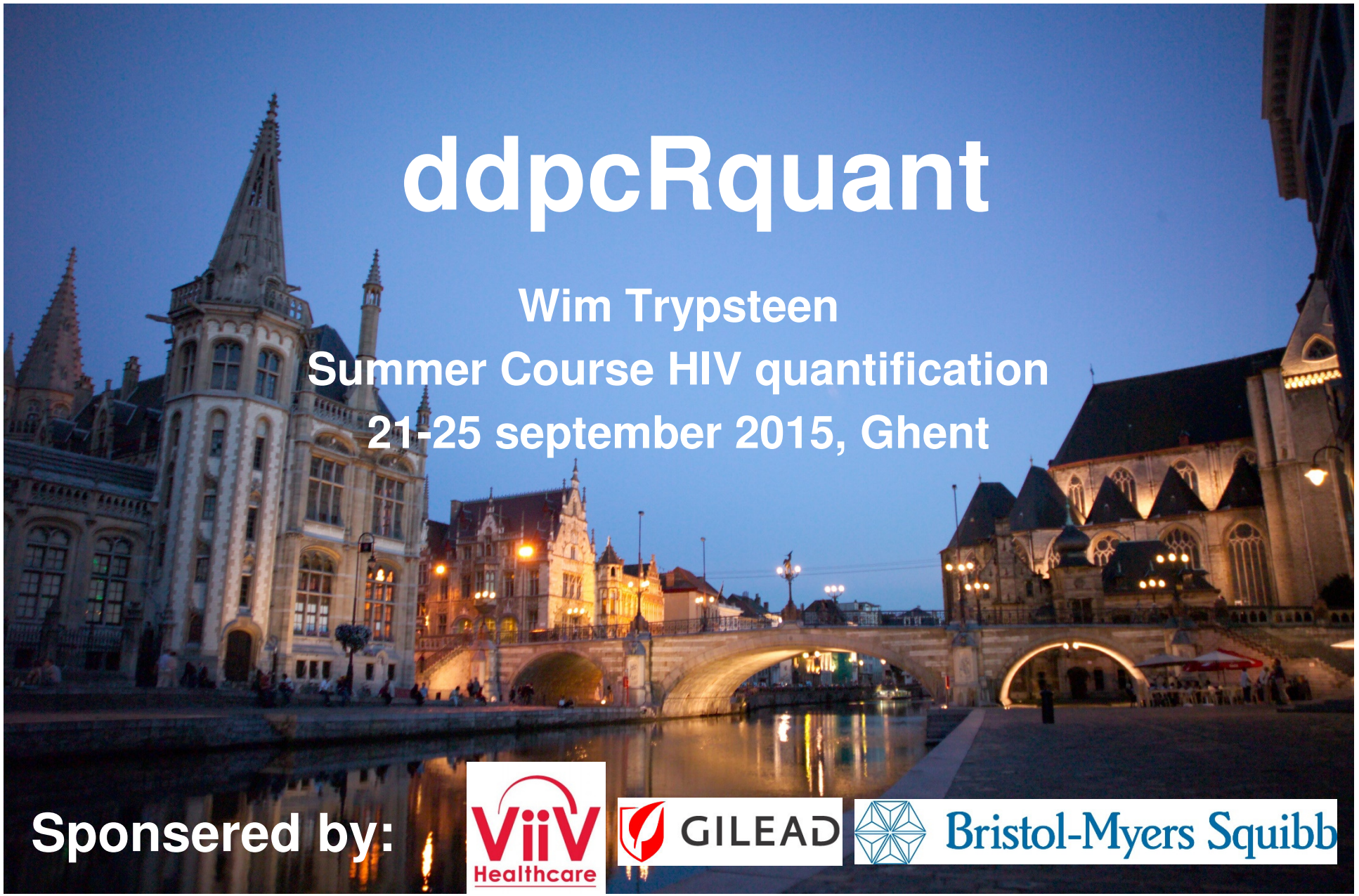
Universitair Ziekenhuis Gent

UNIVERSITEIT GENT



ddpcRquant

Wim Trypsteen

Summer Course HIV quantification

21-25 september 2015, Ghent

Sponsered by: ViiV Healthcare, GILEAD, Bristol-Myers Squibb

Normality & Extreme Value Theory

ddpcRquant overview

- Input & Automation
- NTC Pre-Processing
- NTC Processing
- Sample Processing
- Output

Take Home Messages

ddpcRquant + supplemental material (Trypsteen et al. 2015, Analytical and Bioanalytical Chemistry)

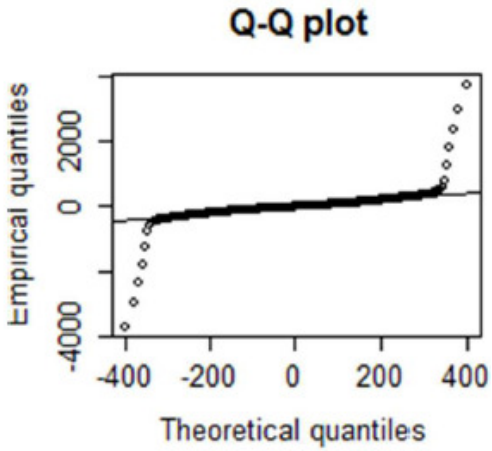
Normality?

135 NTCs: Anderson Darling test

The hypotheses for the Anderson-Darling test are:

H0: The data follows a specified distribution
H1: The data does not follow a specified distribution

P < 0.05: reject H0



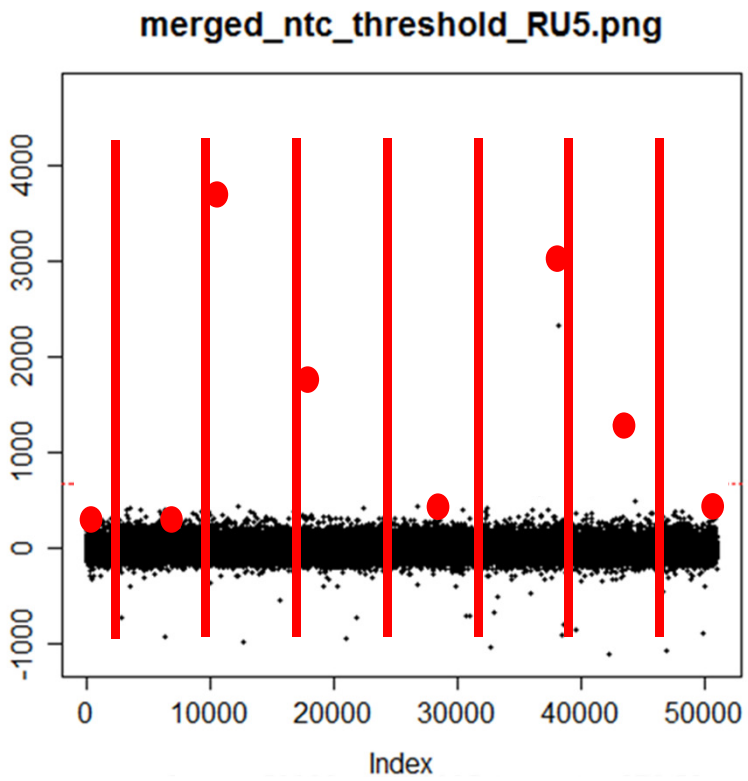
128	Not likely to follow normal distribution (p < 0.000001)
2	Not likely to follow normal distribution (0.05 > p > 0.001)
5	Likely to follow normal distribution (p > 0.05)

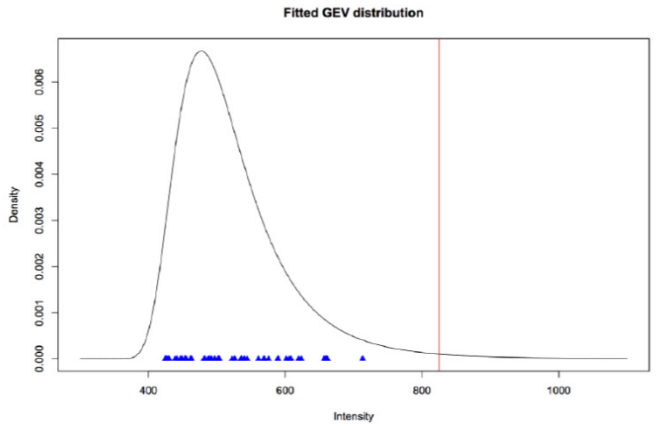
Tails => other distributions?
extreme value theory

Extreme Value Theory: concepts

Fisher-Tippet theorem: The distribution of block maxima is given by the Generalized Extreme Value distribution (GEV) and no other!

= Block maxima (extremes) follow the same distribution





Extreme Value Theory: concepts

Fisher-Tippet theorem: The distribution of block maxima is given by the Generalized Extreme Value distribution (GEV) and no other!

Block: subsample of the datapoints

Block maxima: extreme values (per block droplet with highest fluorescence)

GEV model: Fits 3 parameters (location, scale and shape) to build the distribution model

