

Functional Data Analysis – “What to do when your data are a curve or spectra”

Ryan Parker, PhD

Research Statistician Developer

ryan.parker@jmp.com

Clay Barker, PhD

Research Statistician Developer

clay.barker@jmp.com

Tom Donnelly, PhD, CAP

Systems Engineer

tom.donnelly@jmp.com

DATAWorks 2025

Short Course

April 22

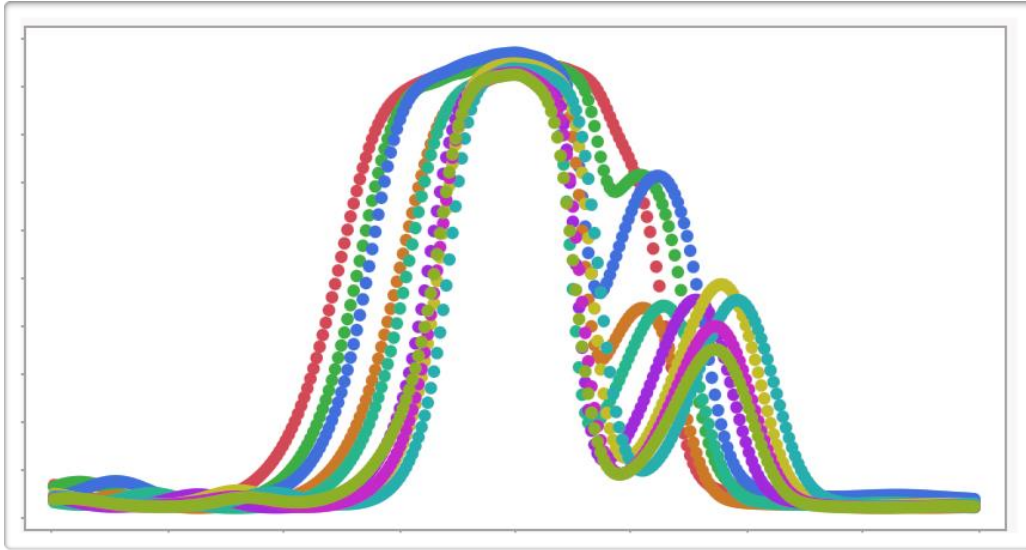
IDA



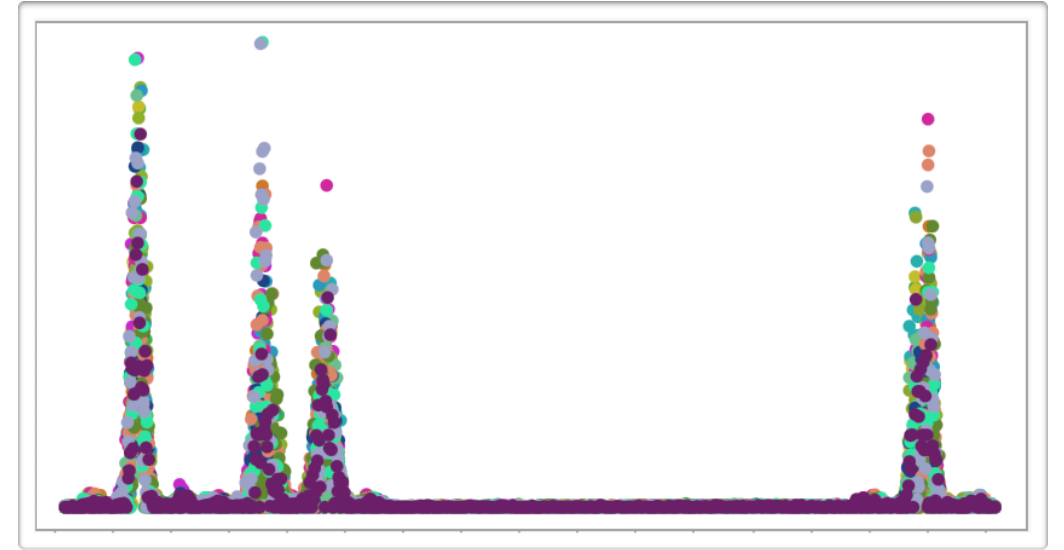
Understanding a Collection of Curves Through Functional Decomposition

