

# Peak Detection with Varian Star Workstation for Varian 3800 and 450 Gas Chromatographs

by:

**Randall Bramston-Cook**

Lotus Consulting

5781 Campo Walk

Long Beach, Ca 90803

310/569-0128

[ebamstoncook@msn.com](mailto:ebamstoncook@msn.com)

December 10, 2009

Copyright 2009 Lotus Flower, Inc.



Gone are the days when rulers, planimeters and “cut-n-weigh” were used to determine detector responses from gas and liquid chromatographs. Now virtually every chromatograph has a computer attached to gather raw data, compute areas or heights, convert these into concentrations and issue a report. Fortunately, much of these processes are automatic, but generated results can be invalid due to improper settings for peak detection, inappropriate baseline assignments and parameter entry errors by the operator. A computer can always generate a number, but that number may not be correct. Understanding how the workstation processes data can help yield better quality for analyte concentrations. A critical factor in maintaining good reproducibility and accuracy of measurements is proper assignment of peak responses above baselines. If peak start and end points are not set appropriately, the peak response can be too large or too small, compared with its true value, and can vary dramatically with replicate runs of the same sample.

Proper assignment of a peak response involves some judgments to distinguish a real peak from noise, rising backgrounds, change in peak shapes from temperature and flow programming, unexpected disruptions in detector response from instabilities, and flow disruptions (from valve switches, for example). The data system, with preset parameter settings, can usually find “peaks” and assign an area or height, but often these initial values do not match the criteria that an experienced operator might have given. This reported value may not properly represent the true concentration of an unknown in every case.

Varian Star Workstation<sup>1,2</sup> for data from gas chromatographic detectors uses a peak detection algorithm similar to one originally developed in 1968 with the Varian 200 Data System set up for data collection from multiple gas chromatographs, and has been continually enhanced through the years in Varian CDS 101/111, Varian Vista 401/402, and Varian DS601/651/654 and now in the Varian Star Workstation<sup>3</sup> with many enhancements over the older programs.

Varian Star Workstation provides a number of user selectable parameters to tune peak detection to specific conditions in use. For example, when a high performance capillary column is employed that generates some very sharp peaks, the user can set up parameters readily to fully detect these sharp peaks and ignore baseline drift from temperature programming. At the other extreme, if peaks are broad from use of packed or megabore columns, settings can be employed to expect these broader peaks and ignore detector spikes. Fortunately, preset parameters in this Workstation are usually sufficient to pick out many peaks, albeit not always optimal, and provide a starting point for fine tuning settings to accurately work on subsequent runs of the same ilk without further adjustments.

Varian Star Workstation performs automatic full storage of every chromatogram, with all raw data points. Parameters involved in the initial data collection (such as data rate and chromatographic conditions) cannot be adjusted after data for the run is gathered. Others can be changed, and their effects on a chromatogram can be visually inspected. Then this modified method can be saved and made available for use with subsequent analyses.

---

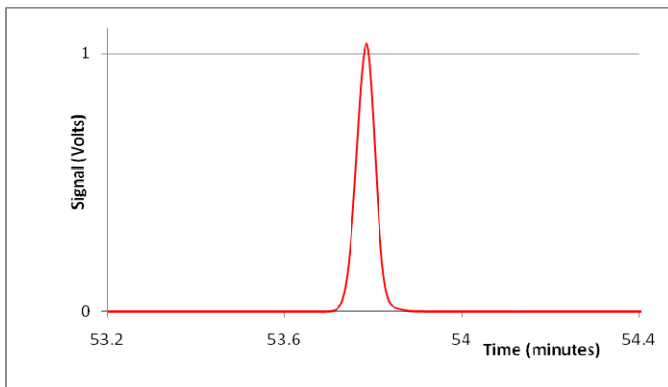
<sup>1</sup> Varian MS Workstation has been undergoing improvements over the years, and some significant changes related to peak detection have been implemented in later versions. All discussions here apply to Star Version 6.4.1, and with GC detectors in MS Workstation Version 6.4 and later.

<sup>2</sup> Varian Star Workstation does not have drivers to operate the Varian 450 as a stand-alone GC, but MS Workstation can collect GC detector data from the Varian 450 in conjunction with MS data. Operations described here are valid for both Workstations, for GC data only.

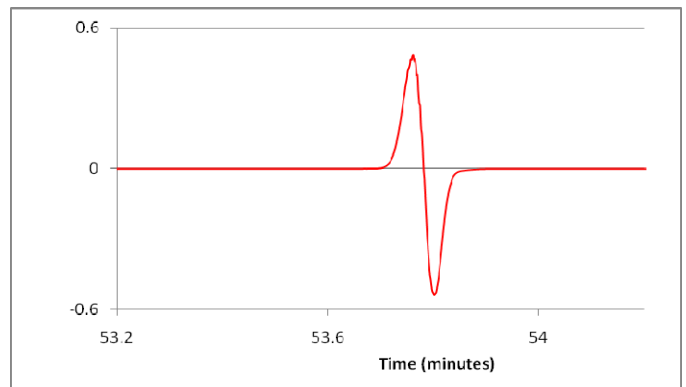
<sup>3</sup> This document presumes that the reader has some familiarity with various screens and operations within Varian MS Workstation. Hints are provided for locating specific operations within the Workstation for the following discussion points, but precise steps are not listed and are assumed to be obvious to the reader.

## Determination of Retention Time

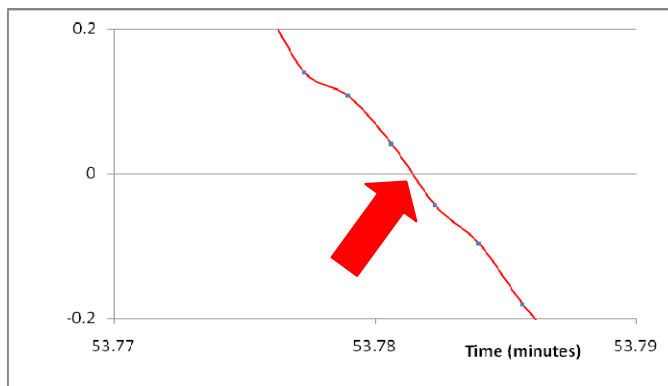
Varian Star Workstation establishes retention time of a peak by computing the zero crossover of the first derivative of the peak. If this value does not correspond to a collected data point, the precise time is determined by linear interpolation of point before and after the zero crossing.



**Retention time is determined by zero-crossing of the first derivative of a chromatographic peak.**



**The derivative is generated by an appropriate digital filter across the peak (see later discussions on Filtering).**



**The precise retention time is set by linear interpolation of the points before and after the zero crossing of the derivative.**

## Retention Time Adjustments with Reference Peaks

Measurements in chromatography cannot be considered exacting, as random and systematic errors can introduce some variability in retention times (along with area measurements). To help correct for systematic deviations in retention times, certain compounds can be assigned as Reference Peaks. Any movement in time of these peaks is detected and a correction based on this change is applied proportionally to all non-reference or "Other" peaks, based on their relative retention to reference peaks. If multiple Reference Peaks are assigned, then the new expected Retention Times are proportioned based on their proximity to these reference peaks.

	Retention Time	Peak Name	Ref
1	5.474	2-Octanone	<input checked="" type="checkbox"/>
2	6.719	1-Octanol	<input type="checkbox"/>
3	7.232	n-Decane	<input type="checkbox"/>
4	8.668	n-Dodecane	<input type="checkbox"/>
5	10.000	2,4-dimethyl...	<input type="checkbox"/>
6	10.461		<input type="checkbox"/>

**Define Peak Windows...**

**Define Peak Windows**

Define Reference Peak Windows

Width (minutes): 0.50

Retention Time %: 0.0

Define Other Peak Windows

Width (minutes): 0.05

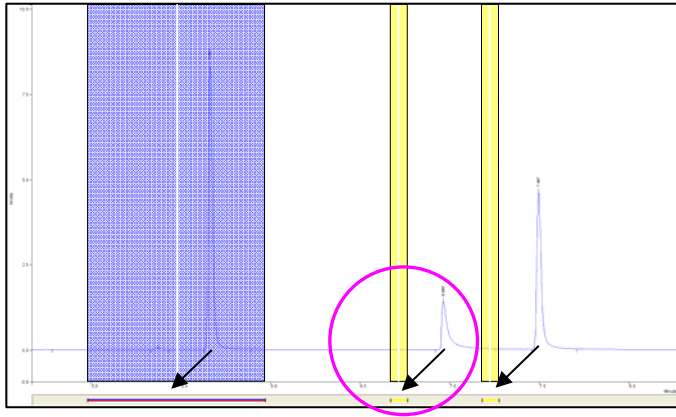
Retention Time %: 0.0

Unretained Peak Time

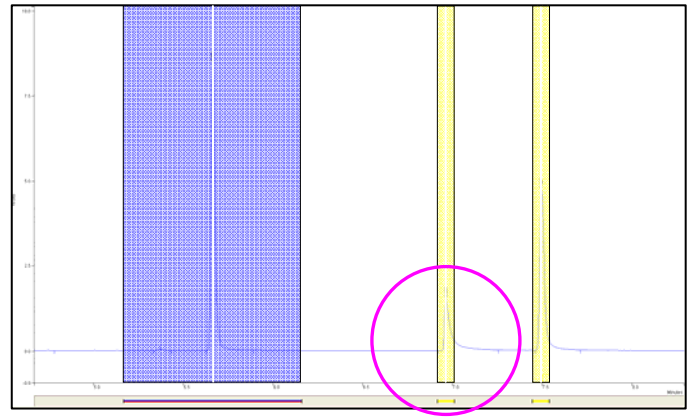
Time (minutes): 0.00

Save Cancel

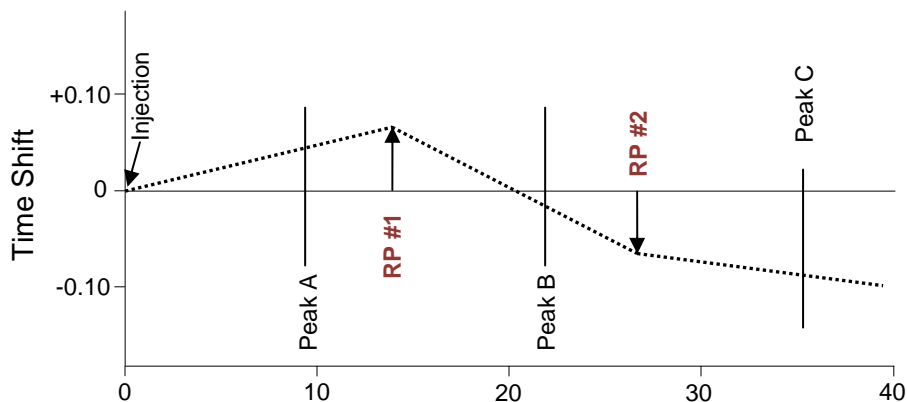
Reference Peaks **always** must be present in **every** sample and standard in a series, and must be the **largest** peak within the Window; typically internal standards are perfect candidates. Non-reference peaks are picked out as the **closest** peak to the non-reference target Retention Time, after time adjustment from a Reference Peak correction. For a Calibration run type, retention times in the Peak Table are updated based on the newly located peak positions. With a run type of "Analysis", the Peak Table in the Method is not altered as a result of these operations and Reference Peaks are only used to pick out peak movement and properly identify analyte peaks.



Normal variations in chromatographic conditions can shift retention times away from the expected locations. And adjacent peaks can be mislabeled and wrong results are reported.