Block size: number of droplets per block

- Needs to be large enough for the Fisher-Tippet theorem to hold
- Important factor to optimize (too small, too large (bad model))

ddpcRquant

Normality & Extreme Value Theory

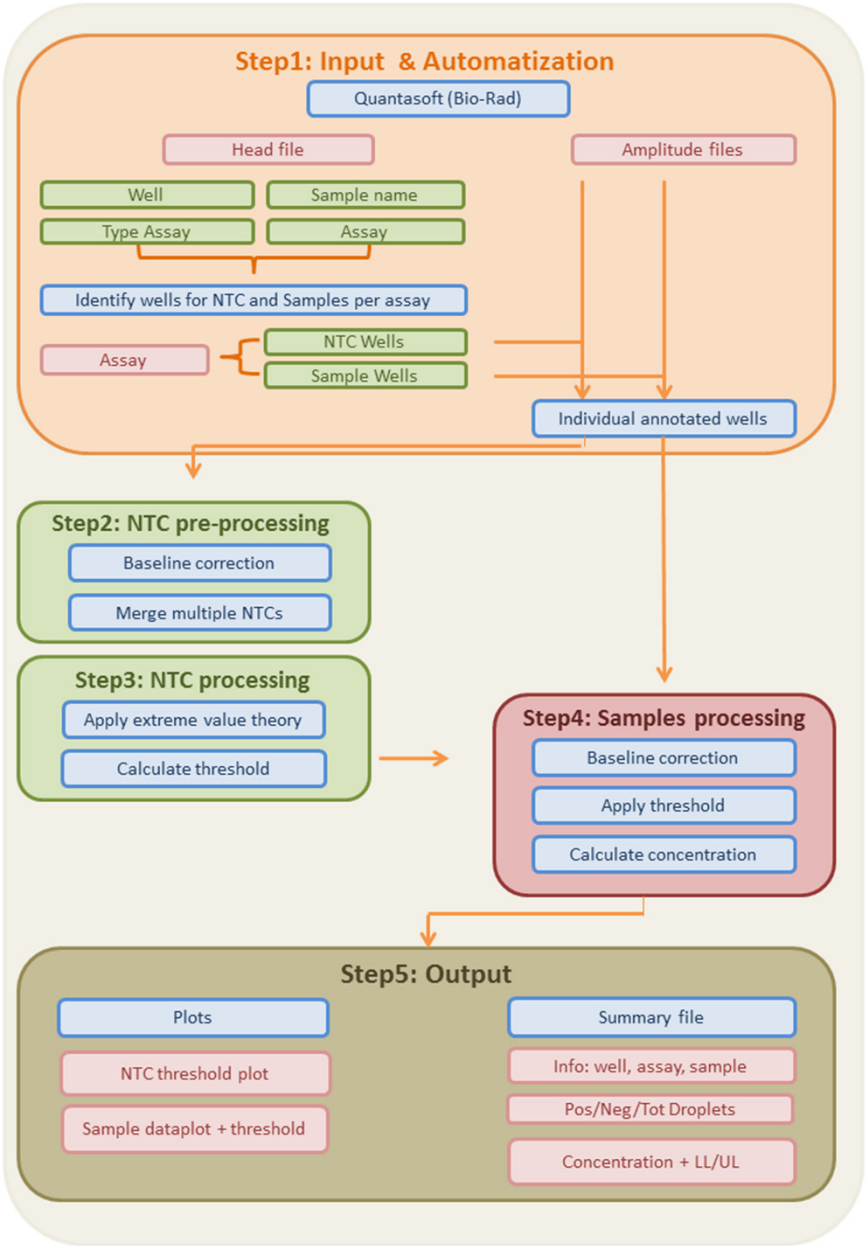
ddpcRquant overview

- Input & Automation
- NTC Pre-Processing
- NTC Processing (Treshold calculation)
- Sample Processing
- Output

Take Home Messages

ddpcRquant + supplemental material (Trypsteen et al. 2015, Analytical and Bioanalytical Chemistry)

ddpcRquant



1-D data = 1 channel at a time

Set single threshold based on NTC negative droplet population

ddpcRquant

Normality & Extreme Value Theory

ddpcRquant overview

Input & Automation

NTC Pre-Processing

NTC Processing

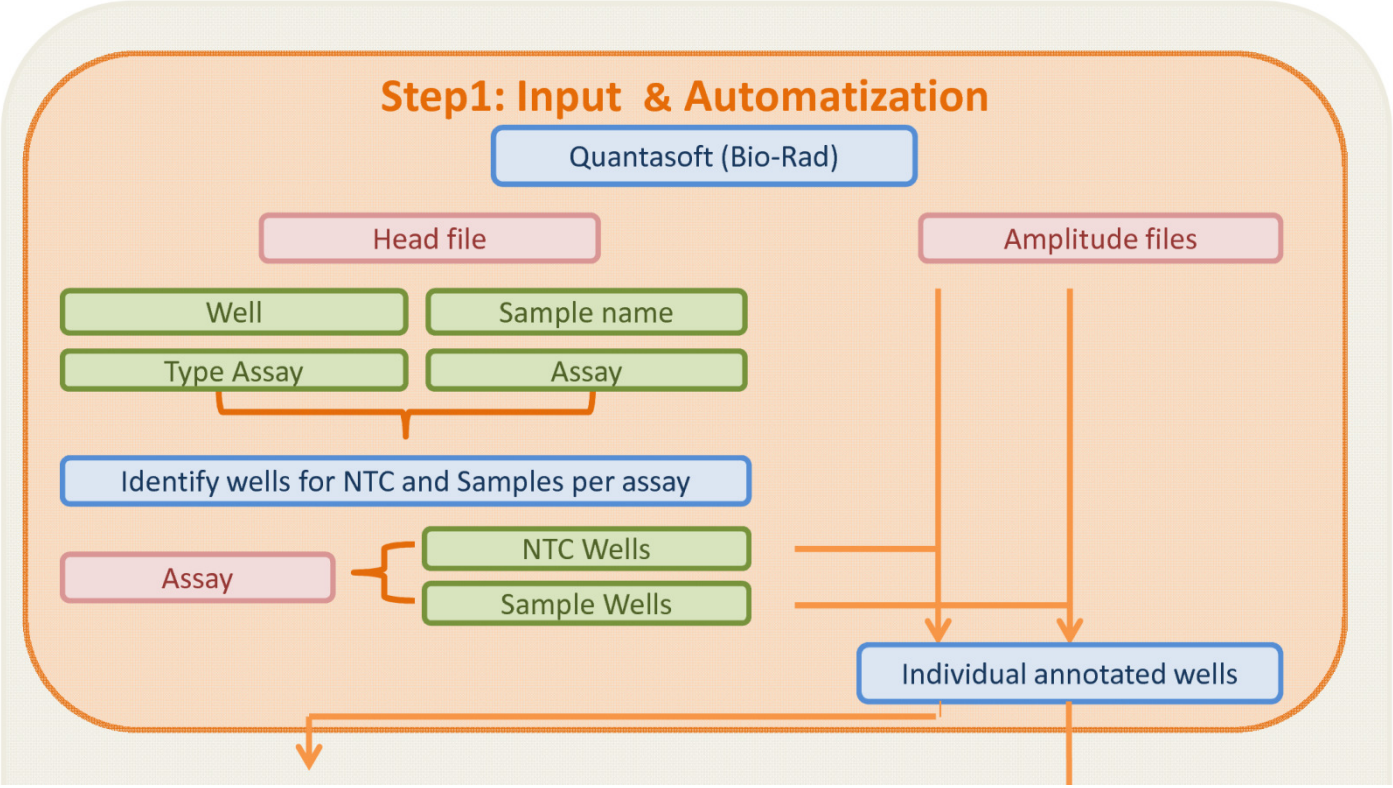
Sample Processing

Output

Take Home Messages

ddpcRquant + supplemental material (Trypsteen et al. 2015, Analytical and Bioanalytical Chemistry)

ddpcRquant > Input & Automation



Goal: Annotate individual wells with sample & assay info
Retrieve Fluoresence Data per sample