What's new in
Functional Data Explorer
in JMP Pro 18

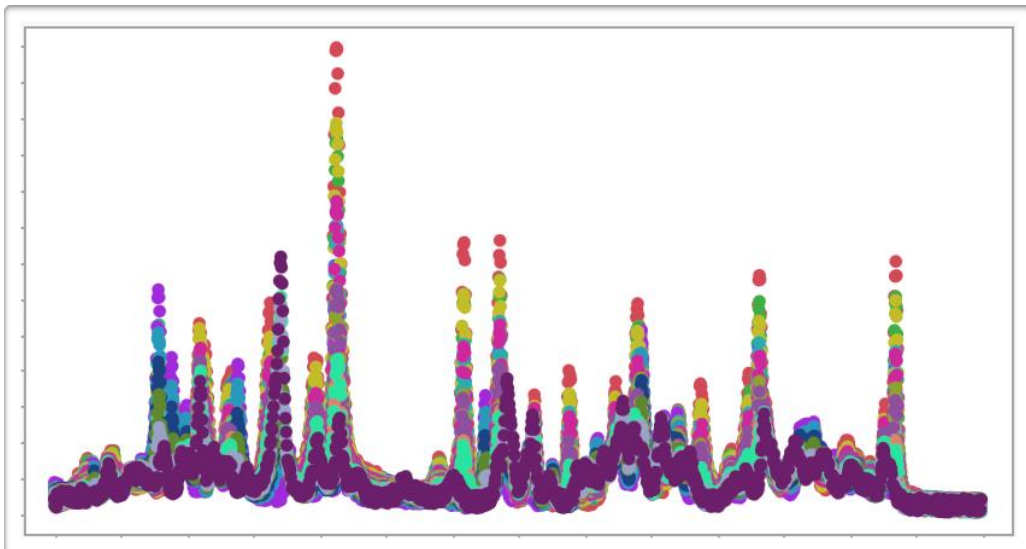
Ryan Parker and Clay Barker



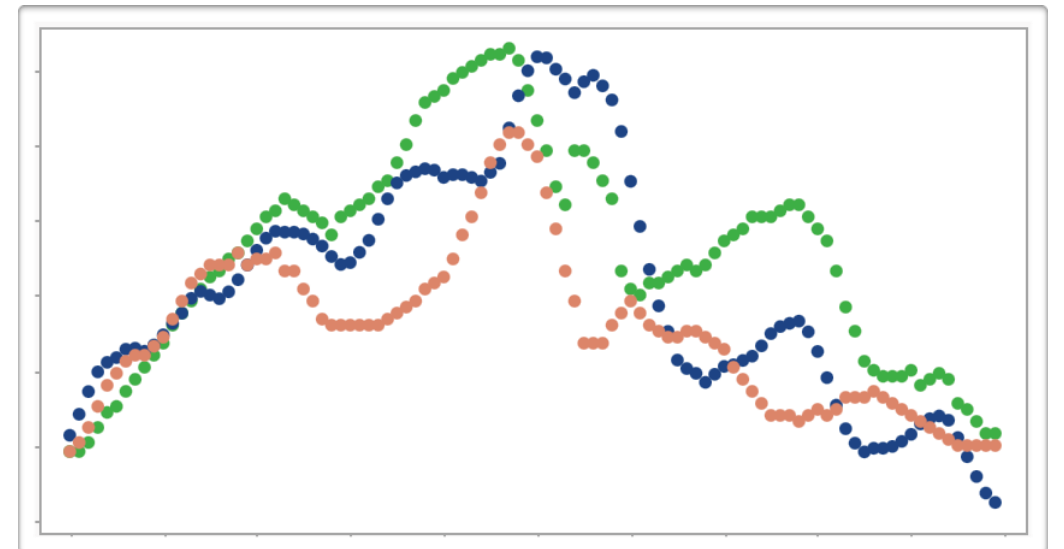
Chromatography (HPLC)



NMR Spectroscopy



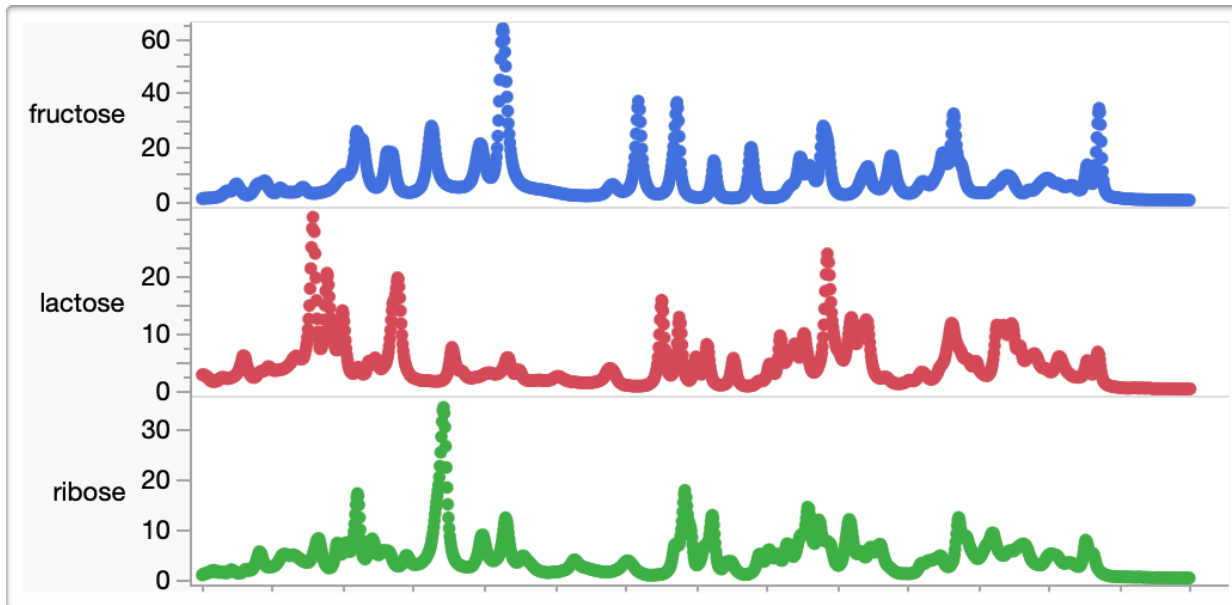
Raman Spectroscopy



Theme Park Wait Times

Direct Decompositions (MCR)