Proper use of Reference Peaks can sense a shift in retention times and allow corrections to move the expected retention times over to the correct peaks. Peak Table is only updated if run is a Calibration type.



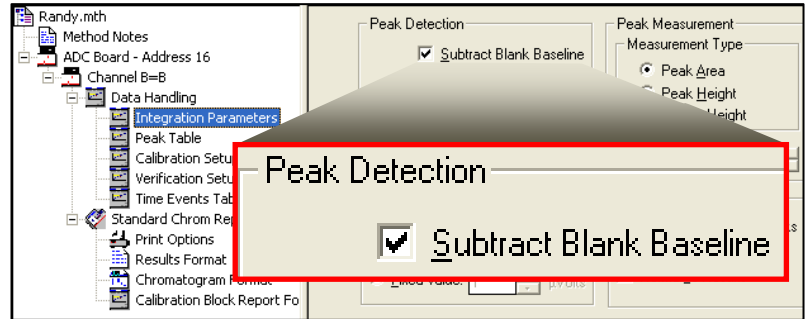
Multiple Reference Peaks (for example RP #1 and RP #2) can provide time corrections over the whole chromatogram. New expected Retention Times for unknown peaks are adjusted proportionally to their positions relative to adjacent reference peaks. This synthetic example illustrates the corrections involved with two reference peaks.

Compound	Expected Retention Time	Measured Retention Time	Measured Shift	Correction	New Expected Retention Time
Peak A	9.374	--		+0.045	9.419
<b>IRP #1</b>	<b>13.841</b>	<b>13.907</b>	<b>+0.066</b>	--	--
Peak B	21.844	--		-0.015	21.829
<b>IRP #2</b>	<b>26.752</b>	<b>26.816</b>	<b>-0.064</b>	--	--
Peak C	35.395	--		-0.064	35.331

## Baseline Subtraction

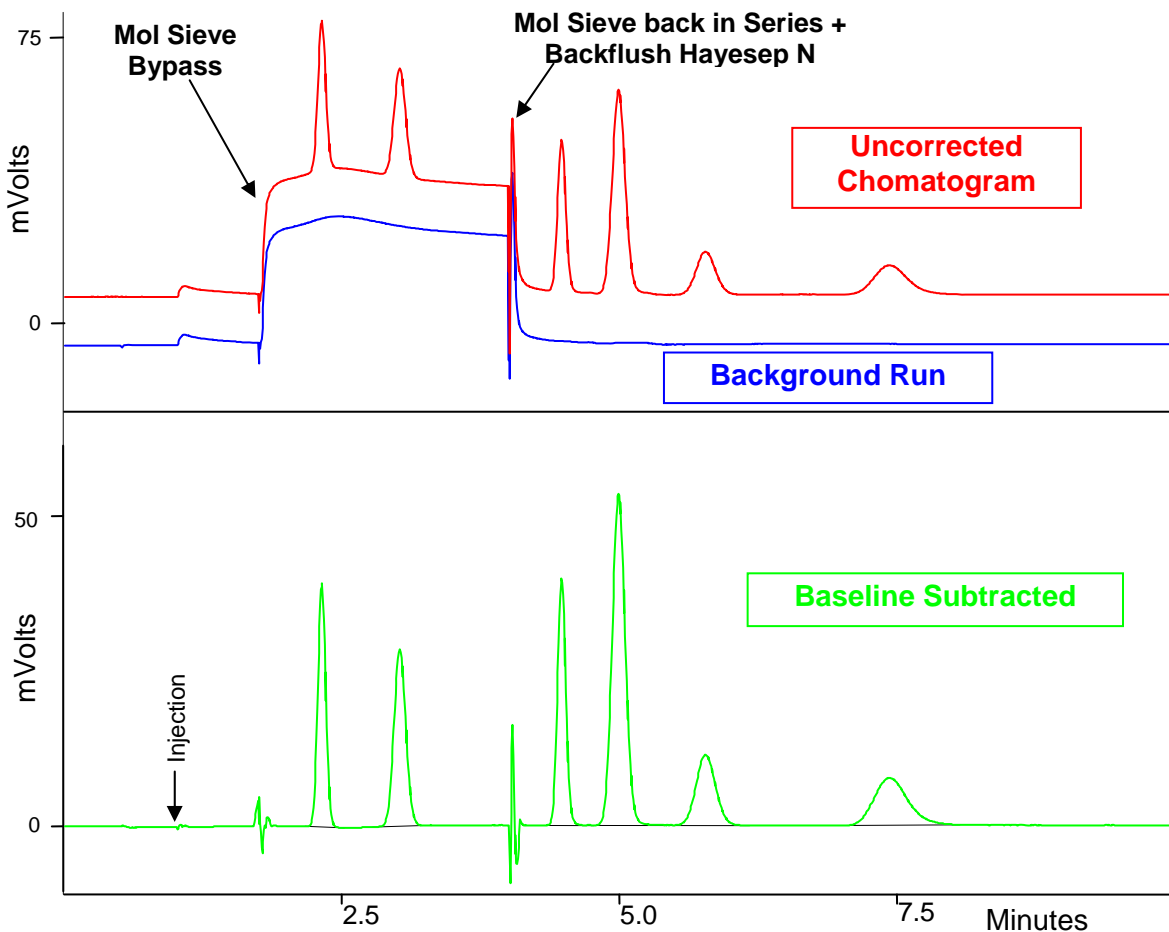
Measurements involving temperature programming can often generate reproducible baseline changes from column bleed. Another common situation is with valve switching where severe pressure surges can introduce dramatic baseline upsets, especially with flow-sensitive detectors such as thermal conductivity detector. These

disturbances can directly affect peak area assignments. These effects can be diminished by performing measurements with a system blank with no injection and then subtracting point-by-point those baseline variations to yield a better chromatogram that can be processed more accurately into areas by the Workstation. The baseline run is a reduced set of 257 to 512 data points across the chromatogram run time by a bunching process, and is stored within the method used for its collection and is subsequently subtracted from sample runs prior to peak detection.



Untitled.smp - Generic SampleList

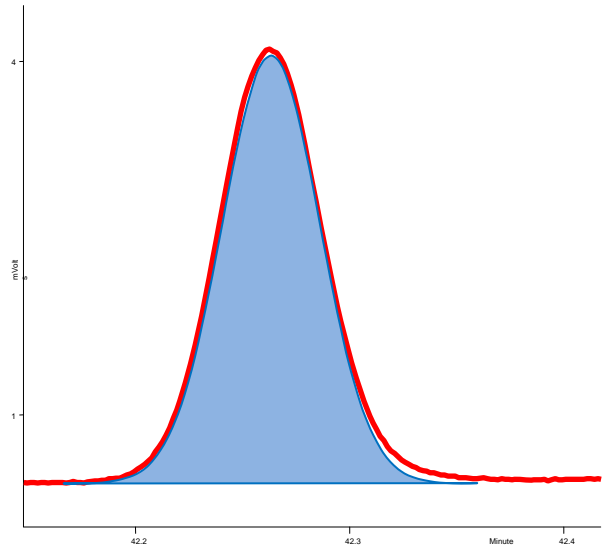
	Sample Name	Sample Type
1	System Blank	Baseline
2		



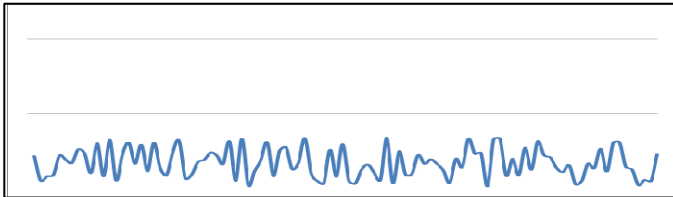
**Valve actions can produce baseline upsets, which can be properly corrected by baseline subtraction, to yield a flatter baseline for better peak area assignments.**

## Peak Simulations

Shape of typical chromatographic peaks (red trace) can be approximated by a Gaussian distribution (shown in blue), as displayed to the right. Using this mathematical model, various data treatments available in Varian Star Workstation, such as data collection rates and bunching can be simulated. In discussions below, 600 points on a Gaussian curve are generated in Microsoft Excel to study various effects on peak processing.



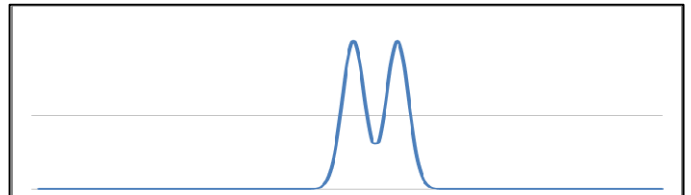
**Normal chromatographic peaks are nearly Gaussian in shape.**



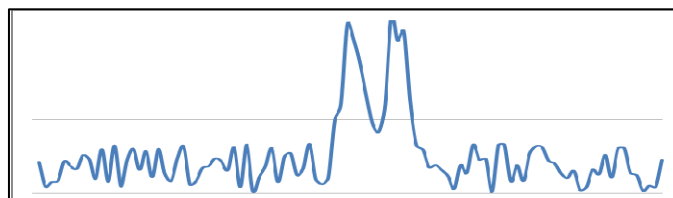
**Noise is simulated by plotting a group of random numbers.**

Noise is imitated by generating a series of random numbers within a specified range and then plotted.

Artificial peaks are created by adjusting Gaussian points to relate to the noise numbers produced. Two closely spaced peaks are used to look at effects of various parameters on resolution of the two, especially when these peaks are combined with random noise numbers.



**Overlapping peaks are created from Gaussian distributions, with a 70% valley penetration.**



**The doublet peak structure is mathematically combined with the random noise to create a data set closely resembling real peaks near their detection limits.**

## Definition of Terms

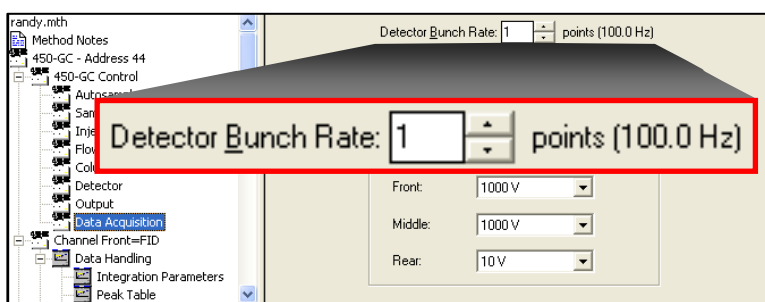
**Conversion Point** – This is raw data from the gas chromatograph detectors, and has some digital filtering applied to it at the chromatograph.

**Detector Bunch Point** – Conversion points are averaged together to suppress random noise and form this Bunch Point. This is data that is transferred from the chromatograph to the Workstation.

**Cluster Point** – Often the Bunch Points are more frequent than is needed to optimally define a peak shape and can make peak detection more difficult. Bunch Points are averaged together to provide 10 to 20 Cluster Points across the top of a peak.

## Detector Bunch Rate

Maximum detector signal conversions occur at 40 Hz for the Varian 3800, and 100 Hz for the Varian 450. To adequately define a peak shape, a peak should have between 10 and 20 points across the top of the peak above the width at half height. For the Varian 3800, this allows peaks as narrow as 0.25 seconds to be well defined, and 0.1 seconds for the Varian 450.