See hands-on for details

ddpcRquant

Normality & Extreme Value Theory

ddpcRquant overview

Input & Automation

NTC Pre-Processing

NTC Processing

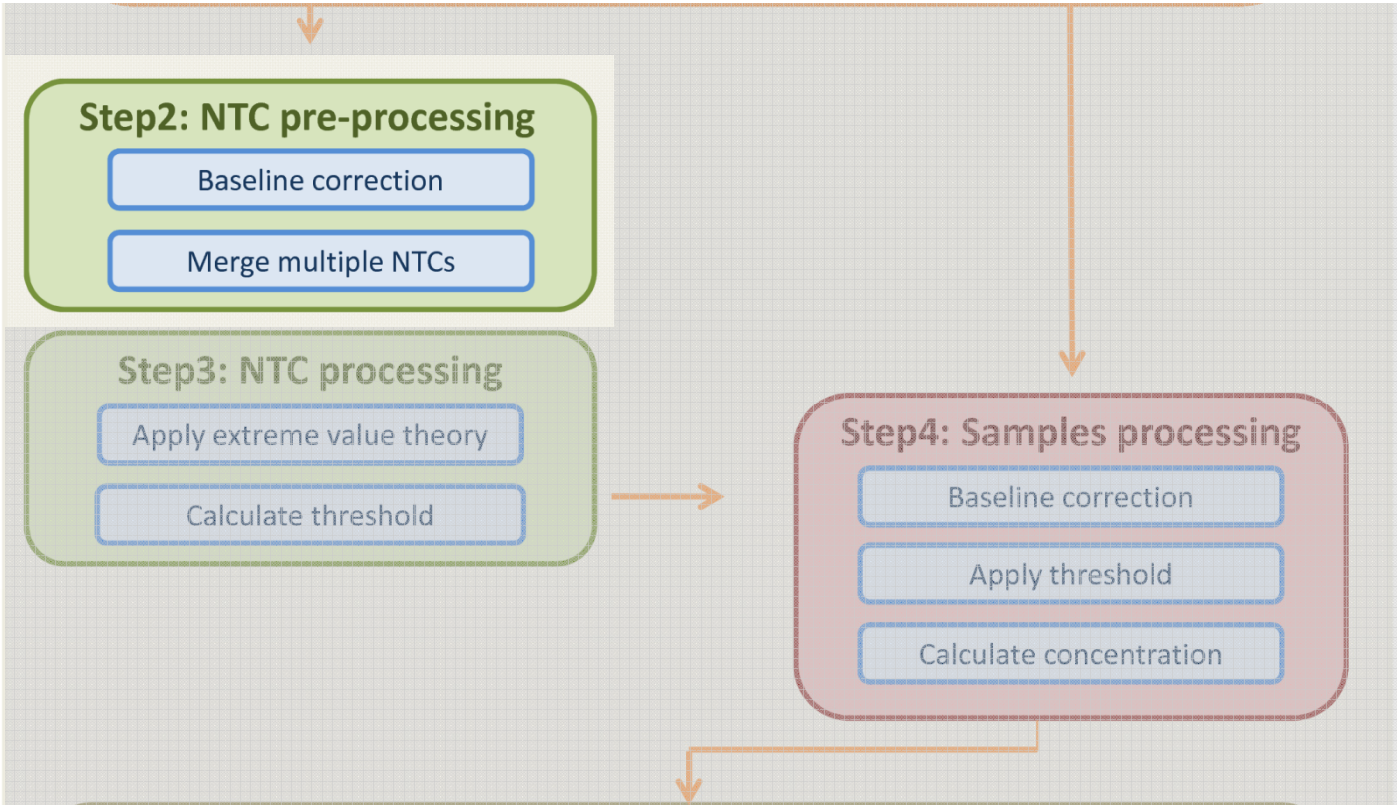
Sample Processing

Output

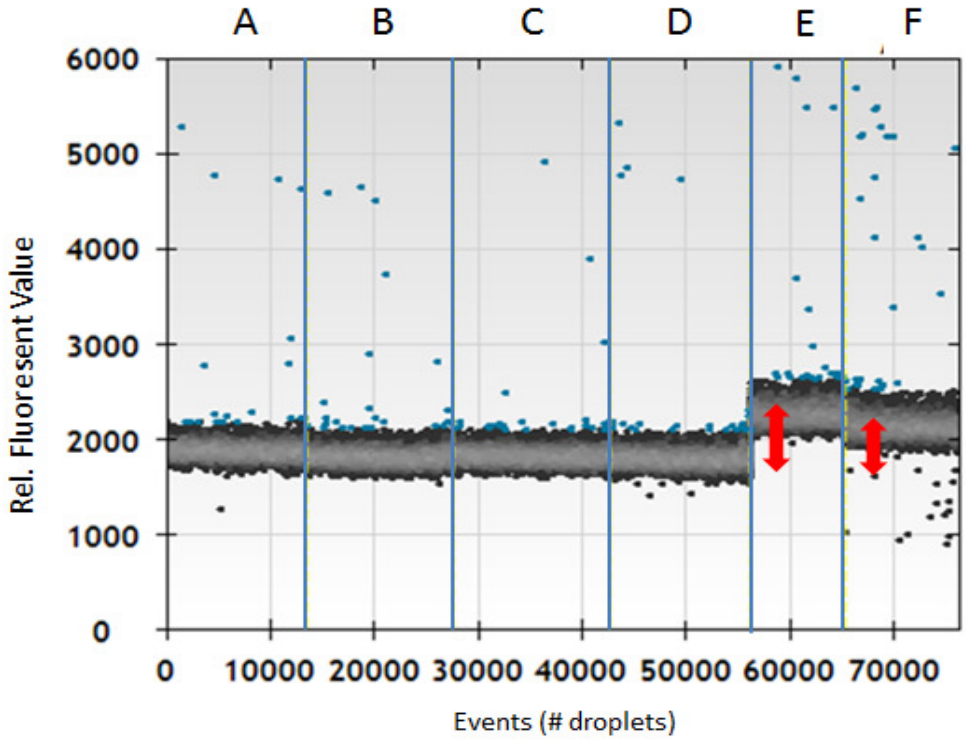
Take Home Messages

ddpcRquant + supplemental material (Trypsteen et al. 2015, Analytical and Bioanalytical Chemistry)

ddpcRquant > NTC pre-processing



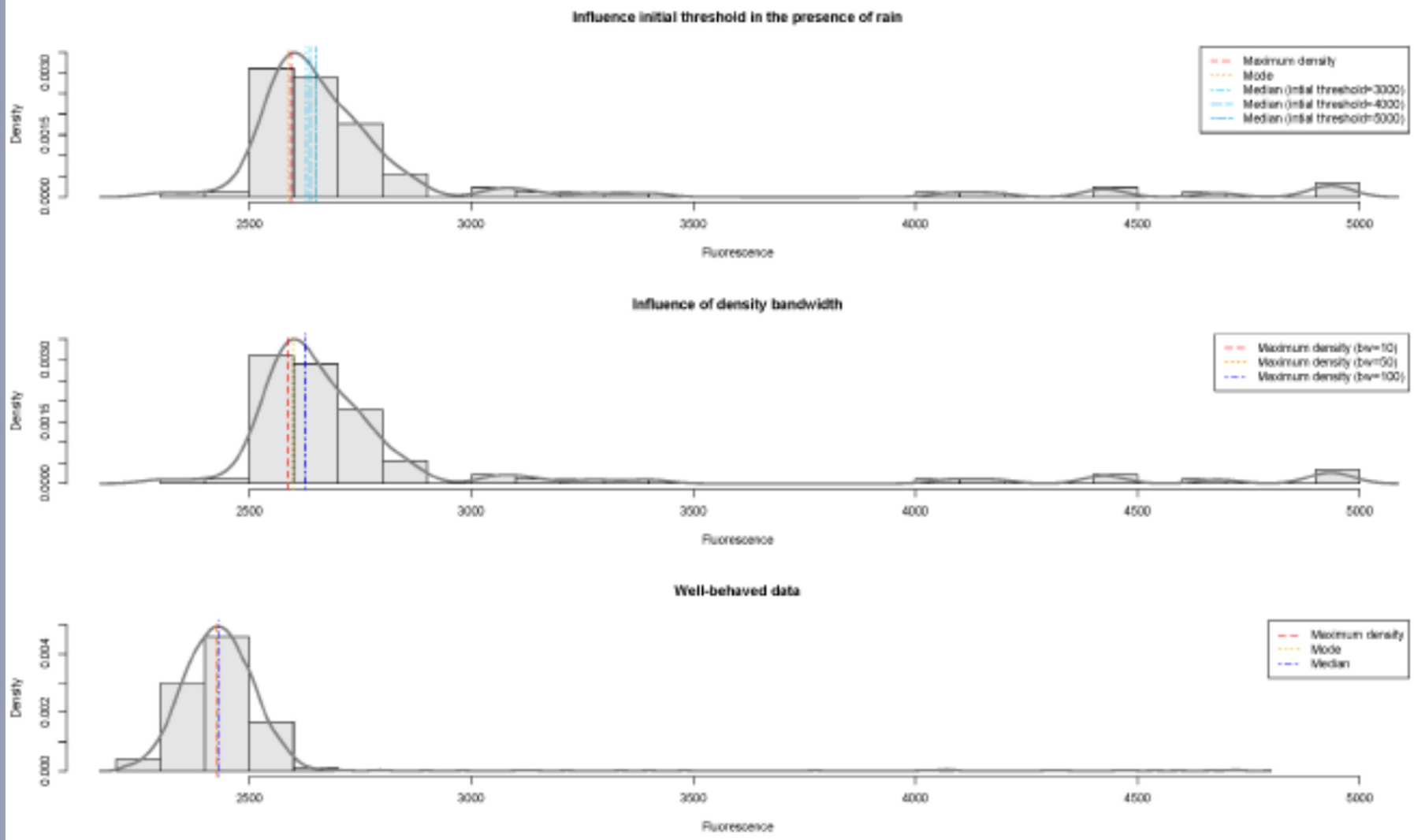
ddpcRquant > NTC pre-processing



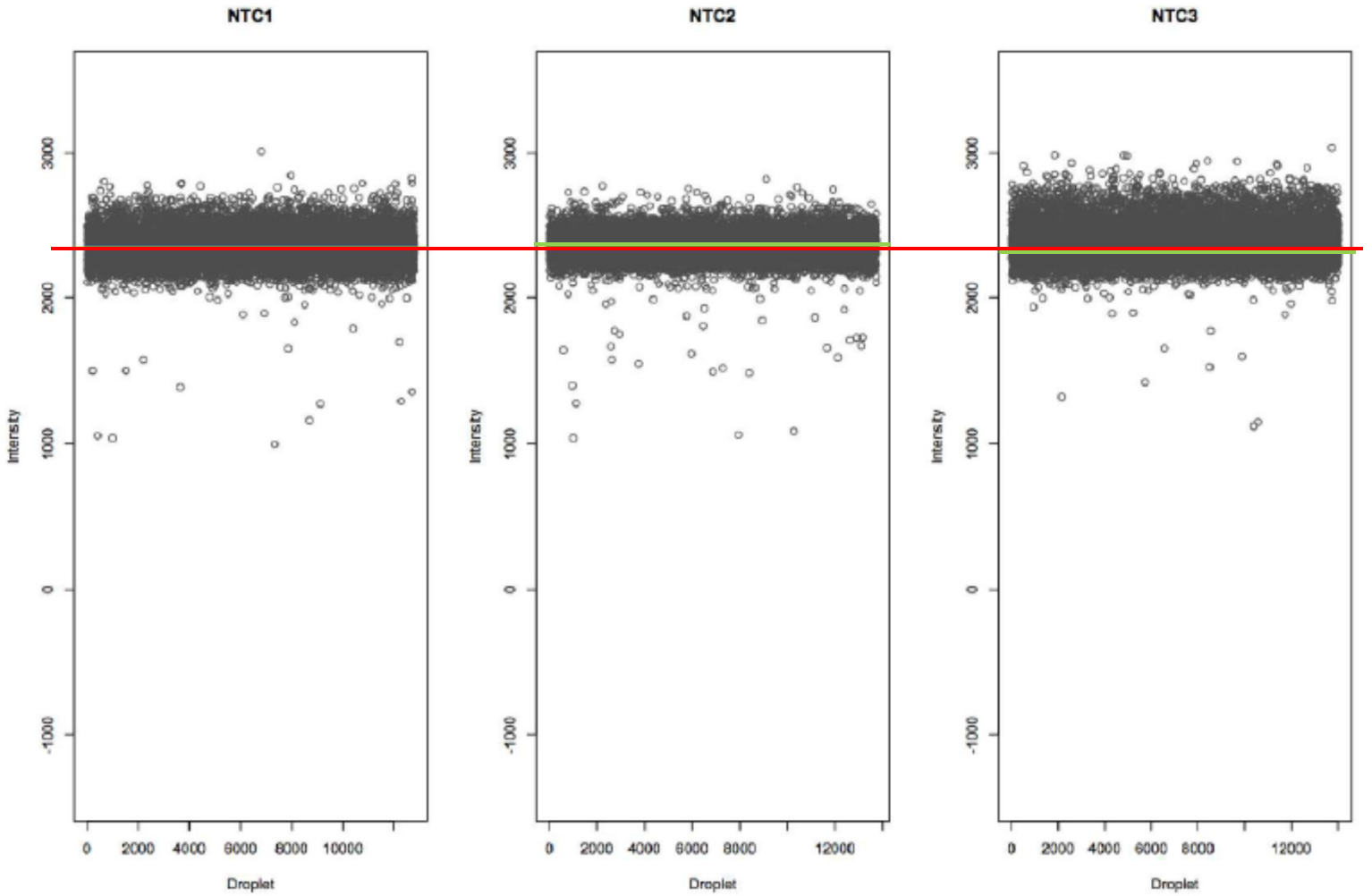
Baseline correction (shift) + merge (increase datapoints)

- Find a value to substract from all droplets
- Evaluation of median, maximum density and mode