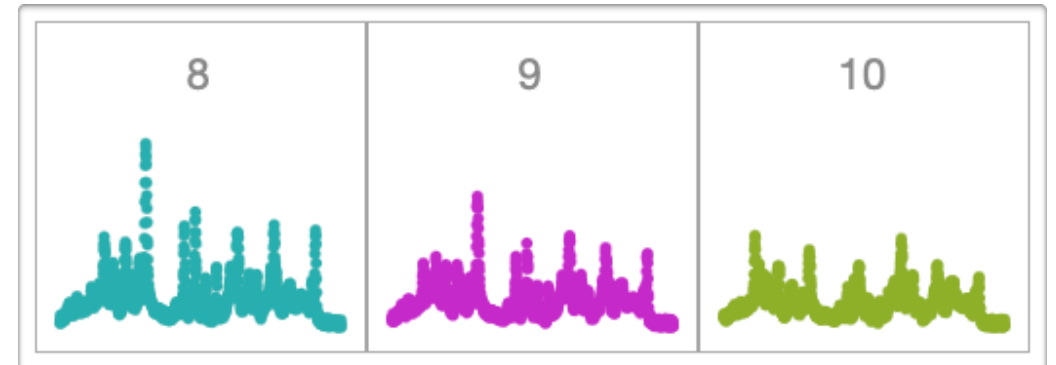
Carbohydrate Raman Spectroscopy



Component Functions

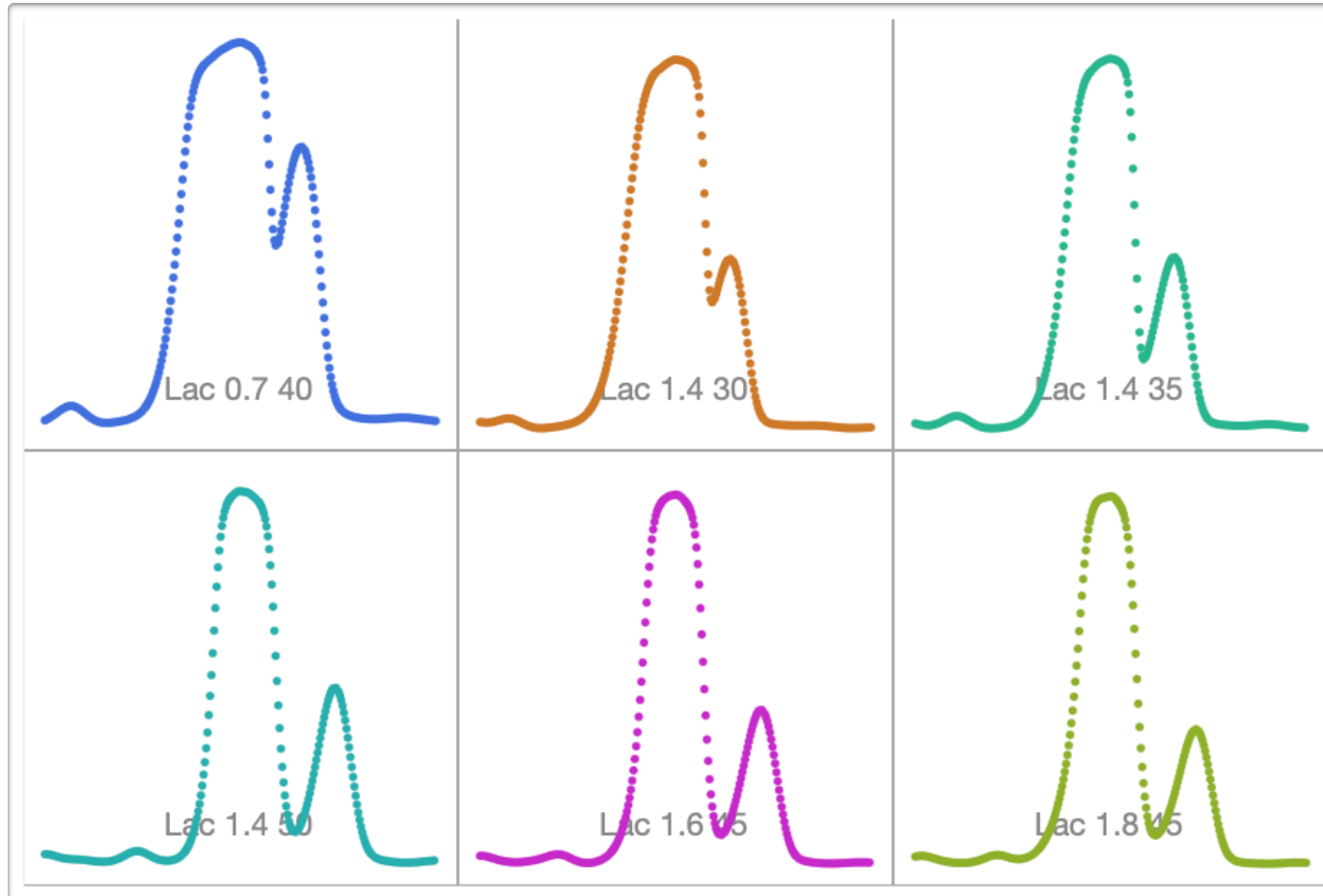
	component 1	component 2	component 3
8	0.6	0.2	0.2
9	0.4	0.4	0.2
10	0.2	0.6	0.2