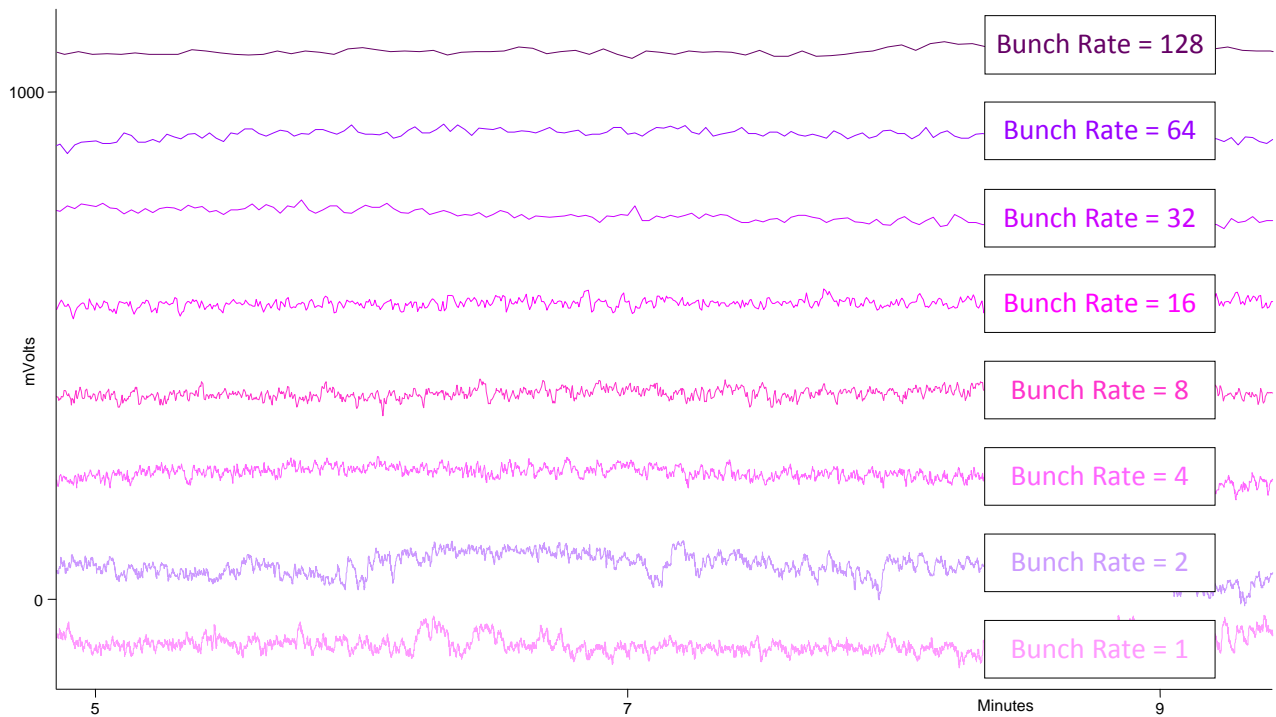
Since most chromatographic peaks are typically much wider than these, signal conversions can be averaged into “bunched” points to still properly describe the peak and allow some suppression of noise through this averaging. A Detector Bunch Point is an arithmetic average of the specified number of raw data points. This point is then transmitted from the gas chromatograph to the Workstation. Its value remains fixed through the run and cannot be changed after initial data collection. The value for this parameter should be set to correspond to one-tenth of the narrowest expected peak in the chromatogram.

Data Collection Rate for Varian 3800								
Detector Bunch Rate	1	2	4	8	16	32	64	128
Data Rate (Hz)	40	20	10	5	2.5	1.3	0.6	0.3
Optimal Minimum Peak Width (sec)	0.25	0.5	1.0	2.0	4.0	8.0	16.0	32.0

Data Collection for Varian 450								
Detector Bunch Rate	1	2	4	8	16	32	64	128
Data Rate (Hz)	100	50	25	12.5	6.3	3.1	1.6	0.8
Optimal Minimum Peak Width (sec)	0.1	0.2	0.4	0.8	0.16	0.32	0.64	1.28



A value too fine will display more noise and larger data files, and a setting too large can cause wanted peaks to be melded into the baseline. The preset rate is 4 for the Varian 3800, and 16 for the Varian 450, and is likely to be appropriate for most chromatograms. Further averaging takes place with expected Peak Widths, as discussed in another section.



[This space is reserved for upcoming data from 450]

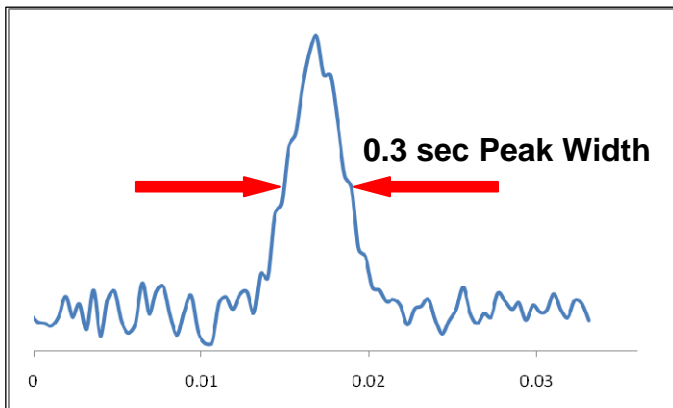
## Digital Filtering

Data collected at 40 Hz (Varian 3800) and 100 Hz (Varian 450) will display some very noisy chromatograms and make sorting out legitimate peaks from baseline noise more difficult. To reduce high frequency noise without significant distortion of peak shapes, a digital filter is applied to the raw data points in the gas chromatograph prior to transmission to the Workstation. Filtered data is put on view in all chromatogram displays, including Interactive Graphics and Reports. All filters discussed here use a common formula:

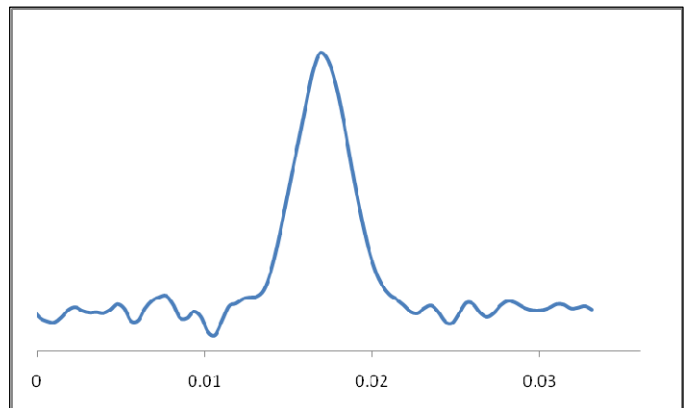
$$y_n = \sum_{i=0}^N b_i x_{n-i}$$

Coefficients,  $b_i$ , and number of points,  $N$ , depend on the function of the filter, including Savitsky-Golay, Finite Impulse Response and derivative. The classical boxcar filter has equal values for all of its coefficients.

For data generated with the Varian 3800, a 4-point Finite Impulse Response (FIR) filter is employed to noticeably quiet down noise while still maintaining original peak shapes.

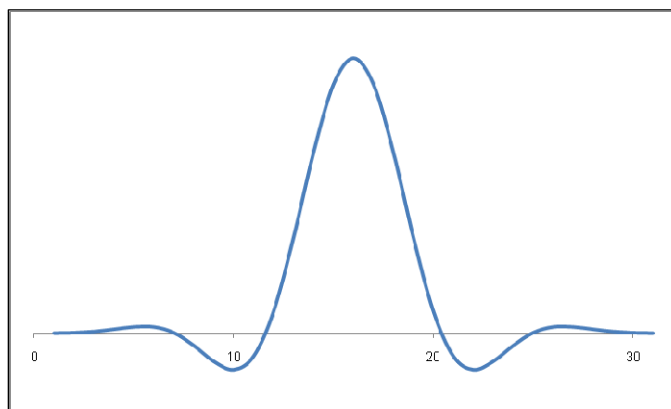


**Raw data from Varian 3800 detectors are simulated by random noise added to a Gaussian peak.**

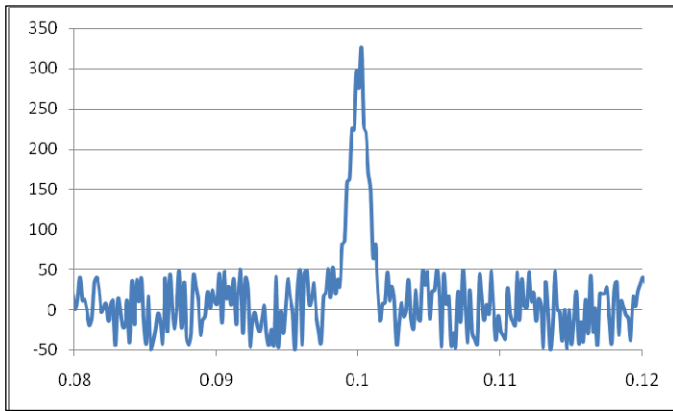


**Synthetic data for Varian 3800 are filtered with a 4-point Finite Impulse Response filter, yielding similar peak shape to the original Gaussian Peak, yet reduced noise.**

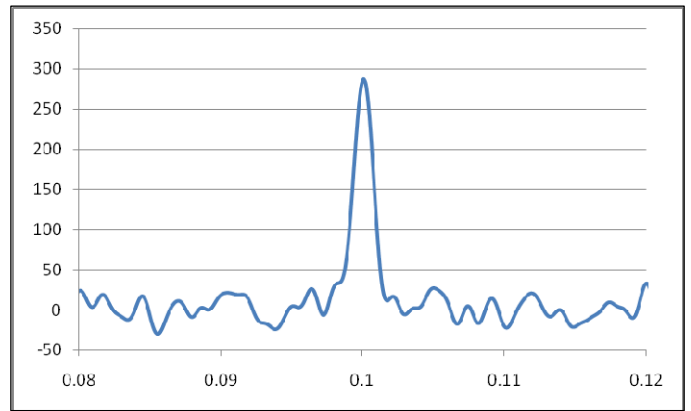
For the Varian 450, a 31-point Finite Impulse Response filter is used. The faster data rate (100 Hz) mandates a bit more digital filtering to suppress higher frequency noise and still leave peaks intact as much as possible.



**31-point Finite Impulse Response (FIR) filter.**



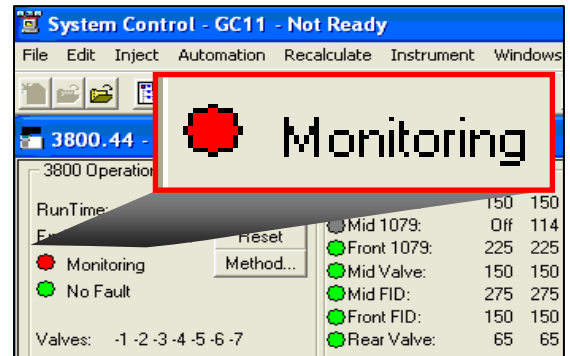
**A created chromatogram imitates raw data generated from a Varian 450 at 100 Hz with a Gaussian peak superimposed over random noise.**



**The “chromatogram” at left is passed through a 31-point Finite Impulse Response filter to suppress high frequency noise and yet keeps the original peak nearly intact.**

## Noise

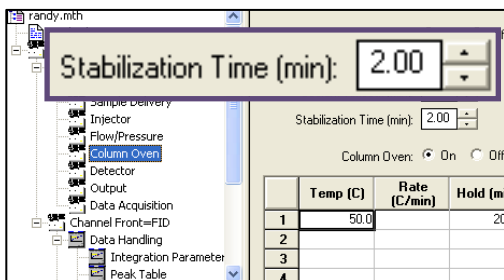
Most of a typical chromatogram has no useful information and only “occasionally” do peaks arise out of the chaos to yield useful data for identification and quantitation of analytes. To separate this noise from peaks, this peak detection process makes a judgment on start and end points of peaks based on what noise is detected just prior to the beginning of the chromatographic run, during a “Monitoring” phase of the data collection process.