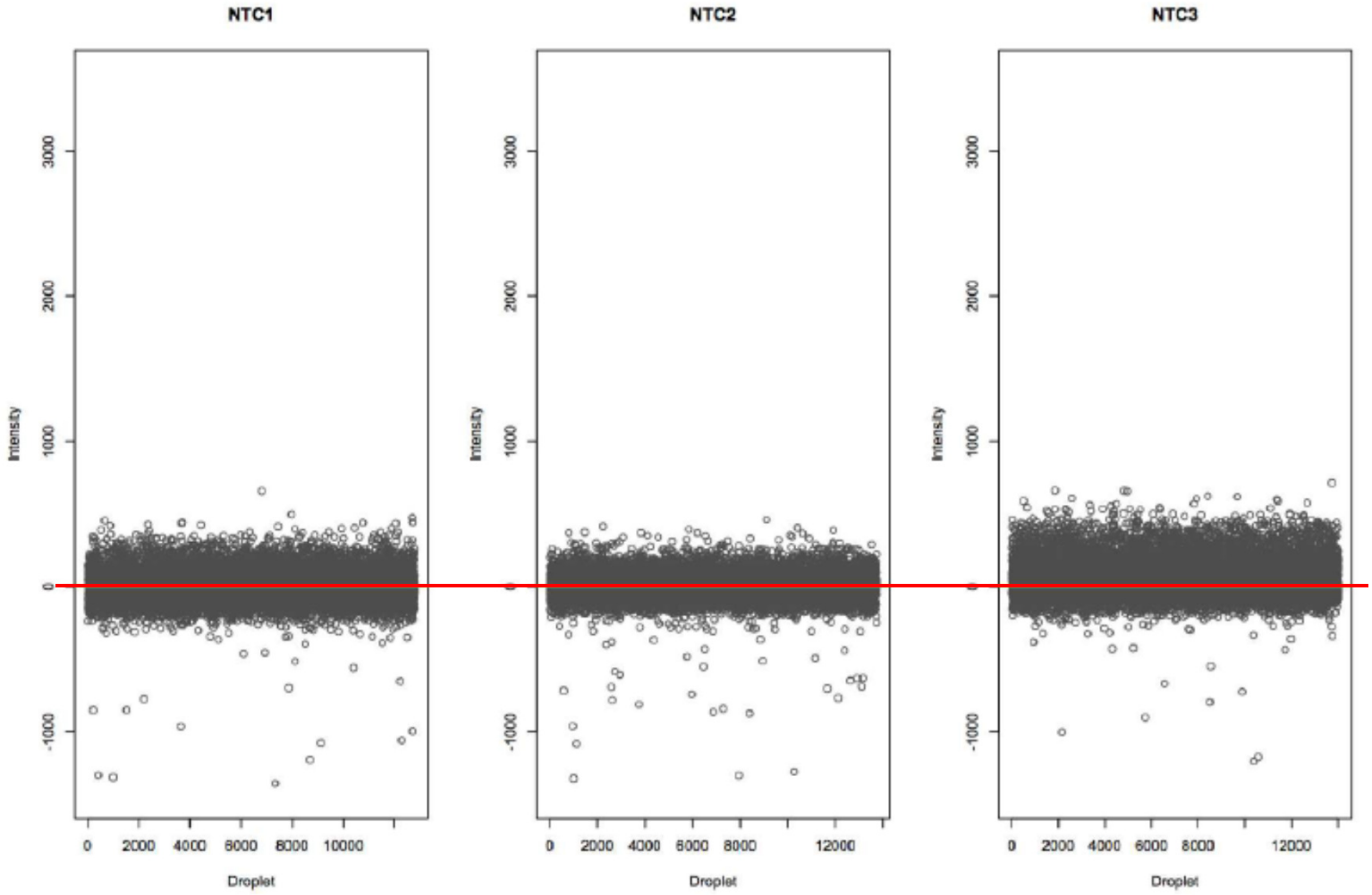
ddpcRquant > Baseline correction NTC



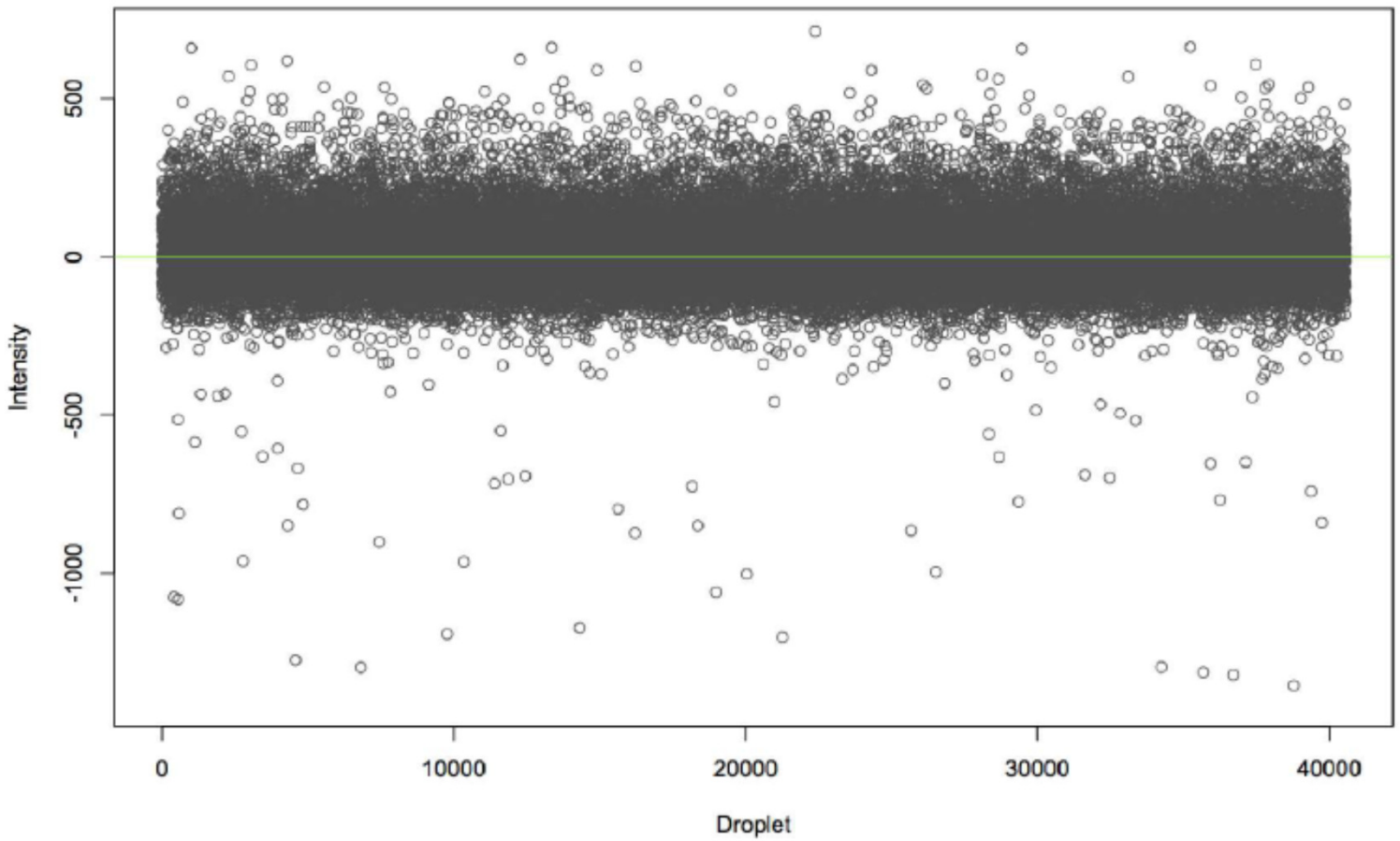
ddpcRquant > Baseline correction NTC



ddpcRquant > Baseline correction NTC



ddpcRquant > Baseline + Merge NTC



ddpcRquant

Normality & Extreme Value Theory

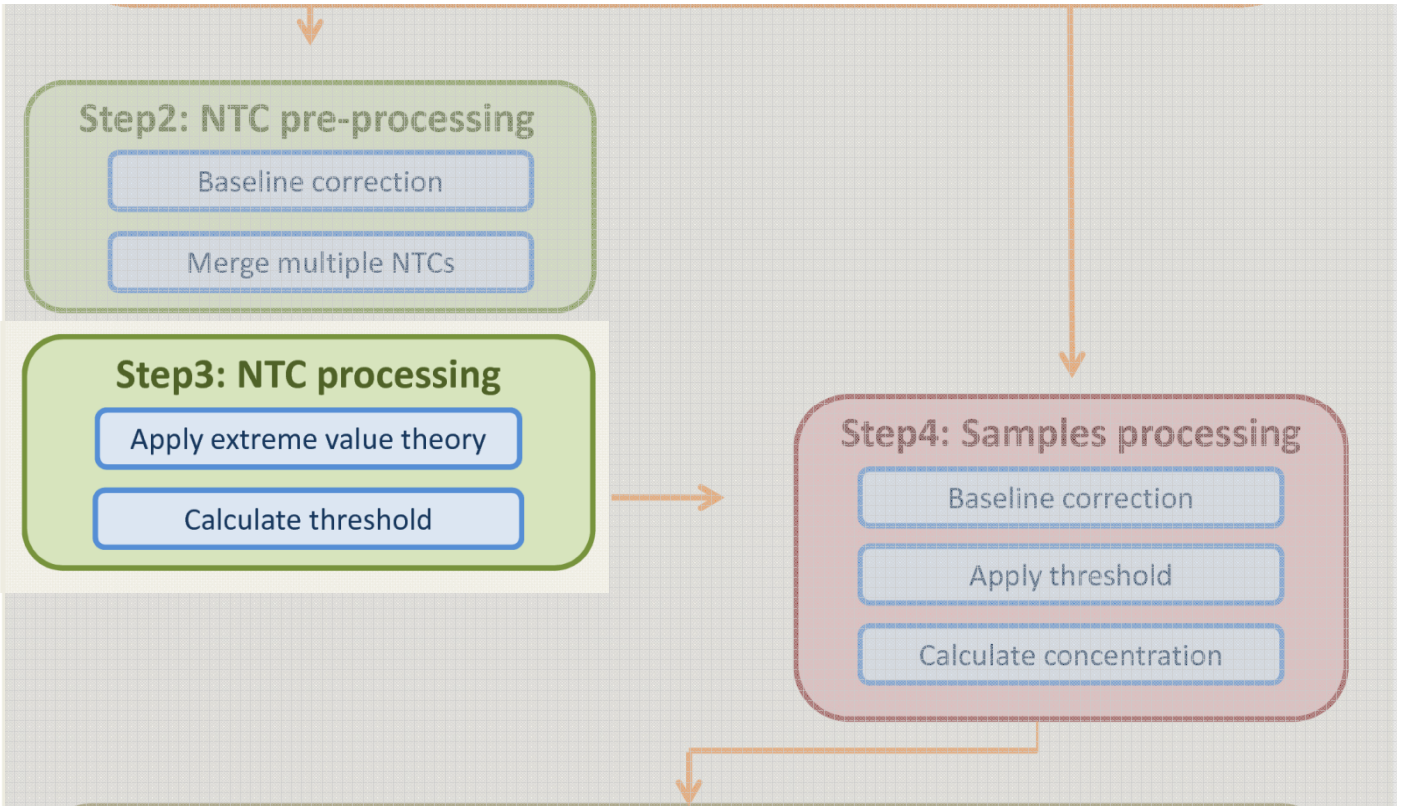
ddpcRquant overview

- Input & Automation
- NTC Pre-Processing
- **NTC Processing**
- Sample Processing
- Output