Mixture Components



Observed Spectra

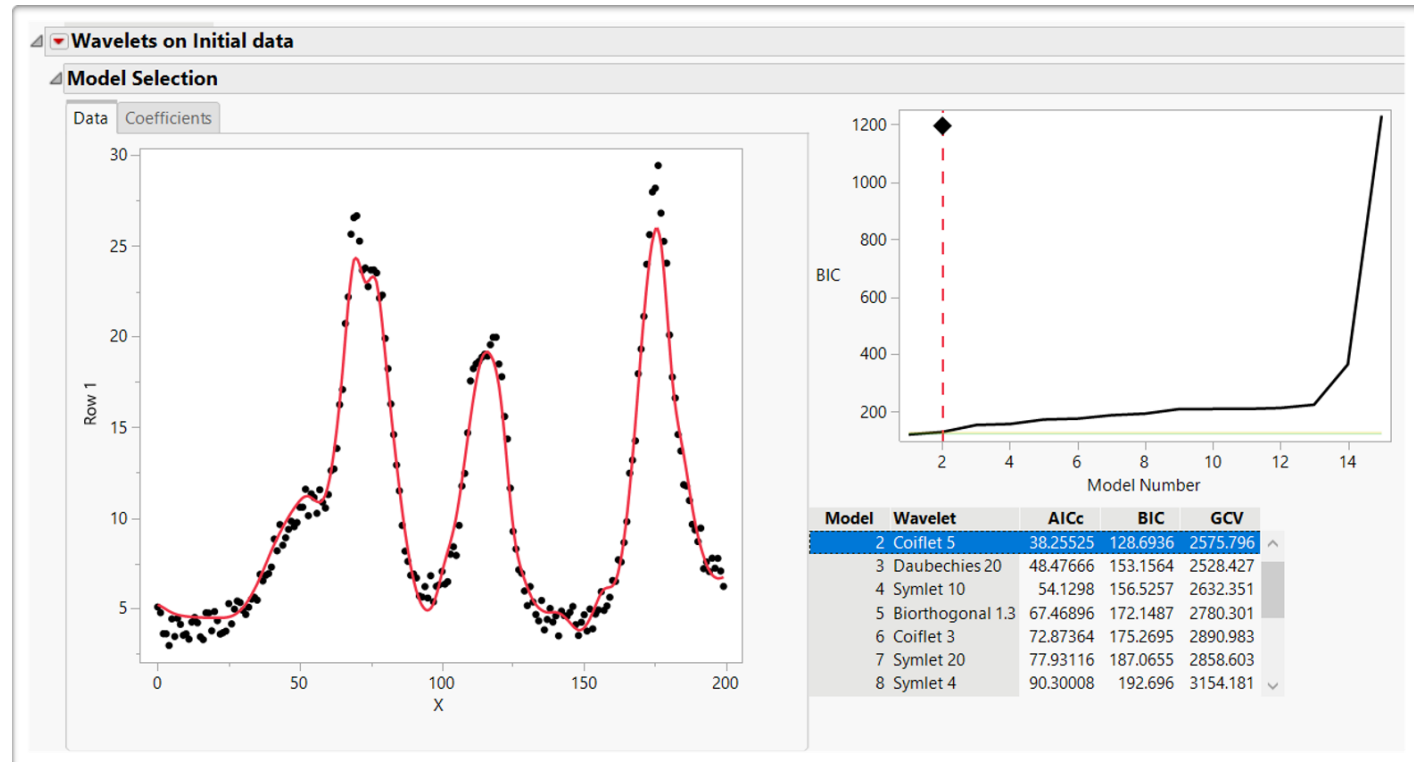
Peak Finding



Chromatography (HPLC)

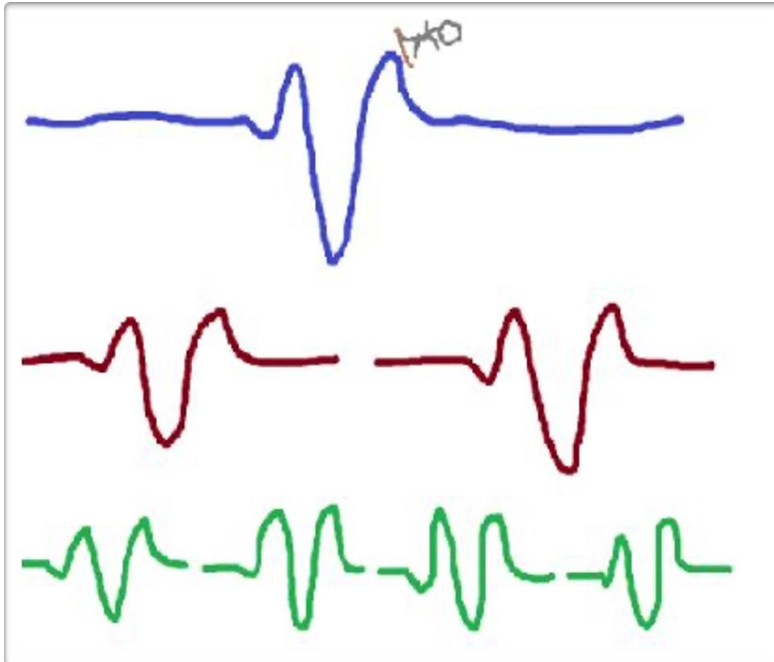
Wavelets and Peaks

- In JMP 17, we added wavelets as a modeling/smoothing option in FDE.
- Smooth
- Flexible



Wavelets and Peaks

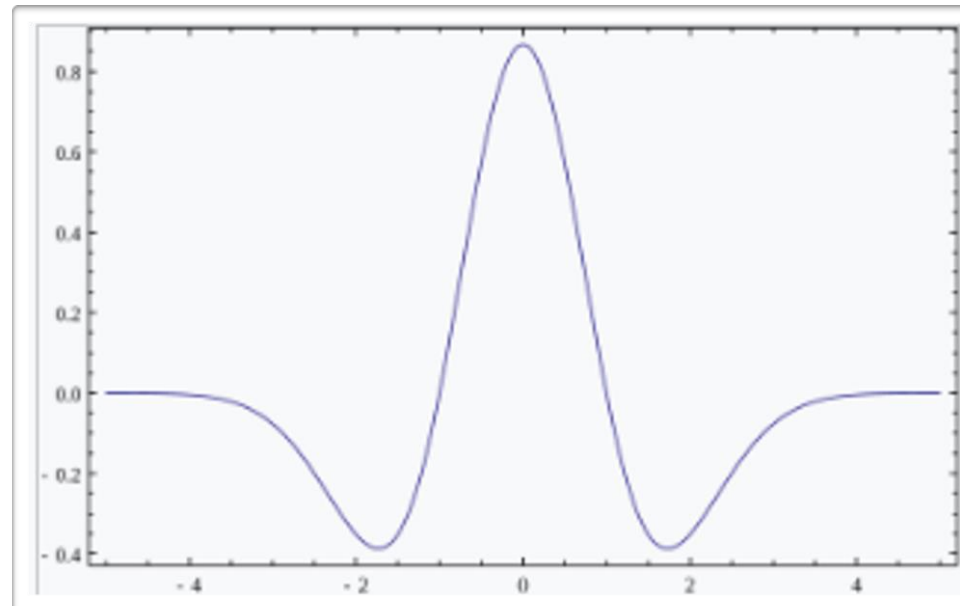
- The discrete wavelet transform repeatedly breaks the data into sections.