The time interval of the noise monitoring is set by the Detector Bunch Rate selected and the number of bunched points selected by the user. The possible entries for Noise Monitor Length is 16 to 128 bunched points. The time interval is computed by:

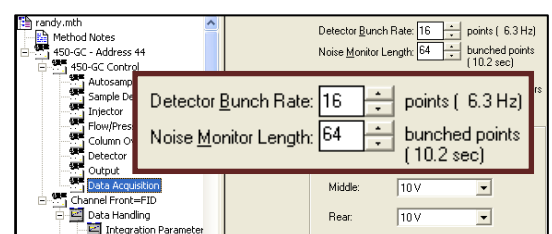
$$\text{Monitoring Time} = 0.01 \text{ s} \times \text{Detector Bunch Rate} \times \text{Noise Monitor Length}$$

The interval should be selected to be sufficient to see enough noise to assess a realistic value, but not too long to delay the start of the run. This interval must represent the typical noise at the start of the run and must not be subject to drifts from column temperature instability and detector wavering after valve upsets and parameter changes. Column Stabilization Time sets an interval for



**Column Stabilization Time should be set long enough to ensure detector signal stability during Noise Monitor interval**

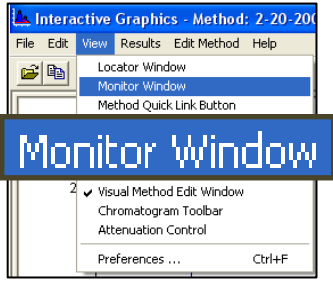
all instrument hardware to achieve stability prior to this noise screening. If column temperature programming is involved in the protocol, then the preset column Stabilization Time of 2.00 minutes should be preserved.



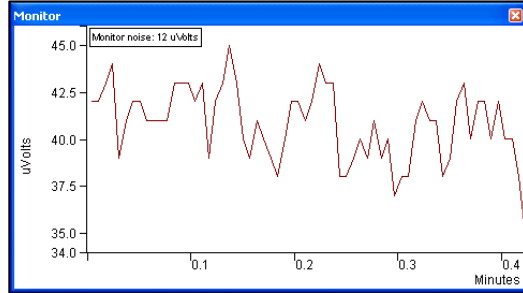
**Parameter entry screen in method for Detector Bunch rate and Noise Monitor Length.**

If final chromatographic conditions and valve operations at the end of the previous run match those of the start of the next run are identical, then this Stabilization Time can be reduced to zero, since the baseline noise should already be representative.

Every archived .RUN file possesses noise data collected during the monitoring interval and can be displayed in Interactive Graphics by selecting View → Monitor Window. Displayed noise should exhibit typical expected noise and should not show drift or unusual spikes that are not typical of the chromatographic conditions. If circumstances cannot guarantee acceptable noise levels from this monitoring, then noise can be fixed either at the start of a method, or fixed from a previously established value. If the noise value determined from the original run is out of sync with expected levels, it can be fixed in post-run operations in Interactive Graphics or cluster reprocessing in Recalculation List or Batch Report.

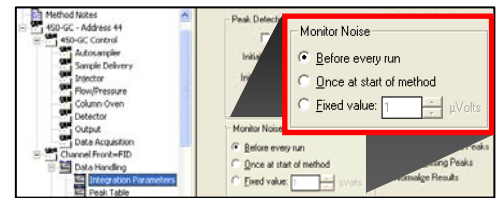


Monitored noise at start of a run can be viewed in Interactive Graphics.

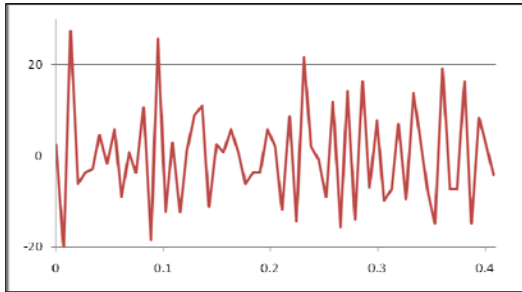


Monitored noise display is actual signal variation, with RMS noise of the first derivative, listed as, in this case, 12 microvolts/second.

Established value. If the noise value determined from the original run is out of sync with expected levels, it can be fixed in post-run operations in Interactive Graphics or cluster reprocessing in Recalculation List or Batch Report.



A choice is provided for how to perform noise monitoring.



The first derivative of the noise displayed above illustrates how a noise value is established.

If the noise level is set too high, real peaks can be integrated improperly with integration baselines set too high on sides of the peak, or peaks can be missed completely. A value too low will allow baseline disruptions to be picked out as peaks, and baseline assignments for peaks can be skewed. An appropriate value should be determined from preliminary runs and then fixed for routine operations to avoid possible disruptions in peak processing from unexpected detector drift or instabilities.

Peak No.	Peak Name	Result	Time (min)	Offset (min)	Area (counts)	Rep.	1/2	Status
1	3-Methylamine	0.0	25.463	-0.040	2949	NS	4.6	C
2	2-Methylamine	0.7	26.786	0.145	519	NS	2.7	
3	1-Heptane	2.7	36.196	0.061	2027	NS	4.9	
4	n-Heptane	1.0	37.090	0.091	782	NS	4.0	
5	Methylcyclopentane	1.2	39.126	0.096	918	NS	5.2	
6	2,2-Dimethylpropane	2.3	39.354	0.070	1849	NS	4.0	
7	Hexane	1.9	41.074	0.055	1244	NS	5.4	
8	Cyclopentane	2.6	41.877	0.079	2257	NS	5.4	
9	2-Methylpropane	1.4	42.222	0.063	1322	NS	4.7	
10	2,2-Dimethylbutane	3.1	42.371	0.071	2481	NS	5.5	
11	n-Heptane	1.5	42.462	0.072	1149	NS	5.3	U
12	2,2,4-Trimethylpentane	2.2	43.823	0.071	1833	NS	5.4	
13	n-Heptane	1.5	44.465	0.065	1054	NS	5.3	
14	Methylcyclopentane	2.7	46.203	0.062	1497	NS	4.1	
15	2,2,4-Trimethylpentane	1.5	47.771	0.054	1213	NS	3.5	
16	Toluene	2.1	48.160	0.046	2092	NS	4.0	
17	2-Methylpropane	1.4	48.484	0.047	1229	NS	3.0	U
18	2-Methylpropane	1.7	48.843	0.050	1299	NS	3.0	U
19	n-Heptane	1.9	49.909	0.048	1930	NS	2.5	
20	n-Heptane	1.6	51.966	0.044	1092	NS	2.3	
21	n-Heptane	1.5	52.189	0.041	1539	NS	4.0	
22	Styrene	2.3	52.705	0.036	1376	NS	2.2	
23	n-Heptane	1.5	52.844	0.041	1133	NS	2.3	
24	n-Heptane	1.5	53.090	0.033	1045	NS	2.0	
25	n-Heptane	2.5	53.490	0.034	1911	NS	2.1	
26	n-Heptane	1.9	54.432	0.032	1289	NS	2.1	
27	n-Heptane	1.5	54.590	-0.027	1307	NS	2.0	
28	1,2-Dichloroethane	1.4	54.758	0.031	1206	NS	2.3	
29	n-Heptane	2.0	55.058	0.026	1897	NS	2.3	
30	1,2-Dichloroethane	3.0	55.210	0.028	1463	NS	2.0	
31	n-Heptane	3.4	55.379	-0.024	2716	NS	4.0	
32	1,2-Dichloroethane	1.8	56.040	0.034	1134	NS	2.1	
33	n-Heptane	2.6	56.415	0.031	1774	NS	2.1	
34	n-Dichloroethane	1.7	56.821	0.030	1125	NS	2.2	
35	n-Heptane	1.5	57.365	0.033	940	NS	2.2	
36	n-Heptane	3.7	59.093	0.035	2214	NS	2.2	

Noise (used): 72 microVolts - monitored before this run

Noise (used): 72 microVolts - monitored before this run

Manual Injection

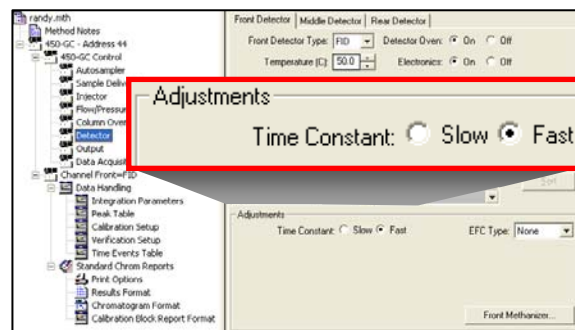
Calib. out of range; No Recovery Action Specified

Every report contains Noise (used) to detect peaks. The value is Root Mean Square of the first derivative for noise sensed just prior to the start of every run, unless another fixed value is chosen.

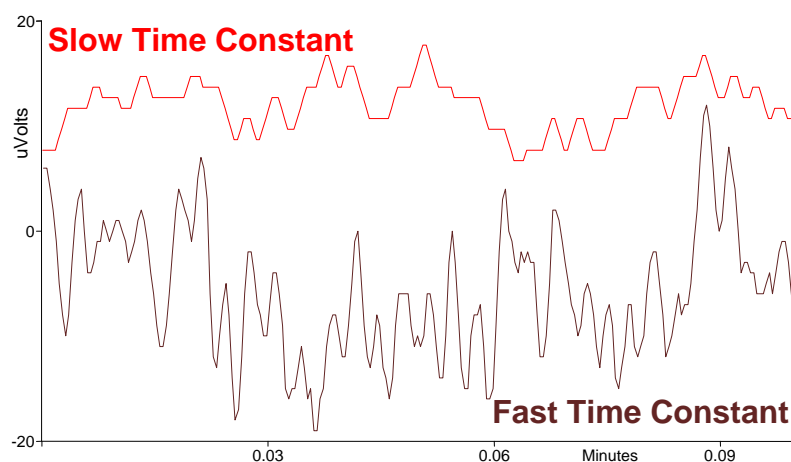
Noise cannot be predetermined and varies dramatically from one detector type to another. Even within the same detector type, operating conditions can greatly alter the characteristic noise. Reported Noise should be tracked, via a daily log, for specific operating conditions to ensure that consistent readings are achieved from run to run. Variations from an established norm can indicate deviations from normal circumstances and can point toward operational difficulties.

## Detector Time Constant

All available detectors in Varian 3800 have a user-selectable parameter for time constant. The “slow” setting helps to dampen noise, but can end up distorting sharp peaks and is intended for use with packed columns and broader Megabore peaks. The “fast” setting allows sharp peaks to maintain their original shape, with a sacrifice in noise. This is helpful with narrow bore columns, especially with high injection split operations.



The “fast” setting for Varian 3800 detectors activates a 4-point Finite Impulse Response filter, as discussed above. A “slow” setting adds in a 16-point moving boxcar average with this filter.