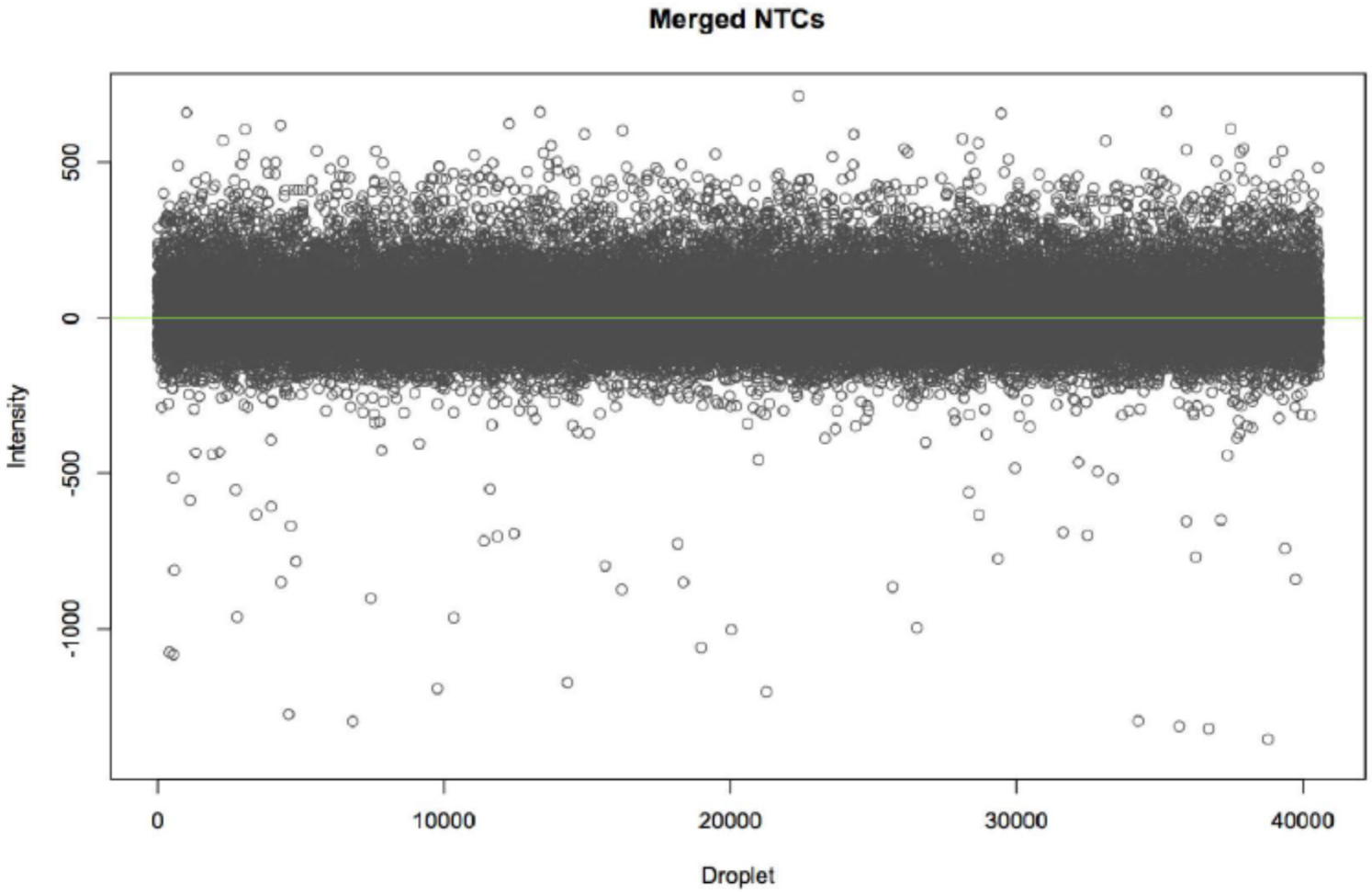
Take Home Messages

ddpcRquant + supplemental material (Trypsteen et al. 2015, Analytical and Bioanalytical Chemistry)

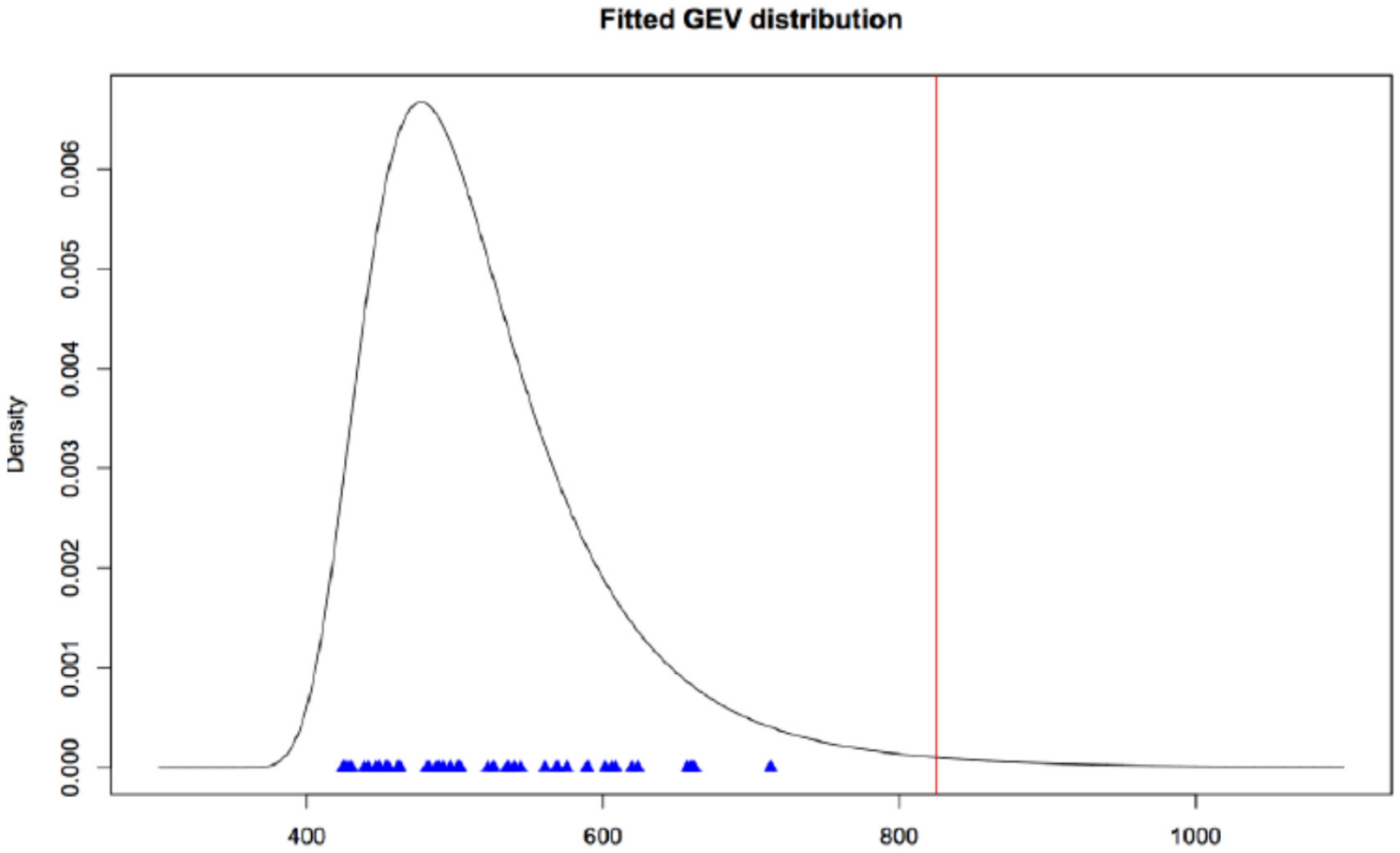
ddpcRquant > NTC processing



Apply Extreme Value Theory



Apply Extreme Value Theory



GEV(μ, σ, ξ)