- Each section gets a corresponding coefficient, some can be zero.
- To get our smoothed function: multiply each section by its coefficient and add them up.
- Large coefficients correspond to areas where a function changes a lot.

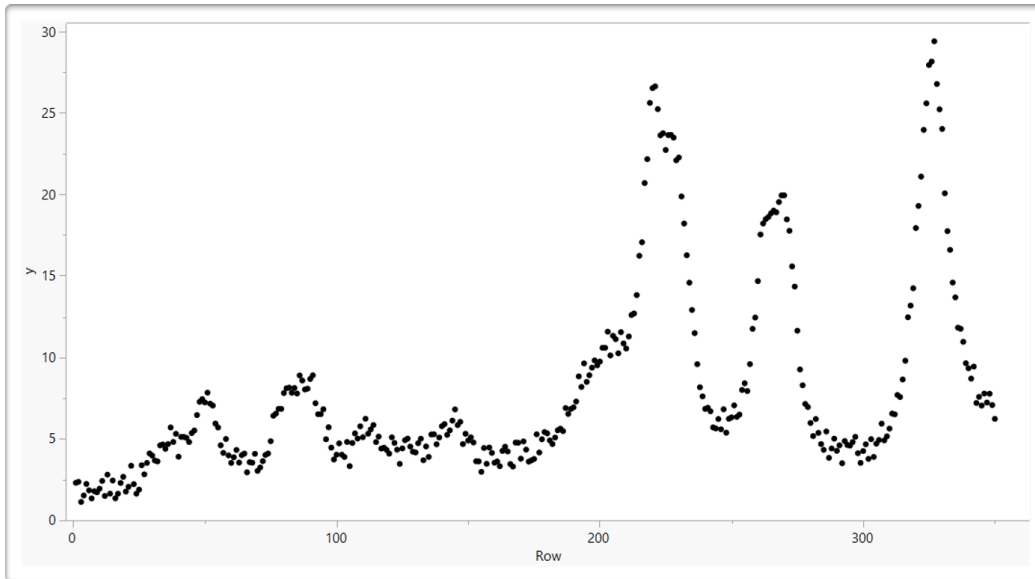
Wavelets and Peaks

- There is also a continuous wavelet transform.
- We won't get into the details but imagine representing our function by stretching and shifting a function like below.



Wavelets and Peaks

- Scaling and sliding gives us a matrix of coefficients X .



location	b_1	b_2	b_3	...
a_1	x_{11}	x_{12}	x_{13}	...
a_2	x_{21}	x_{22}	x_{23}	...
a_3	x_{31}	x_{32}	x_{33}	...
a_4	x_{41}	x_{42}	x_{43}	...
a_5	x_{51}	x_{52}	x_{53}	...
a_6	x_{61}	x_{62}	x_{63}	...
\vdots	\vdots	\vdots	\vdots	\ddots