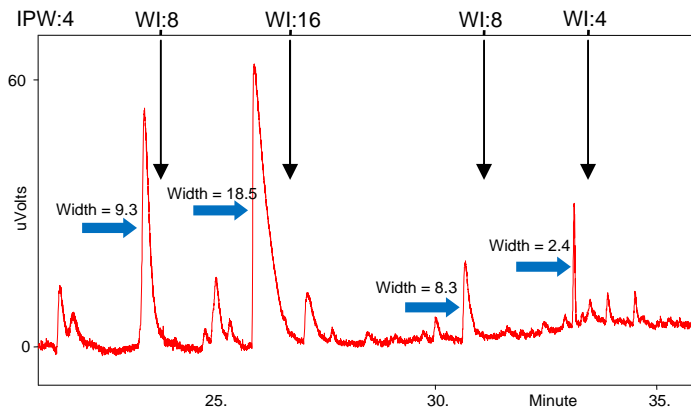
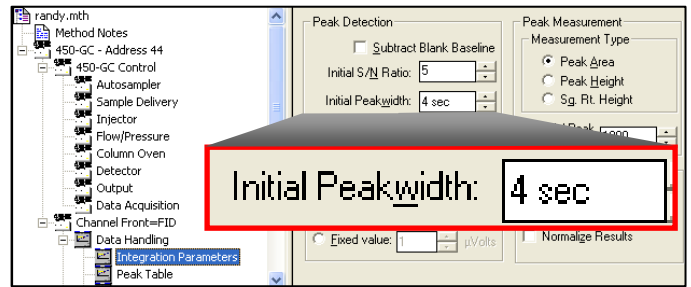
**The choice between Fast and Slow Time Constants with Varian 3800 provides subtle enhancements for broad peaks (slow) and for sharp peaks (fast). Both peak-to-peak noise and Root Mean Square (RMS) noise are decreased by a factor of 2.5 with a change from “fast” to “slow”.**

A change was made in Varian 450 firmware programming with the upgrade to Version 1.07. Prior to this version, filtering was as described above. After Version 1.07,<sup>4</sup> the “Slow” time constants became inactive for all detectors and reverted to the same “filtering” as Fast, and any filtering is now intended to be performed by Data Bunching Rate instead, as discussed above.

<sup>4</sup> Firmware Version 1.07 started shipping from Varian in April 2008, and older versions can be reflashed to bring up to current release.

## Initial Peak Width

The peak detection algorithm is set up to be able to eliminate unexpected peak shapes. For example, if peaks are anticipated to be sharp from “fast” chromatography, then broad baseline drift or background humps should be ignored. At the other end of the scale, if peaks are predicted to be wide, then sharp “spikes” are to be considered as disruptions, and thus disregarded. The operator must set what sort of peaks is expected by choosing an initial peak width. The preset value is 4 seconds, with a range from 0.5 to 256 seconds. This value should be chosen by viewing a report from a typical chromatogram for the specific analysis and noting the peak width at ½ height for the first peak.



Initial Peak Width, the Peak Width (WI) for peak detection is altered as appropriate only by a factor of two (up or down), and continues to be updated as peak widths are sensed to changes. For example, the chromatogram at left has an initial peak width of 4, and after a peak is found with a wider width, this parameter setting is altered to a peak width of 8, and then 16, then 8, and then to 4, as the chromatogram progresses.

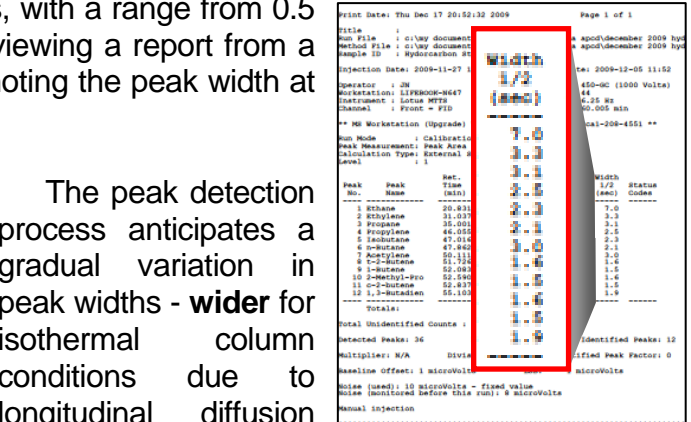
The purpose of this change in WI values is to attempt to maintain 10 to 20 points across the top of the peak, judged to be the optimum number to fully define the shape of the peak. If more than 20 points, then too many are involved and noise is increased. If less than 10, then the peak detection algorithm may not correctly compute this area with insufficient points across the peak. To bring the points across the top of a peak within the optimum range, the expected Peak Width establishes an adjustment to an averaging process to ensure enough points are employed in the peak detection. Bunched Points are combined together into Cluster Points by a simple arithmetic average of Bunched Points, based on the change in expected Peak Width. The process involves clustering bunched points together based on the expected Peak Width and Bunch Rate by the formula:

$$\frac{\text{Bunched Points}}{\text{Cluster Point}} = \frac{\text{Conversions/sec}^5 \times \text{expected Peak Width (sec)}}{\text{Detector Bunch Rate (conversions/bunch point)} \times 10 \text{ Cluster Points}}$$

Conversion to Cluster Points is a simple arithmetic average for the set number of Bunched Points.

The actual number of Cluster Points across a peak is:

$$\text{Number of Cluster Points}_{\text{peak}} = \frac{\text{measured Peak Width} \times 10}{\text{expected Peak Width}}$$

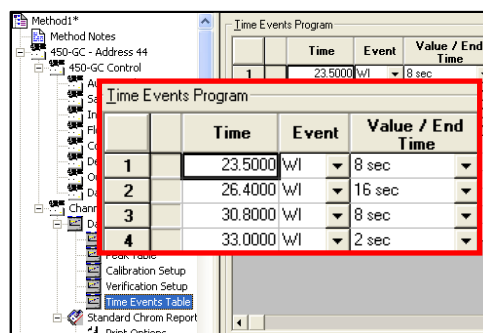


<sup>5</sup> For the Varian 3800, this conversion rate is 40 Hz (conversions/second), and for the Varian 450, it is 100 Hz.

An example of the process is illustrated (from chromatogram on previous page) with an Initial Peak Width (IPW) of 4 seconds and a Bunch Rate of 10 conversions/bunched point (for a Varian 3800). Then a peak of 4 seconds will have 10 points across its top and the Cluster Point remains the same at 4 Bunched Points. Then if the next peak as a width of 9.3 seconds, it will have 23 Cluster Points across its top - more than the optimal range of 10 to 20; as a result, the expected Peak Width doubles for the next peak - to 8 seconds. The next peak, with a width of 18.5 seconds, will have 23 Cluster Points, and now the expected Peak Width doubles again to 16. When the next peak shows up with a width of 8.3 seconds, only 5 Cluster Points define the top of the peak and the Expected Peak Width drops back to 8 seconds.

To be sure spikes do not cause dramatic alterations in the expected Peak Width, a change in its value can be only factors of two for each peak update - either up or down. This change always occurs after the affected peak has come back to baseline.

This process to update peak detection parameters always occurs after a peak width change is sensed, rather than prior to the subject peak. To ensure that all peaks are properly integrated, expected Peak Widths can be time-programmed so that this change can occur before the subject peak elutes, to provide better assessment of peak area. Based on an initial run of a mid-level standard, reported peaks widths can be set to occur prior to elution of the peak, to more accurately represent the sort of chromatography being generated. With this programmed entries for expected Peak Width, the automatic updating is deactivated.



**Time-programming WI parameters sets the expected Peak width for individual peaks and disables the automatic update process.**

The preset value for expected Peak Width is 4.0 sec, with a range from 0.5 to 256 seconds.

The precise setting for this parameter is not very critical, as Varian Star Workstation readily locates a peak even when expected peak widths do not correlate precisely with the actual widths of detected peaks. Two examples from actual data sets are shown in tables below.

**Peak with 0.6 sec Width<sub>1/2</sub> height**

WI Parameter	Raw Area Counts	Relative Area
<b>0.5</b>	11165	<b>100%</b>
<b>1</b>	11146	<b>100%</b>
<b>2</b>	11118	<b>100%</b>
4	10580	95%
8	9919	89%
16	7302	65%
32	3871	35%
64	1979	18%
128	not reported	not reported
256	not reported	not reported

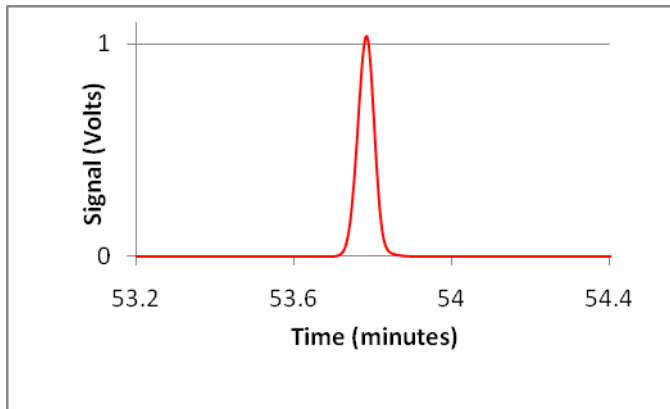
**Peak with 18.7sec Width<sub>1/2</sub> height**

WI Parameter	Raw Area Counts	Relative Area
<b>0.5</b>	34824	96%
<b>1</b>	34824	96%
<b>2</b>	34824	96%
4	34824	96%
8	35928	99%
<b>16</b>	36246	<b>100%</b>
<b>32</b>	36427	<b>100%</b>
<b>64</b>	36365	<b>100%</b>
<b>128</b>	36399	<b>100%</b>
<b>256</b>	36275	<b>100%</b>

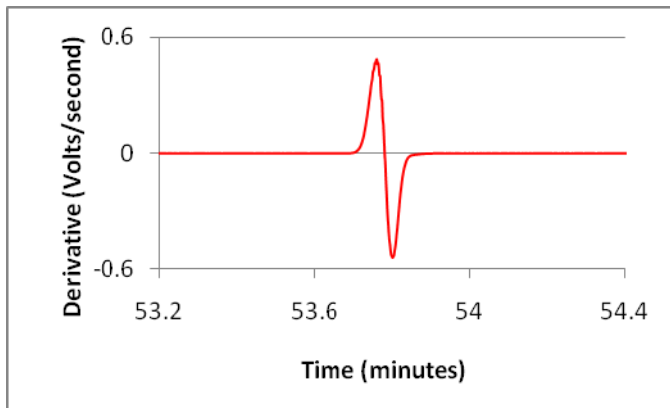
## Detection of Peak Start

Precise details of the peak detection algorithm in Varian Star Workstation are proprietary. The concepts discussed here are intended to provide a basic understanding of this process to improve the quality of the final results. User-selectable parameters for peak detection can have major impacts of the appropriateness of locating peak start and end points.

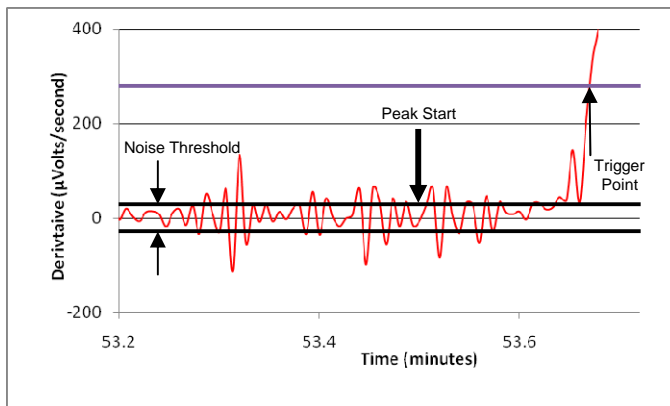
After a chromatogram is run, its first derivative is performed, using Cluster points, to determine peak starts. Based on the Noise threshold from the Noise Monitor interval prior to the start of the run and the Initial S/N Ratio, a peak start point is triggered when the derivative value first exceeds this value.