$$\frac{1}{\sigma}\, t(x)^{\xi+1} e^{-t(x)},$$

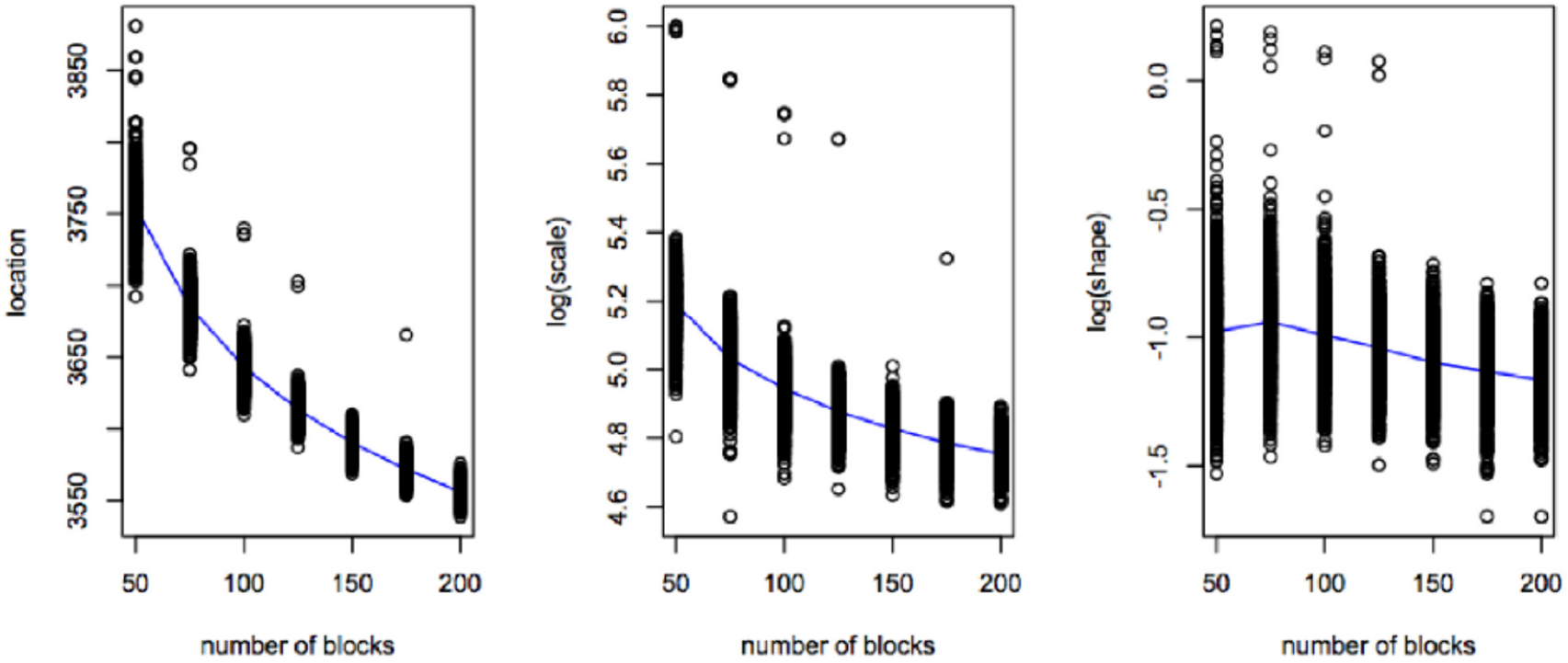
where

$$t(x) = \begin{cases} \left(1 + \left(\frac{x-\mu}{\sigma}\right)\xi\right)^{-1/\xi} & \text{if } \xi \neq 0 \\ e^{-(x-\mu)/\sigma} & \text{if } \xi = 0 \end{cases}$$

This function takes 3 parameters: location, scale and shape to be modelled based on the extremes

Apply Extreme Value Theory

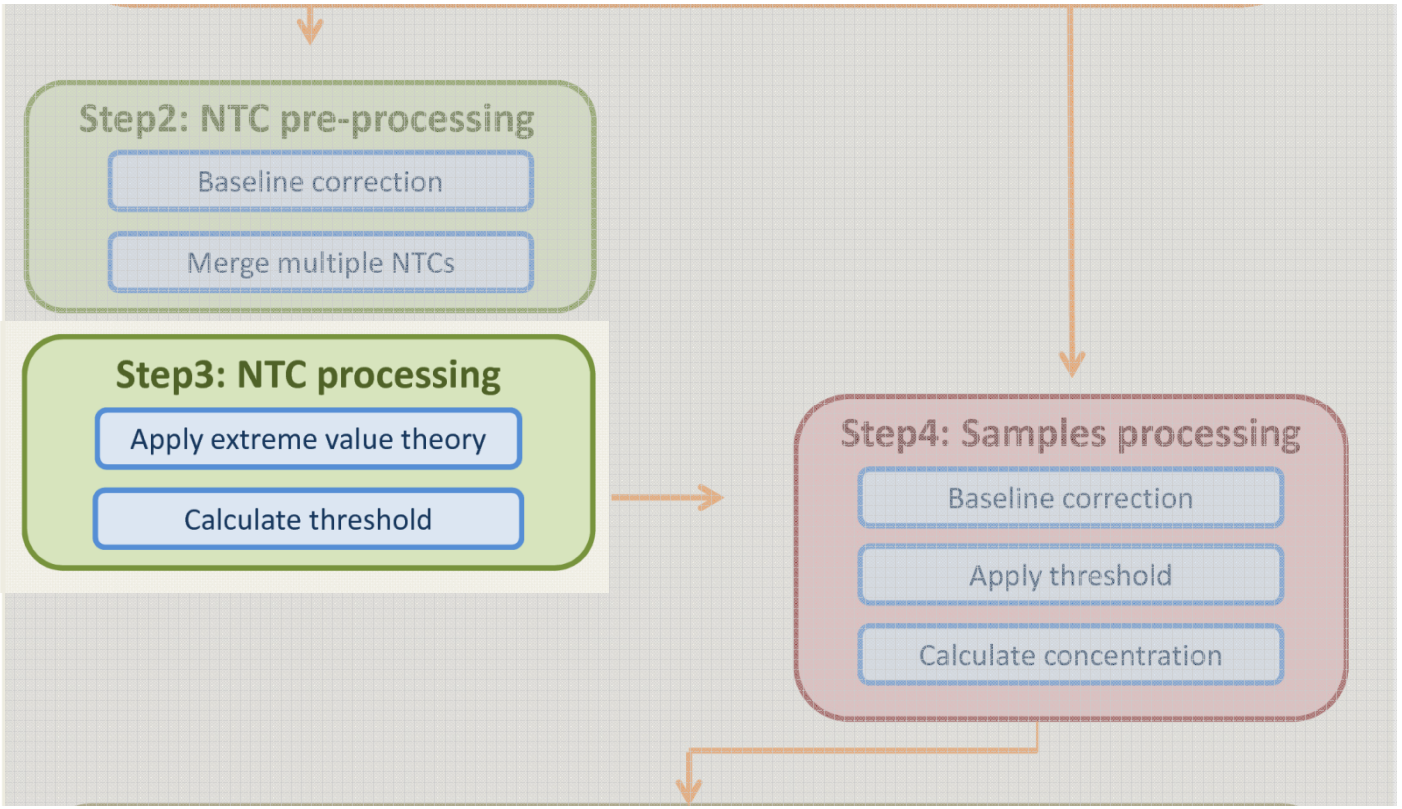
3 parameters to be determined: location / scale / shape (block size!)



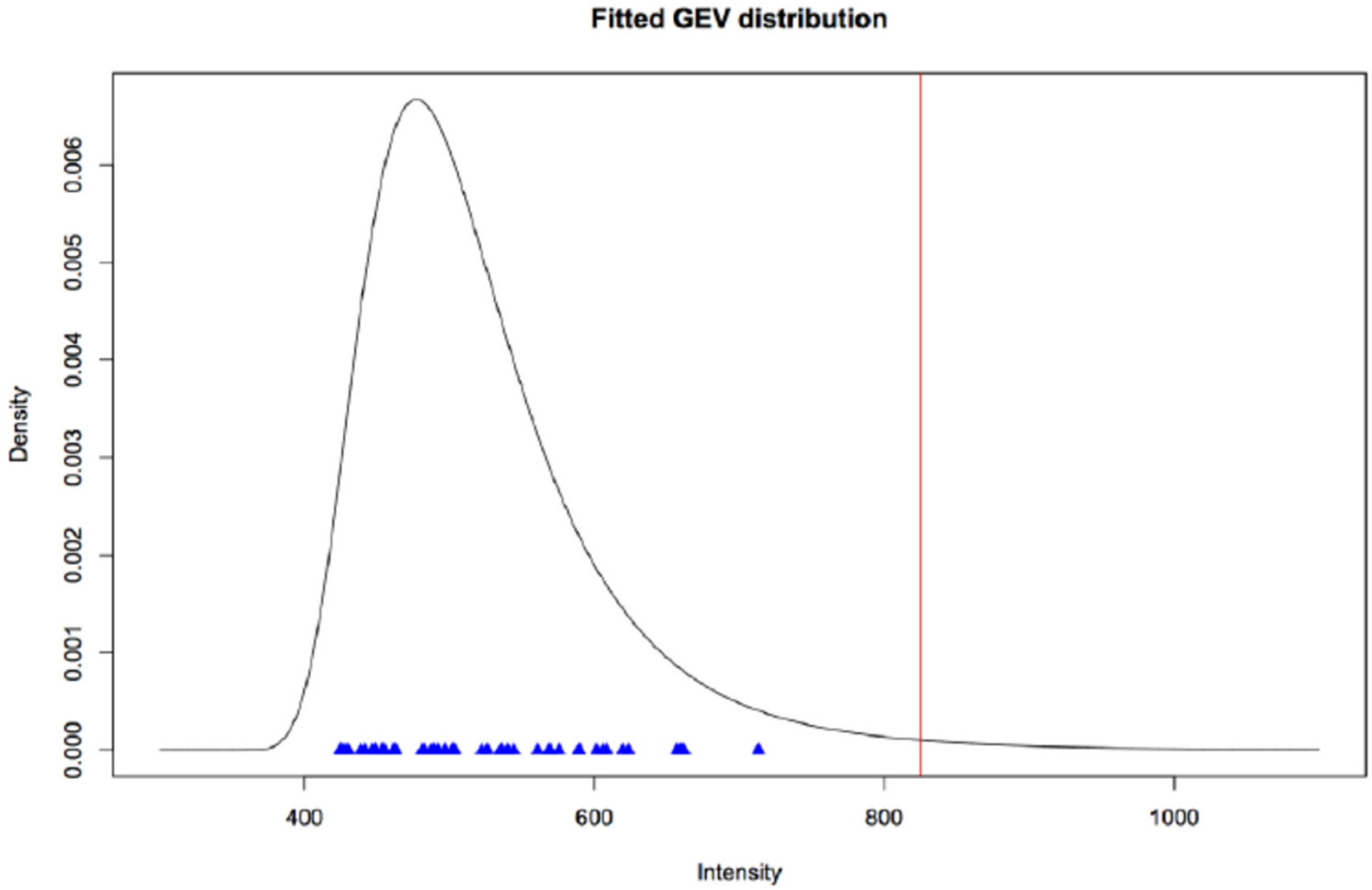
Block number too small variation bigger

Block number too big not anymore looking at extreme values
fisher tippet theorem does not hold