Wavelets and Peaks

BIOINFORMATICS

ORIGINAL PAPER

Vol. 22 no. 17 2006, pages 2059–2065

doi:10.1093/bioinformatics/btl355

Genome analysis

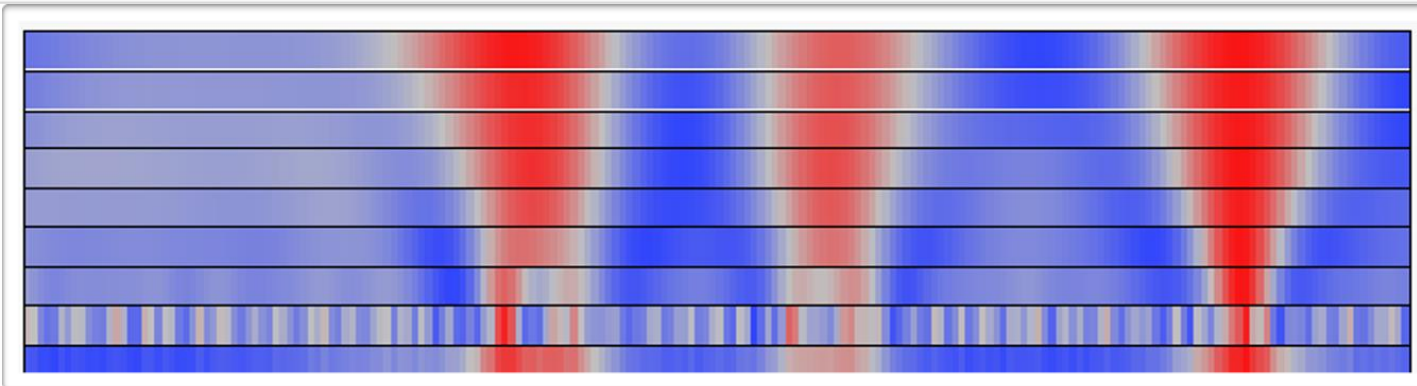
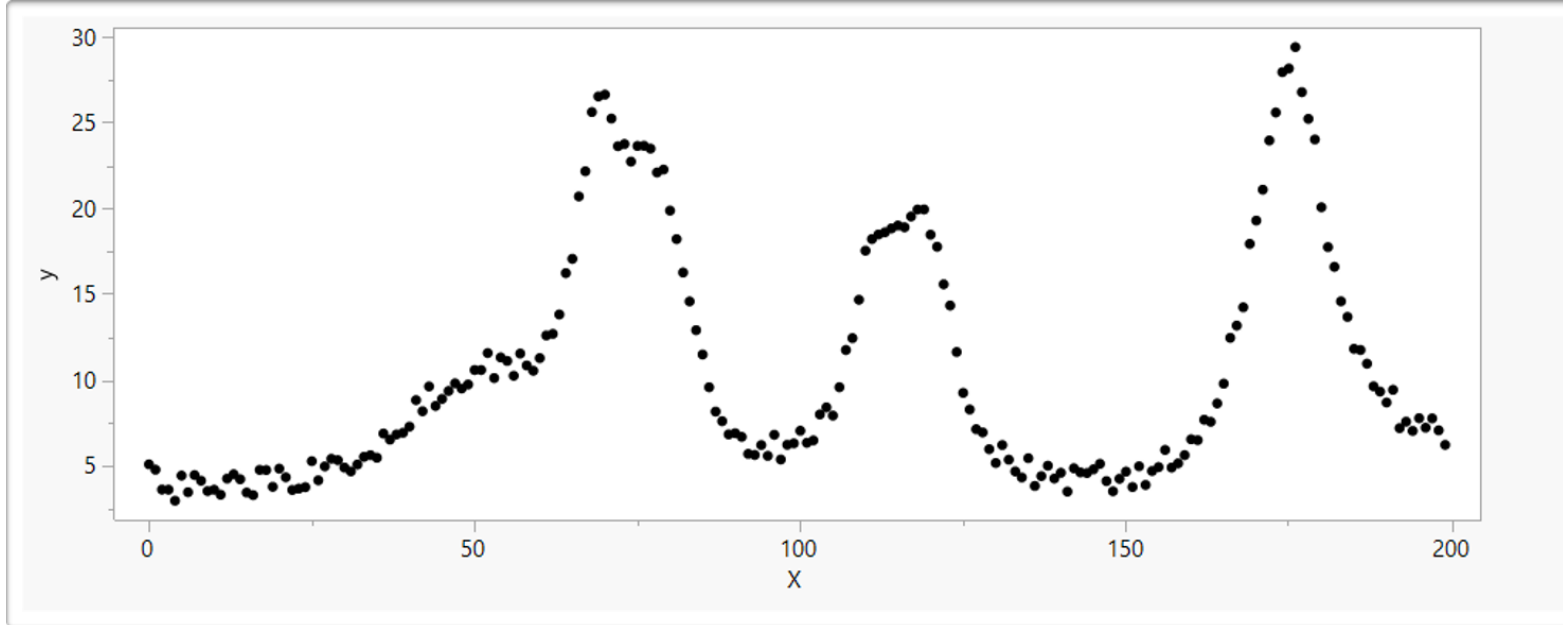
Improved peak detection in mass spectrum by incorporating continuous wavelet transform-based pattern matching

Pan Du¹, Warren A. Kibbe¹ and Simon M. Lin^{1,*}

¹Robert H. Lurie Comprehensive Cancer Center, Northwestern University, Chicago, IL 60611, USA

- Du et al (2006) proposed using CWT coefficients to help locate peaks.
- Main idea: when there is a peak at location ℓ , the wavelet coefficients near ℓ will be large for a range of scales.
- Eg $x_{\ell 1}$, $x_{\ell 2}$, $x_{\ell 3}$, ... will all be relatively large