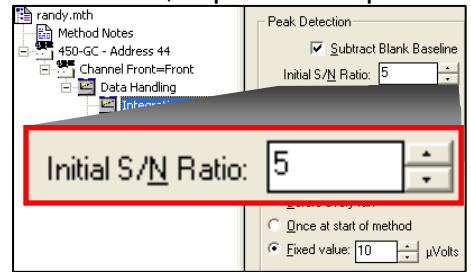
**A representative peak is shown as an example for peak processing.**



**The first derivative is computed for this peak to provide data points for setting the trigger point for Peak Start.**



**An expanded region in the peak front of the derivative shown above illustrates locations for the trigger and actual Peak Start.**



This action point is:

$$Trigger = Noise \times (S/N \text{ Ratio})$$

S/N Ratio can be time-programmed to optimize detection of specific troublesome peaks. The available range of values is 1 to 256, incremented by factors of 2.

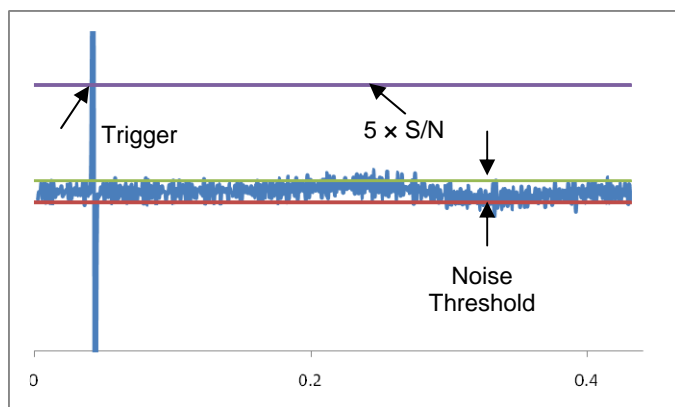
When the expected Peak Width changes, the Noise Threshold value is altered to allow for the change in the trigger point location for narrower or broader peaks. For example, if the Initial Peak Width is 4 seconds, and the automatic update changes the current expected Peak Width to 8, then the Noise Monitor value is cut in by  $1/\sqrt{2}$ , to account for the change in noise. If the expected Peak Width becomes 2, then this value is increased by  $\sqrt{2}$ .

The last 60 Cluster Points prior to this trigger point are examined backwards to find a local minimum, which is then taken as the Peak Start. This step ensures that the true peak start is determined into the preceding noise and the computed area includes an area slice under the whole peak that would have been excluded if the peak start had been set at the trigger point.

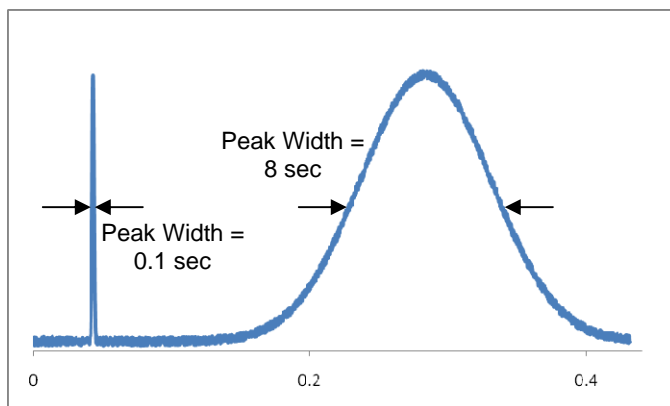


To illustrate the process for peak detection, several case studies demonstrate how expected peaks are reported and undesired peaks are ignored.

If very narrow peaks are anticipated, then appropriate parameters, specifically expected Peak Width and Signal-to-Noise, need to be set to pick out the sharp peaks and disregard baseline drifts and unexpected broad peaks. Then when the derivative of the signal is performed, the target peak is properly detected



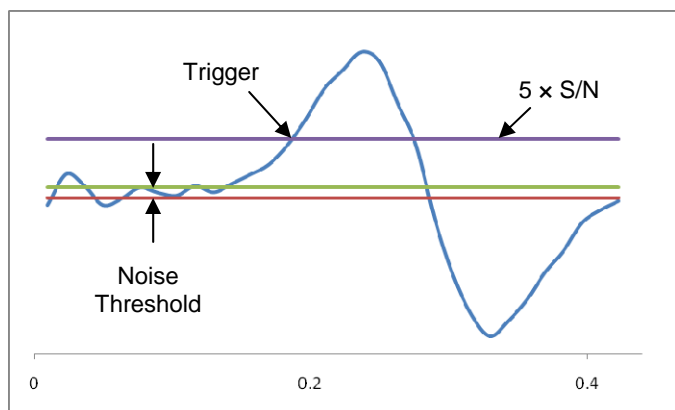
**The derivative of this chromatogram readily picks out the narrow peak as it easily exceeds the trigger point. The broad peak barely breaks the Noise threshold and is not detected.**



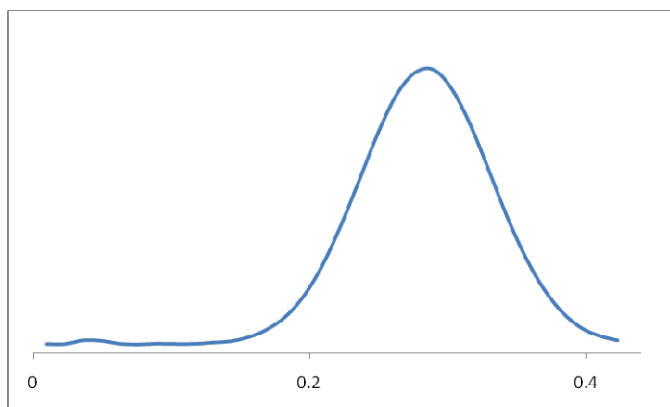
**A very narrow peak is synthesized along with a broad peak. With an expected Peak Width of 0.5 sec, the narrow peak is targeted and the broad peak ought to be treated as “uninteresting”.**

and the broad “baseline drift” is ignored. With judicious choice of Expected Peak Width (0.5 sec) and Signal-to-Noise (5), only the sharp, narrow peak breaks the Signal-to-Noise Threshold; the broad peak does not come close to the Trigger level and is disregarded and not reported.

Then, when broad peaks are anticipated, expected Peak Width is widened to match the desired peak, and the consequential averaging often completely suppresses narrow peaks, as noise, and boosts detection of the wider target peak.



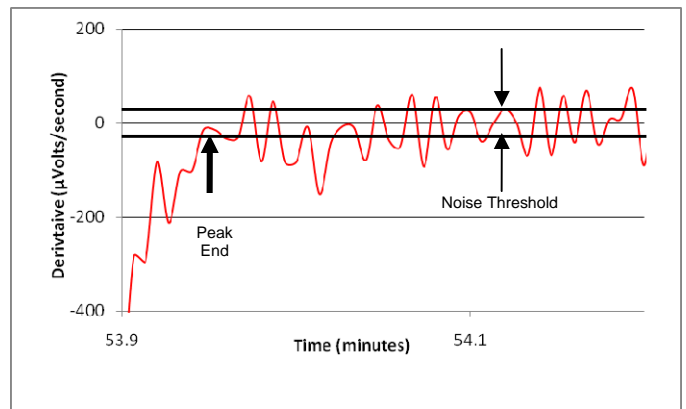
**A derivative of the chromatogram, with a  $WI=8$ , gives a different sense to possible peak starts, as the narrow peak is hardly sensed, and noise on the broad peak is radically suppressed and its derivative readily exceeds the threshold.**



**When broad peaks (width @  $\frac{1}{2}$  height of 8 sec) are expected, the expected Peak Width can be set to match their widths ( $WI=8$ ) and sharp peaks will melt into the noise.**

## Detection of Peak End

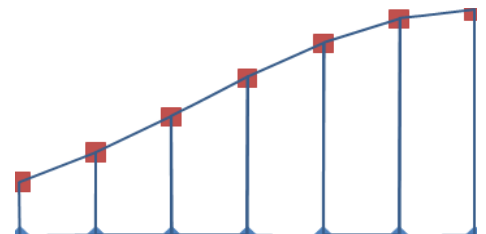
Peak end is detected when the derivative level returns into the Noise Threshold window. If this signal promptly zooms in a positive direction, then a fused peak is detected and a judgment is made with TAN% to determine if the rider peak requires a perpendicular drop to baseline or a tangent skim (see discussions below).



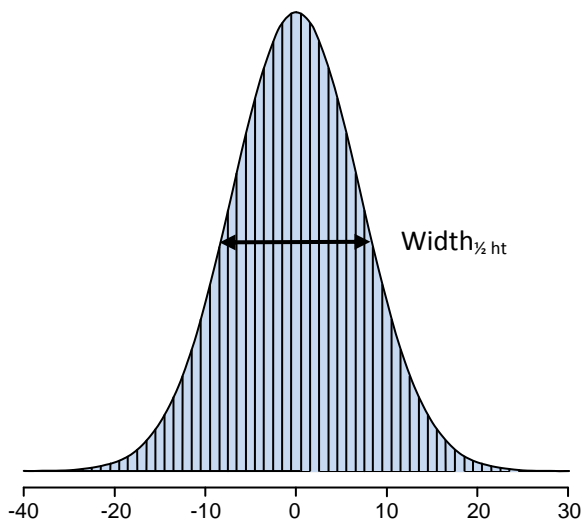
**Peak End point is set when the derivative signal returns into the Noise Threshold window.**

## Computing Peak Areas

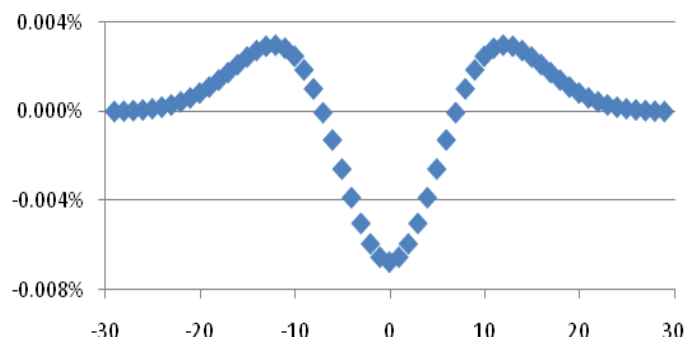
Varian Workstations compute areas for detected peaks by a summation of trapezoid areas found between raw data points in the curve and the assigned baseline. This simple, but elegant, approach represents accurately the true shape of the peak, particularly if sufficient data points are collected to give at least 10 points across the top of the peak above the width at half height.



