner, M.; Zimmermann, R. *J. Chromatogr. A* **2008**, *1200*, 8−16.
(34) Heim, J. *LECO Corporation, Technical Note, Life Sciences and Chemical Analysis Solutions*; Form 203-821-377; LECO Corporation: St. Joseph, Michigan, 2010.
(35) Almstetter, M. F.; Appel, I. J.; Dettmer, K.; Gruber, M. A.; Oefner, P. J. *J. Chromatogr. A* **2011**, *1218*, 7031−7038.
(36) Welke, J. E.; Manfroi, V.; Zanus, M.; Lazzarotto, M.; Alcaraz Zini, C. *Food Chem.* **2013**, *141*, 3897−3905.
(37) Stadler, S.; Stefanuto, P.-H.; Brokl, M.; Forbes, S. L.; Focant, J.-F. *Anal. Chem.* **2013**, *85*, 998−1005.
(38) Pasikanti, K. K.; Norasmara, J.; Cai, S.; Mahendran, R.; Esuvaranathan, K.; Ho, P. C.; Chan, E. C. Y. *Anal. Bioanal. Chem.* **2010**, *398*, 1285−1293.
(39) Li, X.; Lu, X.; Tian, J.; Gao, P.; Kong, H.; Xu, G. *Anal. Chem.* **2009**, *81*, 4468−4475.
(40) Ma, C.; Wang, H.; Lu, X.; Wang, H.; Xu, G.; Liu, B. *Metabolomics* **2009**, *5*, 497−506.
(41) Qiu, Y.; Lu, X.; Pang, T.; Ma, C.; Li, X.; Xu, G. *J. Sep. Sci.* **2008**, *31*, 3451−3457.
(42) Benjamini, Y.; Hochberg, Y. *J. R. Stat. Soc.: Ser. B Stat. Methodol.* **1995**, *57*, 289−300.
(43) Hoggard, J. C. peg2mat3p8; http://depts.washington.edu/synlab/software/, 2011.
(44) Massart, D. L. *Chemometrics: A textbook*; Elsevier Science Pub. Co.: Amsterdam; New York, 1988.
(45) Duda, R. O.; Hart, P. E.; Stork, D. G. *Pattern Classification*; John Wiley & Sons: New York, 2012.
(46) Efron, B. *J. Am. Stat. Assoc.* **2005**, *100*, 1−5.
(47) Efron, B. *J. Am. Stat. Assoc.* **2007**, *102*, 93−103.
(48) Efron, B. *J. Am. Stat. Assoc.* **2010**, *105*, 1042−1055.
(49) Efron, B.; Tibshirani, R. *Genet. Epidemiol.* **2002**, *23*, 70−86.
(50) Leek, J. T.; Storey, J. D. *Stat. Appl. Genet. Mol. Biol.* **2011**, *10*, 1−22.
(51) Storey, J. D. *J. R. Stat. Soc.: Ser. B Stat. Methodol.* **2002**, *64*, 479−498.
(52) Storey, J. D. *J. R. Stat. Soc.: Ser. B Stat. Methodol.* **2007**, *69*, 347−368.
(53) Woo, S.; Leek, J. T.; Storey, J. D. *Bioinformatics* **2011**, *27*, 509−515.

DOI: 10.1021/ac504472s
Anal. Chem. 2015, 87, 3812−3819

(54) Vis, D. J.; Westerhuis, J. A.; Smilde, A. K.; van der Greef, J. *BMC Bioinf.* **2007**, *8*, 322.
(55) Lu, H.; Liang, Y.; Dunn, W. B.; Shen, H.; Kell, D. B. *TrAC, Trends Anal. Chem.* **2008**, *27*, 215−227.

DOI: 10.1021/ac504472s
Anal. Chem. 2015, 87, 3812−3819