Wavelets and Peaks



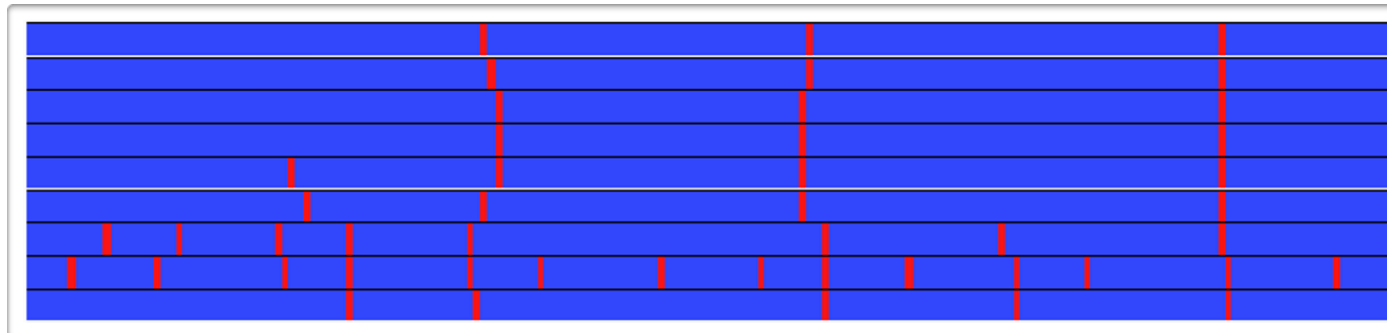
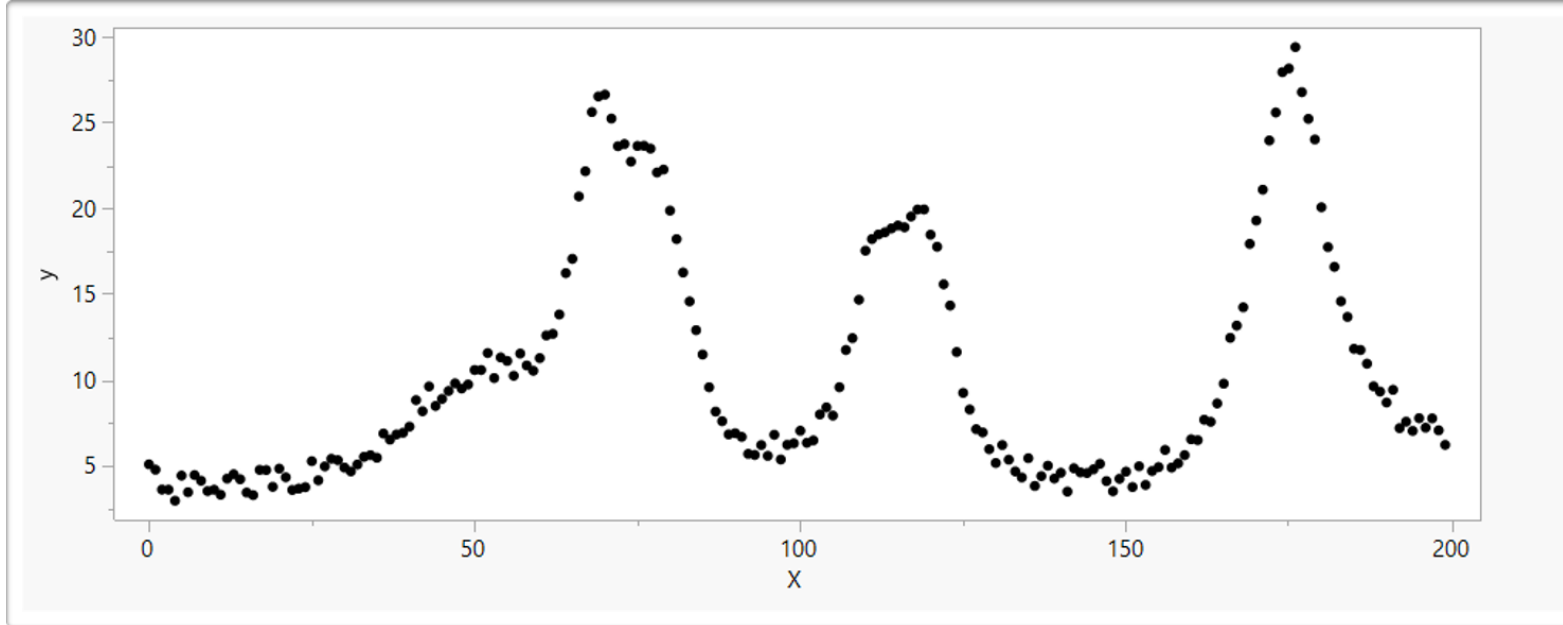
Large scale