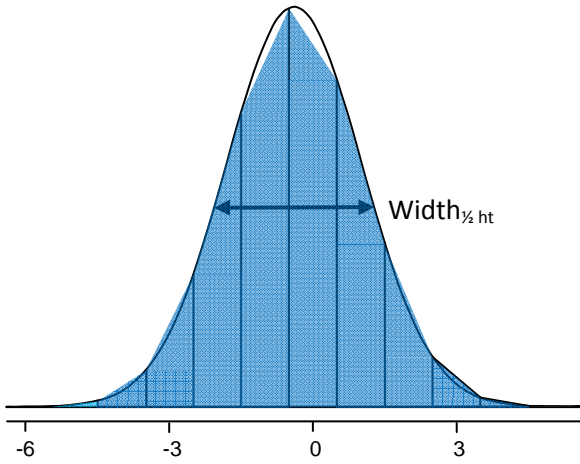
**True area of a Gaussian curve simulated by summation of trapezoids.**



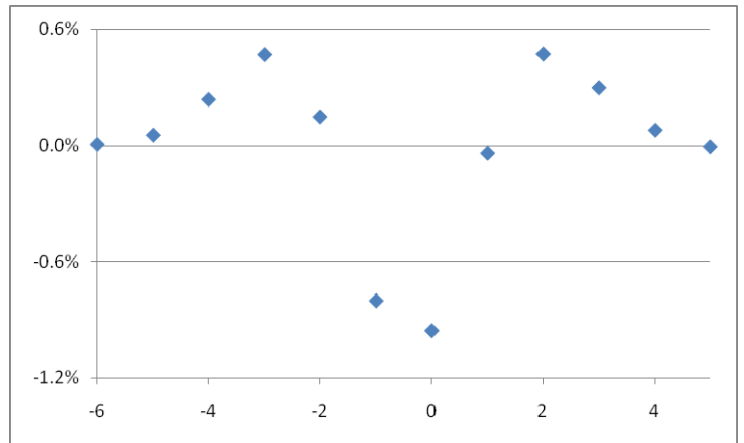
**Summation of Trapezoids with 16 points across top of peak.**



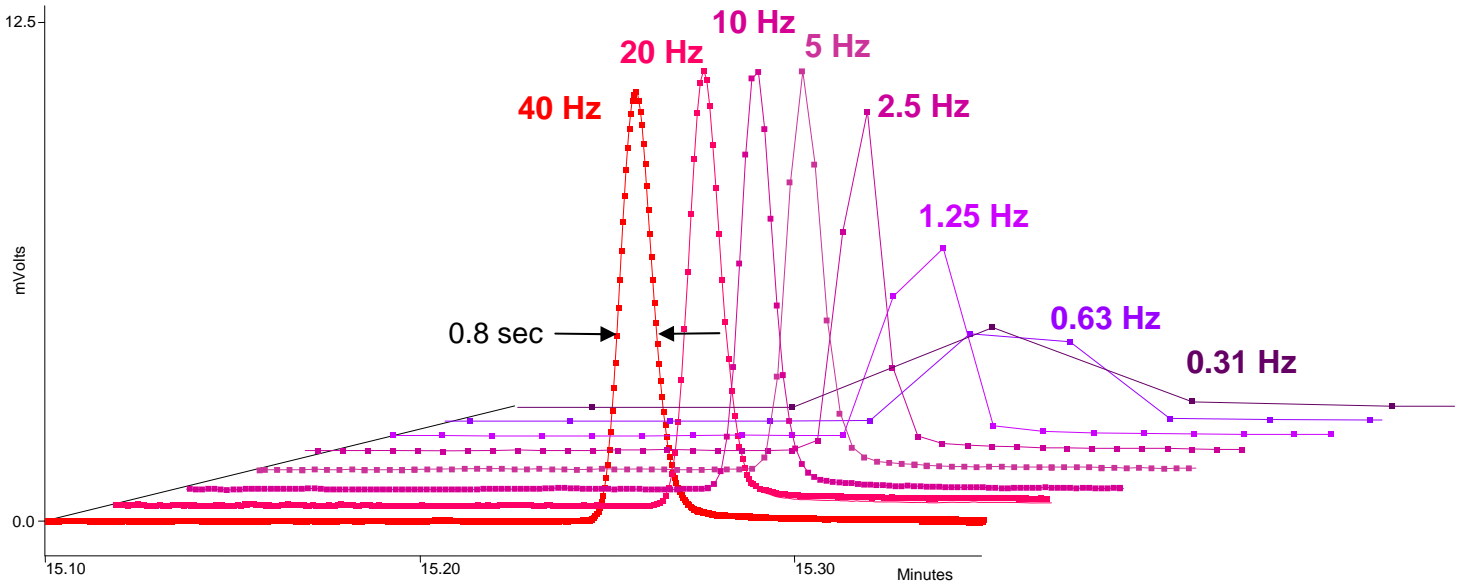
**Percent deviations of area by trapezoids compared with true Gaussian curve, with 16 points across top. Total area by trapezoids is 0.99999 of true value**



**Summation of Trapezoids with 3 points across top of peak, with top data point not at apex of Gaussian peak.**



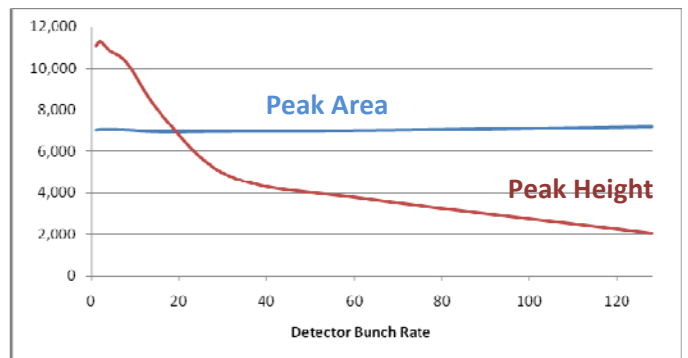
**Percent deviations of area by trapezoids compared with true Gaussian curve, 3 points across top. Total area by trapezoids is 0.99986 of true value.**



**Changes to Detector Bunch Rate have dramatic effects on peak shapes, but remarkably little effect on peak areas.**

Surprisingly, areas are measured quite nicely with dramatically fewer data points than anticipated. Typically, at the apex region of the peak, areas by trapezoids are underreported, whereas areas are overvalued at the leading and trailing edge of the peak. These errors often cancel out, and the area is very close to the true value.

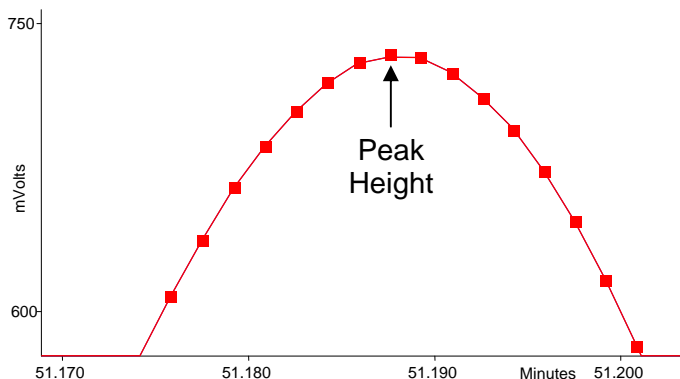
One strong advantage of slower data rates is the related reduction in noise, as more raw data points are represented within a single “bunched” point.



**Plots of peak areas and heights illustrate consistency of peak areas over the range of Bunch Rates.**

## Computing Peak Heights

The amplitude of the peak apex is determined as the higher of the two corresponding bunched data points used in picking out the peak retention time, after adjustment for baseline. This peak height point does not necessarily match the point used for retention time (see Determination of Retention Time above). Unit for the reported value is Microvolts.



## Least Significant Bit (LSB)

Area counts or height counts in a Star Workstation report are based on an arbitrary characteristic of the data generator, called the Least Significant Bit (LSB). Its reported value represents the smallest data increment that can be measured. Height counts are expressed in Least Significant Bit; Area counts are expressed in Least Significant Bit X seconds.

65	71.370	0.000	605	BV	4.7
66	71.770	0.000	677372	BP	2.8
67	72.035	0.000	350	PE	0.0
68					

-----	
Total U	LSB: 1 microVolts
Detecte	
Multiplier: N/A	Div
Baseline Offset: 0 microVolts	LSB: 1 microVolts
Noise (used): 25 microVolts - fixed value	
Noise (monitored before this run): 28 microVolts	

To express the results in microvolts (height) or microvolts-sec (area), reported counts must be multiplied by the reported Least Significant Bit value. The Least Significant Bit value is shown in a Standard Report, and for Varian 3800 and Varian 450 Gas Chromatographs is typically equal to 1 microvolt.

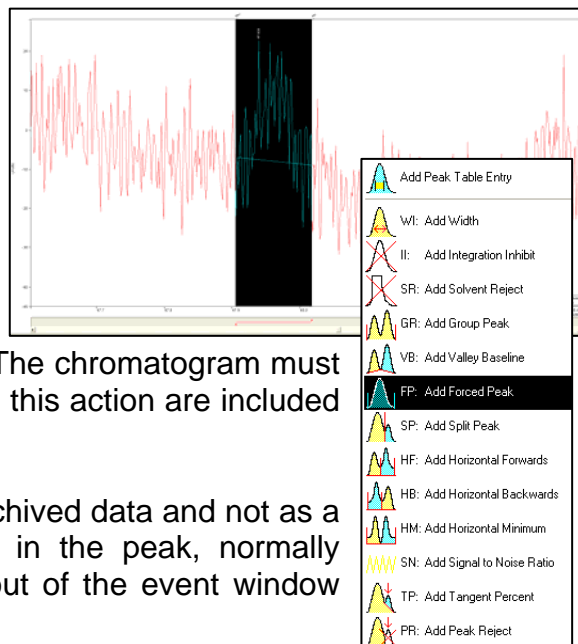
## Automatic Peak Detection

The setting of starting and ending points for a chromatographic peak can have significant impact on the computation of its area and heights, and could result in inaccurate and imprecise quantitative results. A myriad of factors enter into this process, including parameters within the Star Workstation method. Fortunately, preselected values in the method are often correct for most chromatograms. However, occasionally preset choices turn out to be inappropriate for specific peaks. Adjustments can be performed through **Star Interactive Graphics** to visualize the impact of changes, and then the altered method can be saved so that all subsequent runs of a similar type can have these same updates applied. A correct choice of parameters will greatly minimize subsequent manual manipulations of the chromatogram,

Major peaks in a chromatogram normally do not have difficulty in having their start and end points located, or if these spots are not valid, their wrong positioning likely will have very little impact on area assigned. The problem becomes significantly more important when dealing with small peaks approaching detection limits. Subtle movement in the starting and ending sites can yield major changes in the area assigned.

## Forced Peak

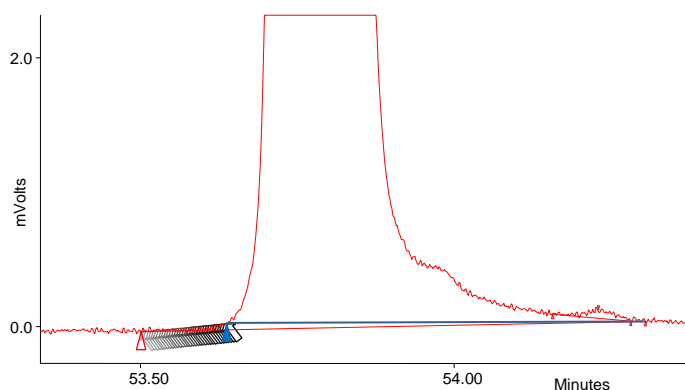
A choice to create areas for peaks not assigned a baseline, and thus not integrated, is use of the time-programmable event “Forced Peak - FP”. This parameter is accessed in **Interactive Graphics** and in **Method Editor**. A right mouse click in the Time Events window opens a menu for all events. Selecting “FP” inserts a bar that can be moved around and resized by the mouse to adjust the peak start and end as needed. The chromatogram must be reintegrated for this event to take effect. Results from this action are included with final results.



This parameter should be employed only singly on archived data and not as a regular part of a method, as any retention time shift in the peak, normally expected in chromatography can move the new peak out of the event window and radically assign its area improperly.