ddpcRquant > NTC processing



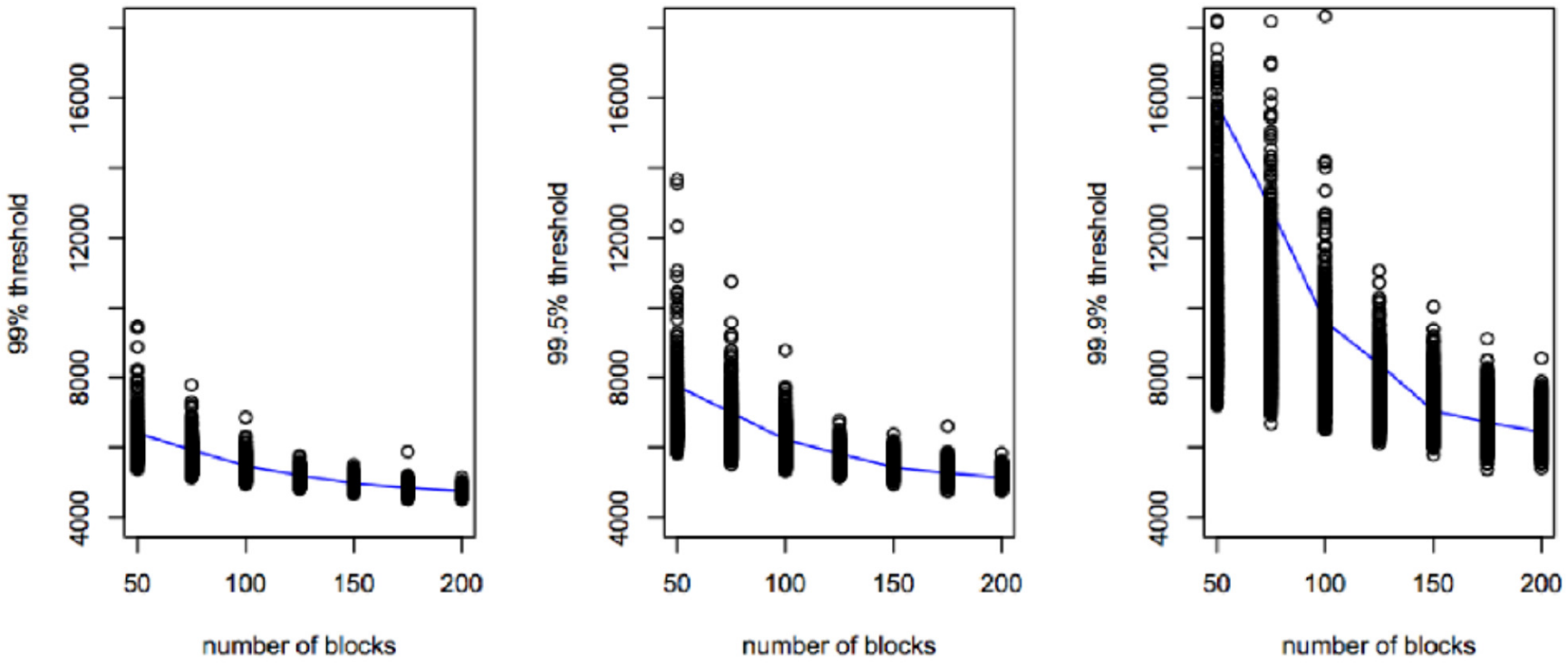
Threshold selection



Treshold **percentile of the modelled GEV distribution**

depends on block size + user defined percentile

Threshold selection



Increase in number of blocks lowers threshold variation

Increase in threshold value with increased percentile

Threshold selection

Treshold rules of thumb

We advice to scale up the blocks with the number of NTCs you have run to have the optimal threshold calculation (standard in ddpcRquant)
=> fisher-tippet theorem + variation increases

We evaluated calculated threshold levels and determined a default value of 99.95 % (but is user defined!)

Threshold interpretation

Based on the extremes of the NTC you can say that if you go above the threshold (0.9999) you have 1 in 10000 chance that the extreme event in the sample is a false positive

ddpcRquant

Normality & Extreme Value Theory

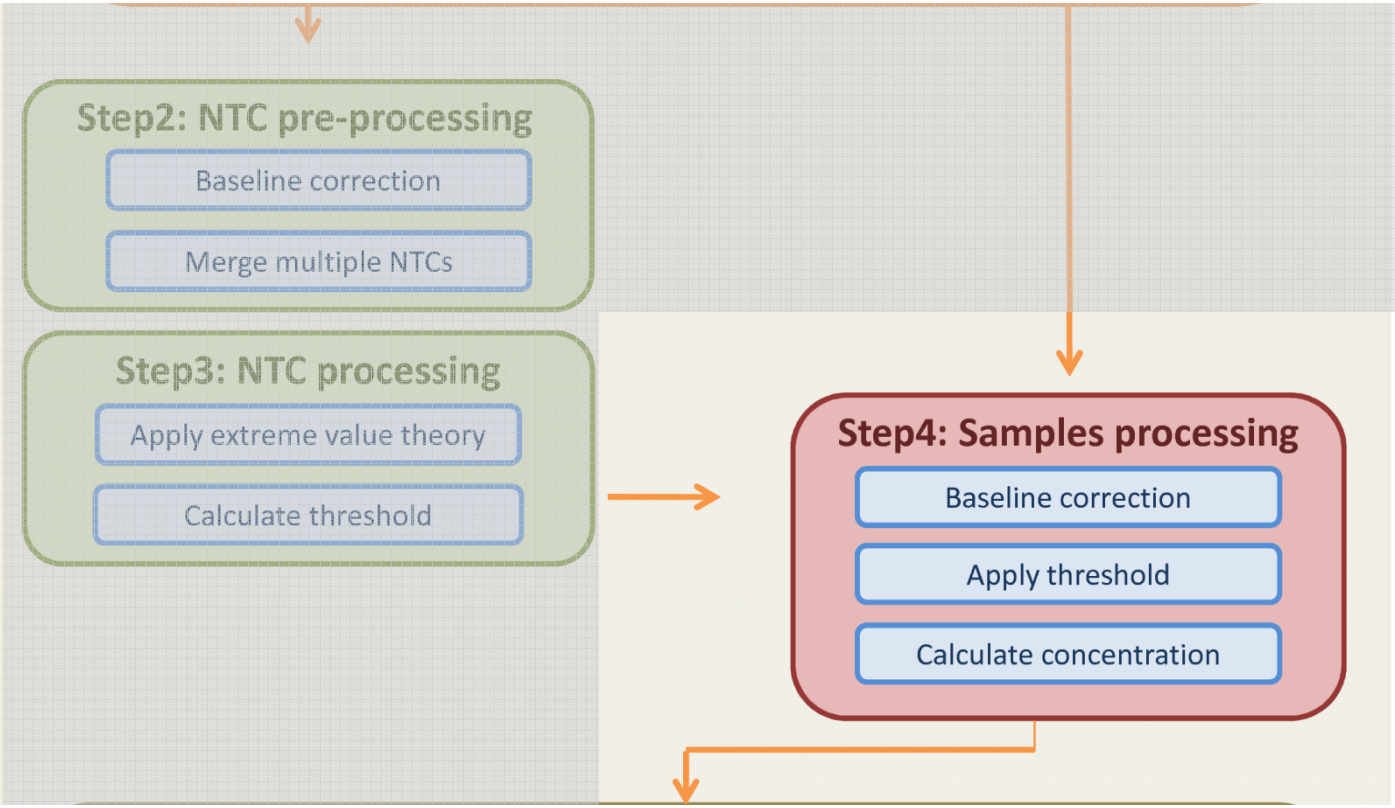
ddpcRquant overview

- Input & Automation
- NTC Pre-Processing
- NTC Processing
- **Sample Processing**
- Output

Take Home Messages

ddpcRquant + supplemental material (Trypsteen et al. 2015, Analytical and Bioanalytical Chemistry)

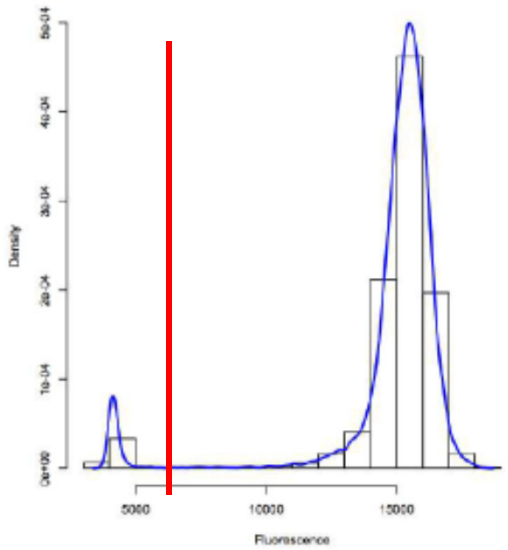
ddpcRquant > Sample Processing



Sample Processing

1. **Determine the mode (cfr. NTC)**

- Difficulty: multiple droplet populations
- We need starting value for mode estimation
- Automated: mean(NTC modes) + threshold