Wavelet coefficients



Small scale

Wavelets and Peaks



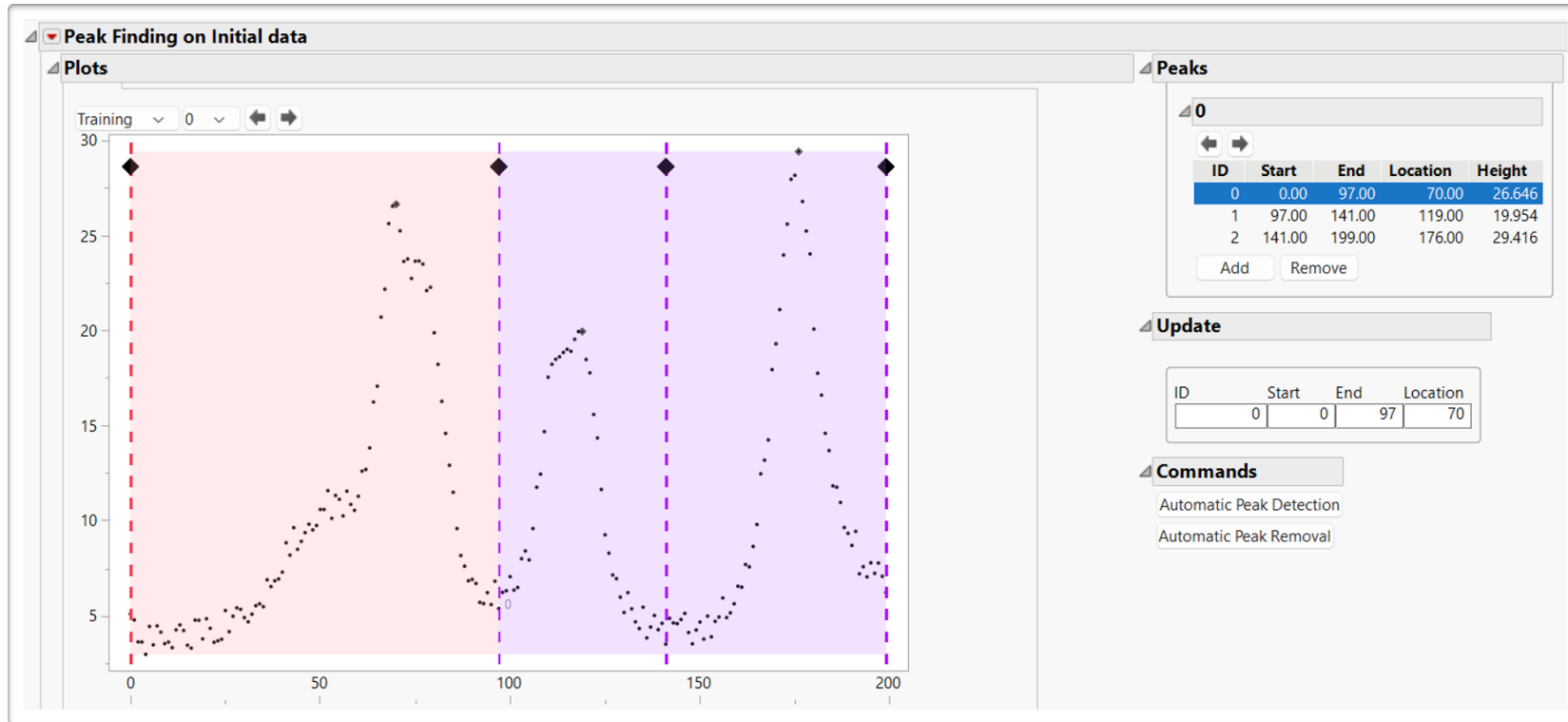
Large scale

Local maxima



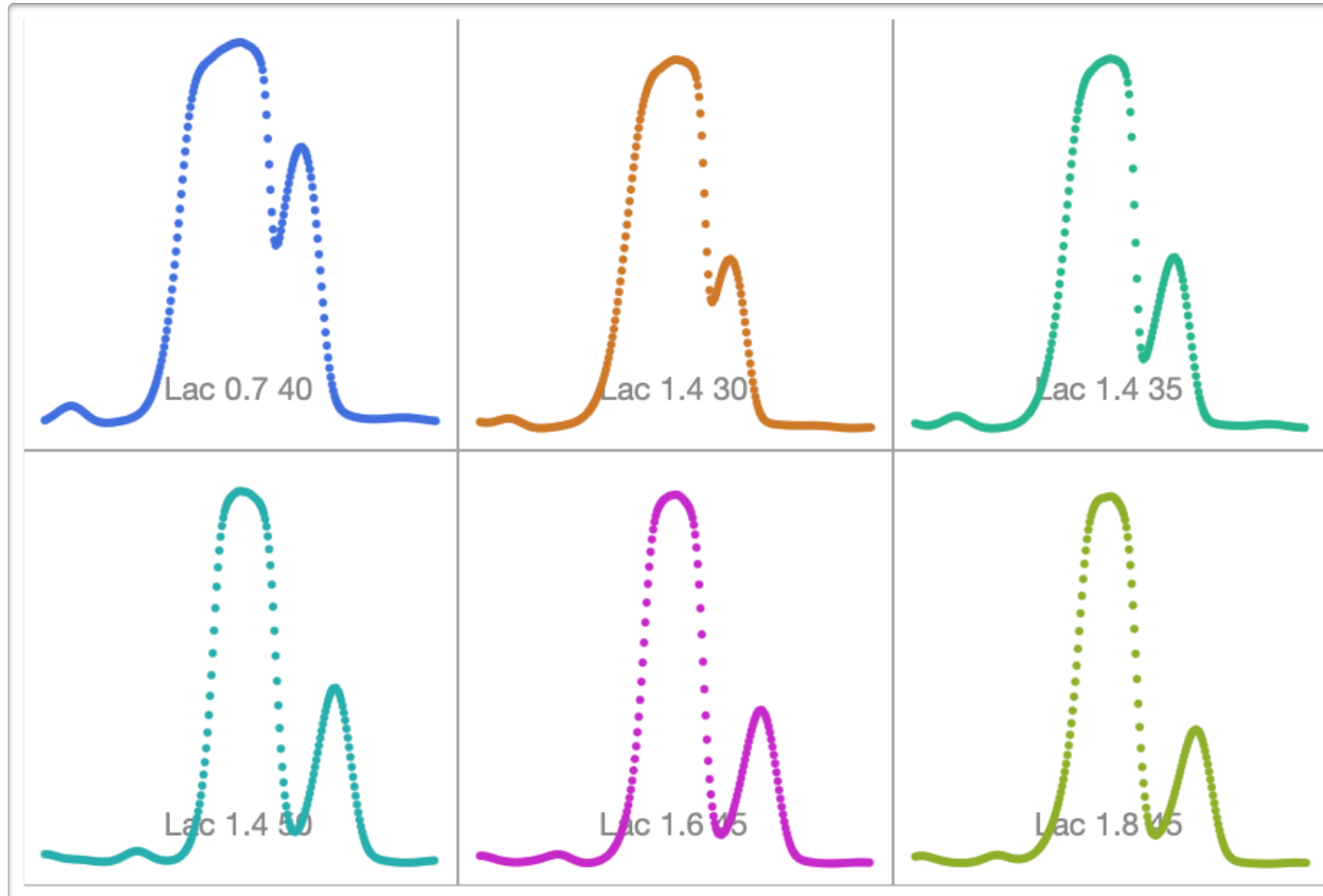
Small scale

Wavelets and Peaks



- Pan Du, Warren A. Kibbe, Simon M. Lin, Improved peak detection in mass spectrum by incorporating continuous wavelet transform-based pattern matching, Bioinformatics, Volume 22, Issue 17, September 2006, Pages 2059–2065, <https://doi.org/10.1093/bioinformatics/btl355>

Peak Finding



Chromatography (HPLC)

Chromatography DOE

Batch ID	Flow rate (mL/min)	Temperature (C)	0.0003333333	0.0070003581
Lac 0.7 30	0.7	30	-404	-402.4998926
Lac 0.7 35	0.7	35	-187	-205.0005909
Lac 0.7 40	0.7	40	-13	-13.25005371
Lac 1.4 30	1.4	30	-906	-885.5987108
Lac 1.4 35	1.4	35	171	197.25225602
Lac 1.4 40	1.4	40	98	117.42698745
Lac 1.4 45	1.4	45	1021	1012.7482274
Lac 1.4 50	1.4	50	54	58.800429718
Lac 1.6 45	1.6	45	-1785	-1778.699033
Lac 1.8 45	1.8	45	566	520.32032681

DOE Factors: Flow rate and Temperature

Demo