## Manually Move Start and End Points

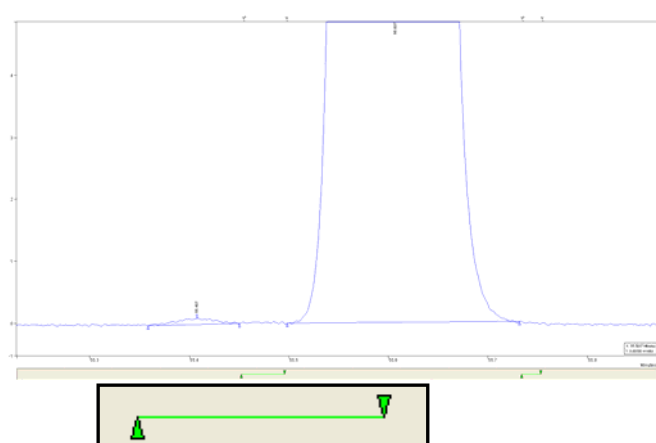
If start and end points are not set appropriately, they can be manually readjusted through **Interactive Graphics** by grabbing the offending triangle with the mouse pointer and dragging it a new desired location. New peak points are indicated by solid red triangles, and original points remain open centers. The new area is generated by “Processing” the run, and results are reported in the final report.



This operation will only apply to the active chromatogram and will not be part of any method and will not impact any other chromatogram.

## Integrate Inhibit

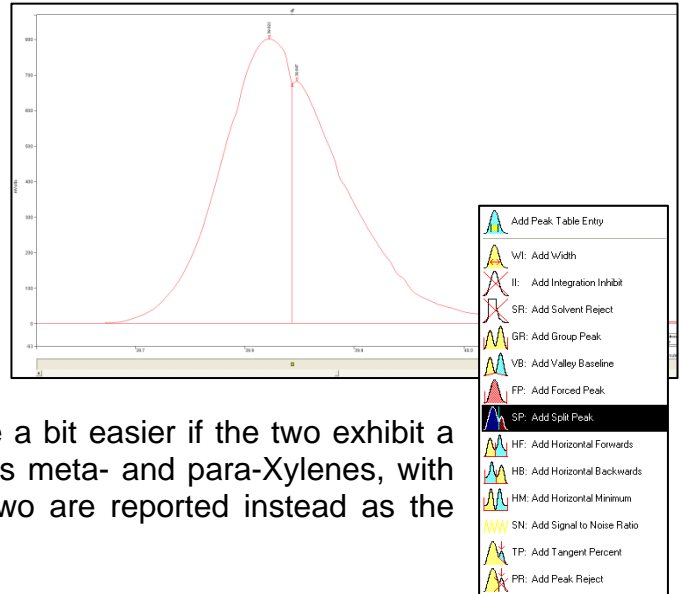
Sometimes the automated process to locate peak start and end points are not quite set as the user would have chosen. One mechanism to force operations is to utilize the time-programmable event “Integrate Inhibit - II”. By placing this event before and after a peak, all data points inside each span are not considered, and start and end points of the event always become possible baseline points. Established peak detection settings still apply before and after this action



If retention shifts are not dramatic, this event can be set into the method and be applied to all subsequent runs, as peak detection rules still apply for the span between two of these events. However, if the chromatography shifts, this event can severely impact the assignment of baselines and yield wrong areas for the peak.

## Split Peak

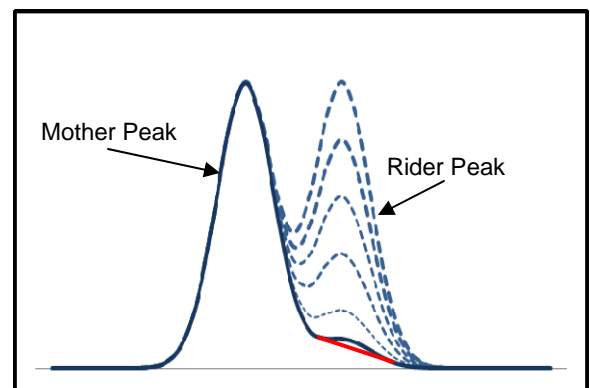
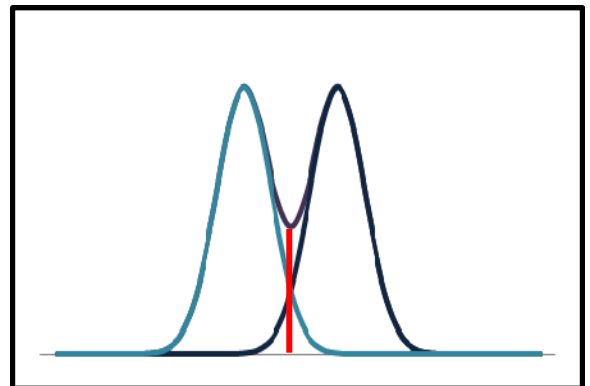
Occasionally peaks occur very close together. The chromatography employed can make their individual areas appear as one peak, maybe with a hint of splitting, but not enough to be picked up as separate peaks by the chosen parameters. The time-programmable event “Split Peak - SP” can be added in to force the peak into two areas. Use of this function can help divvy up areas into the two peaks. This action can be a very subjective call on where to place this event, made a bit easier if the two exhibit a hint of a valley between them. The notable pair is meta- and para-Xylenes, with nearly the same chromatography. Often these two are reported instead as the combined area and labeled as “m+p-Xylenes”.



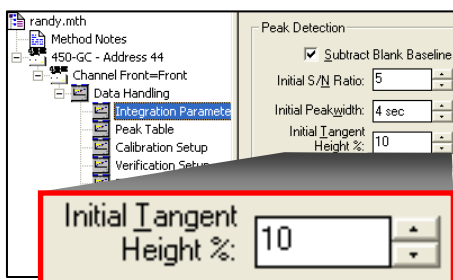
This parameter should be employed only singly on archived data and not as a regular part of a method, as any subtle retention time shift normally expected in chromatography can move the new peaks out of the event window and radically assign their area improperly. The chromatogram must be reprocessed for this event to be executed.

## Tangent Percent

Overlapping peaks can introduce some significant errors during allocation of areas between peaks. If two peaks are equal in size and overlap, a simple drop to baseline at the valley point between the two will allow a reasonable assessment of the areas. However, as the rider peak fades compared with its mother peak, at some point a perpendicular drop does not properly assign areas. The tail of the mother peak adds too much to the area of the rider. Instead, a tangent skim should be applied to the rider peak to provide its “baseline” for computing its area. The triangle under the rider is added to the mother peak’s area. The decision point for switching between perpendicular drop and tangent skim is based on the parameter Tangent Percent - TP. This compares the height of a rider peak to the height of its mother. If the ratio exceeds the Tangent Percent value, a perpendicular drop is executed. If less than this value, a skim is undertaken and marked as “Tangent Skim - TS or Tangent Fused - TF”. This setting can be time-programmed as needed.



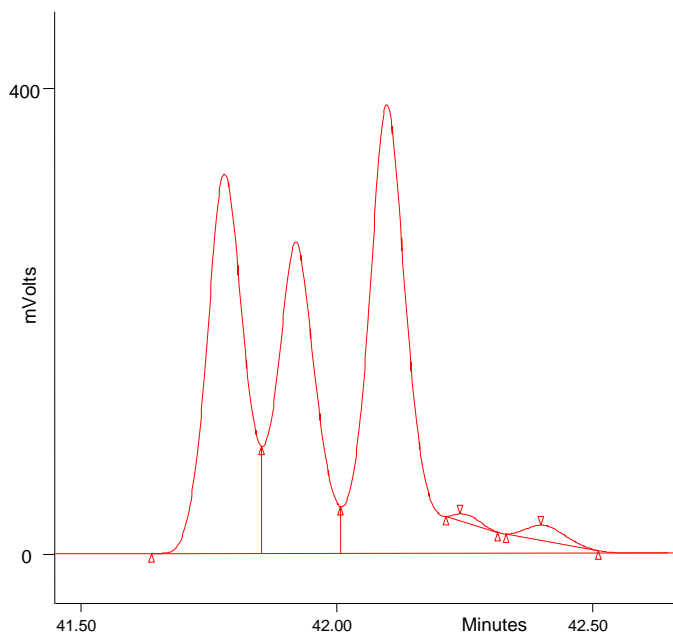
The preset value is 10%.



This allocation of areas, at best, is a crude approximation, but at least small rider peak areas will not be grossly over-exaggerated with inclusion of a portion of the mother peak, and yet when the rider is large, the areas for the two peaks can be reasonably assigned.

How peaks are treated is listed as “Separation Code in **View Results** and in reports.

Baseline Codes	Peak Onset	Peak End
BB	Baseline	Baseline
BV	Baseline	Valley
VV	Valley	Valley
VB	Valley	Baseline
TF	Tangent Fused	
TS	Tangent Skim	
PB	Perpendicular Drop	Baseline
BP	Baseline	Perpendicular Drop
VP		
PV		
GR	Grouped Peaks by Time	
HF	Horizontal Baseline Forward	
HB	Horizontal Baseline Backwards	

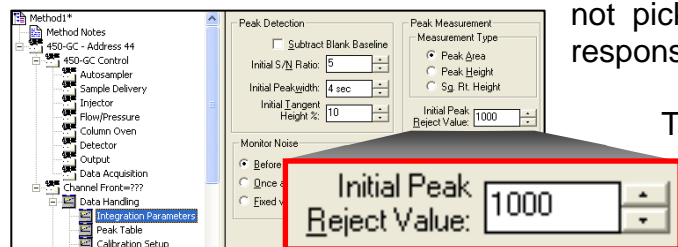


**Representative case where overlapping peaks have their areas assigned from a perpendicular drop to baseline and tangent skimming**

## Peak Reject Value

Peak areas below an established detection limit are not likely to be reportable and can be excluded from consideration in quantitation. As the Varian Star Workstation can eject low areas, this parameter becomes a criterion for eliminating small peaks from results. Preset value is 1000 and has a range from 0 (no rejection) to 1,000,000 [sic].

A valid entry should be slightly lower than the method detection limit for that compound. However, when performing a multi-point calibration, especially with Reference Peaks, this Reject Value may need to be increased only for the calibration series to ensure that small noise peaks are not picked out as the analytes and retention times and response factors are mistakenly altered.



This parameter, labeled “PR” can be time-programmed to suit particular chromatographic conditions.

## Guidelines for Setting Optimum Parameters for Peak Detection

1. Collect data for a midrange standard with chromatography fully optimized.
2. Set expected Peak Widths for each compound to match peak widths at  $\frac{1}{2}$  height from the report for midrange standard.
3. Set data rate to enable at least 10 data points across top of the narrowest peak in the chromatogram.
4. Set Identification Reference Peaks for compounds present in every sample, typically Internal Standards and Surrogates.
5. Visually check every peak in the midrange standard to ensure that peak starts/ends and peak baseline assignments are appropriate.
6. Set Peak Reject Value just below lowest standard area counts.
7. Rerun or recalculate a midlevel standard to reset expected Retention Time windows.
8. Then SAVE method.
9. Reset Peak Reject Value to just below established detection limit and resave method.

**This author thanks Les Ehrlich of Varian, Inc.  
for assistance in preparing this monograph.**

Screens are copyrighted by Varian Inc.,  
and are reprinted (reproduced) with the permission of Varian, Inc. All rights reserved.  
Varian and the Varian logo are trademarks or registered trademarks of Varian, Inc.  
in the U.S. and other countries.

---

---

**Lotus Consulting**

310/569-0128 Fax 714/898-7461  
email ebramstoncook@msn.com



5781 Campo Walk  
Long Beach, California 90803