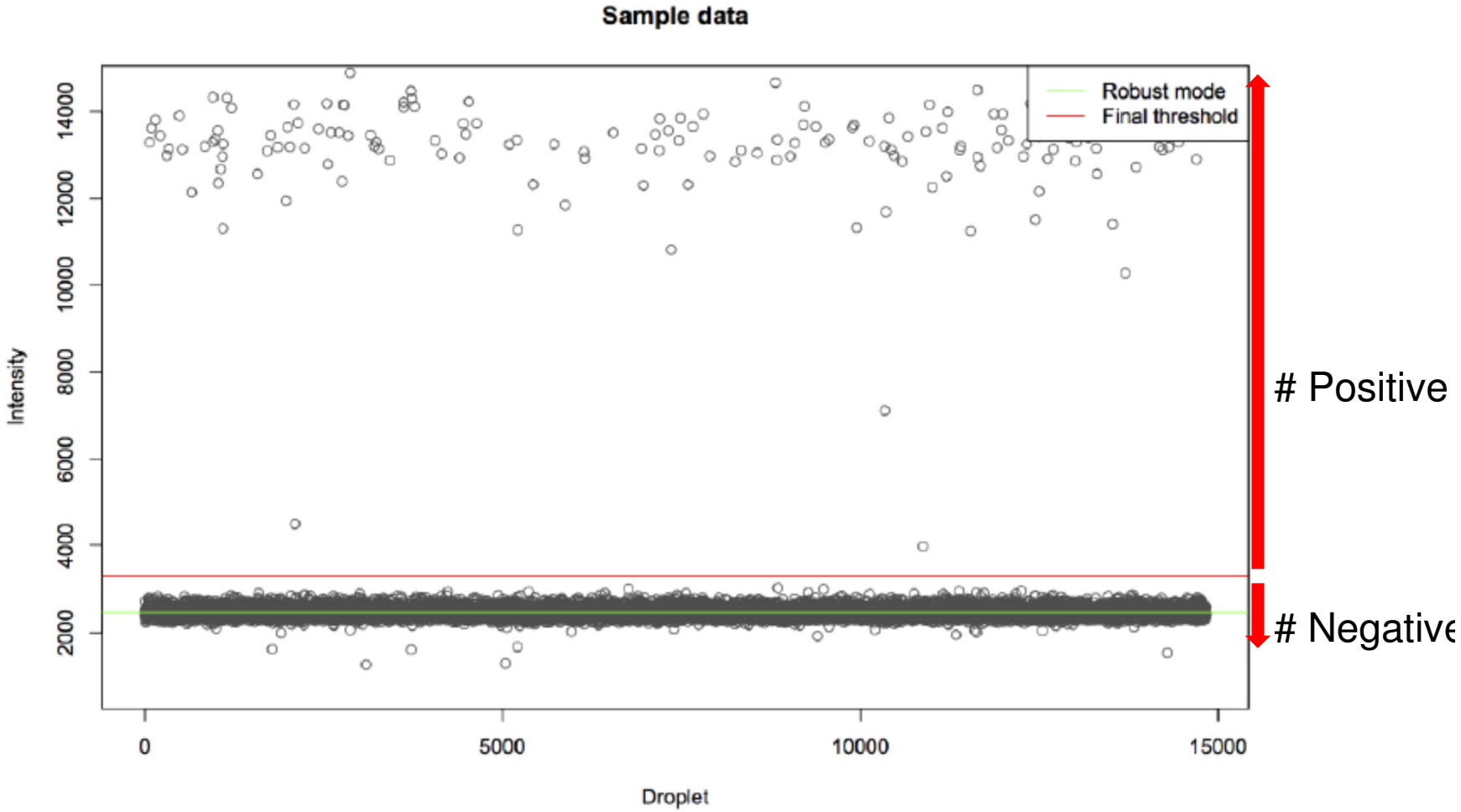
Important: Run NTCs with template free gDNA (background fluo)

2. **Apply threshold**

Sample Threshold = Calculated Sample Mode + Threshold value

3. **Quantification by Poisson distribution calcalutions**

Sample Processing



$$C = -\ln\left(\frac{Nneg}{N}\right) * \frac{1000}{Vd} * D$$

ddpcRquant

Normality & Extreme Value Theory

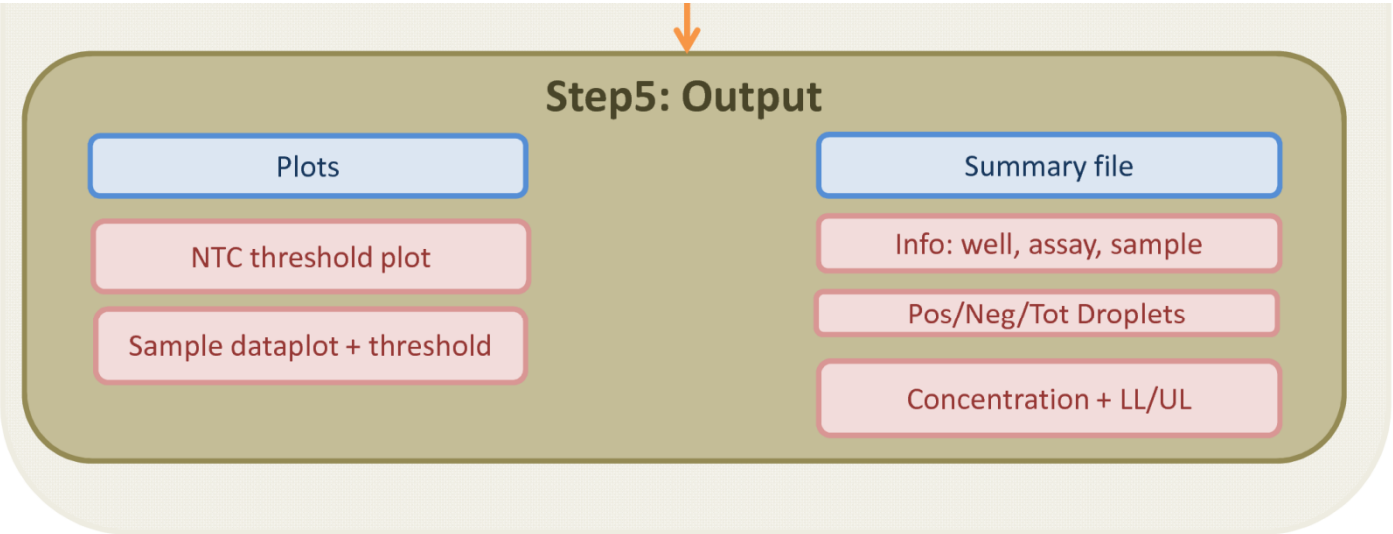
ddpcRquant overview

- Input & Automation
- NTC Pre-Processing
- NTC Processing
- Sample Processing
- **Output**

Take Home Messages

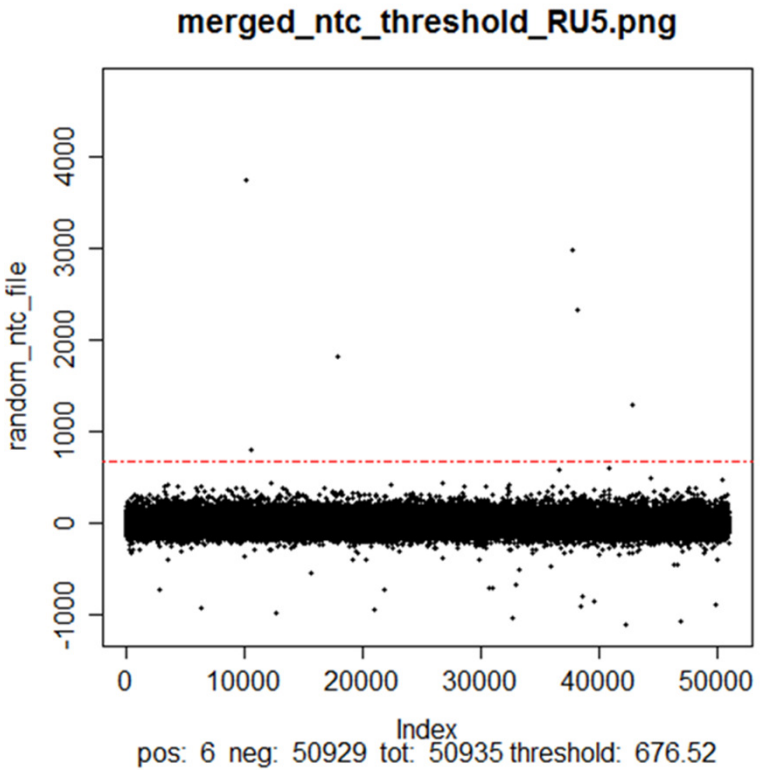
ddpcRquant + supplemental material (Trypsteen et al. 2015, Analytical and Bioanalytical Chemistry)

ddpcRquant > Output

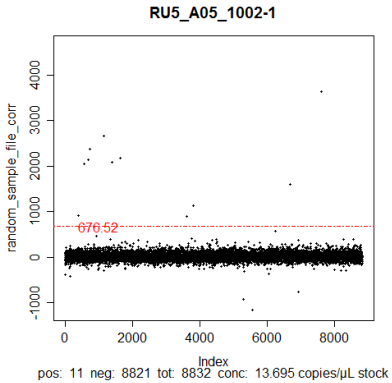


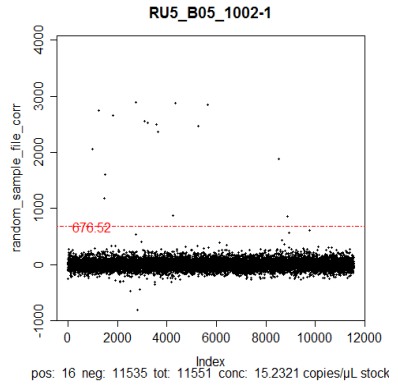
Output > Plots

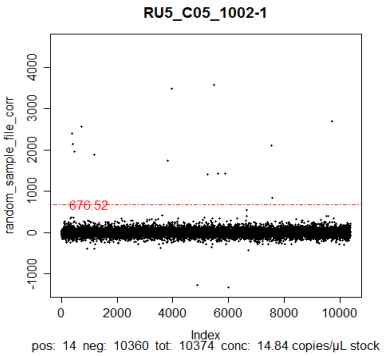
MERGED NTC + threshold

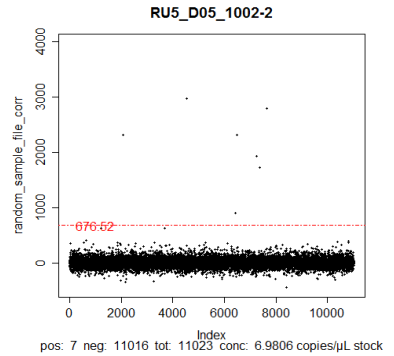


SAMPLES









Output > summary files

Csv file

	Well	assay	name	type	positive droplets	negative droplets	total droplets	concentration	lowerCI	upperCI
1	merged	2 LTR	merged_NTC	ntc	1	19376	19377	0.5671	0.0666	4.8298
2	E01	2 LTR	gDNA 1	sample	116	10957	11073	115.7274	91.1503	146.887
3	F01	2 LTR	gDNA2	sample	163	10391	10554	171.0427	139.8199	209.1717
4	G01	2 LTR	plas 1	sample	1	12425	12426	0.8844	0.1038	7.531
5	H01	2 LTR	plas 2	sample	9	7758	7767	12.7409	5.5316	29.3334

HTML file: http://www.ddpcrquant.ugent.be/output/RnaseP.html

ddpcRquant

Normality & Extreme Value Theory

ddpcRquant overview

- Input & Automation
- NTC Pre-Processing
- NTC Processing
- Sample Processing
- Output

Take Home Messages

ddpcRquant + supplemental material (Trypsteen et al. 2015, Analytical and Bioanalytical Chemistry)

Take Home Messages

Extremes follow GEV distribution and no other (Fisher-Tippet)

Block size has an effect on the parameter estimation of the GEV

Threshold = percentile of the GEV model based on the extremes (data points)