

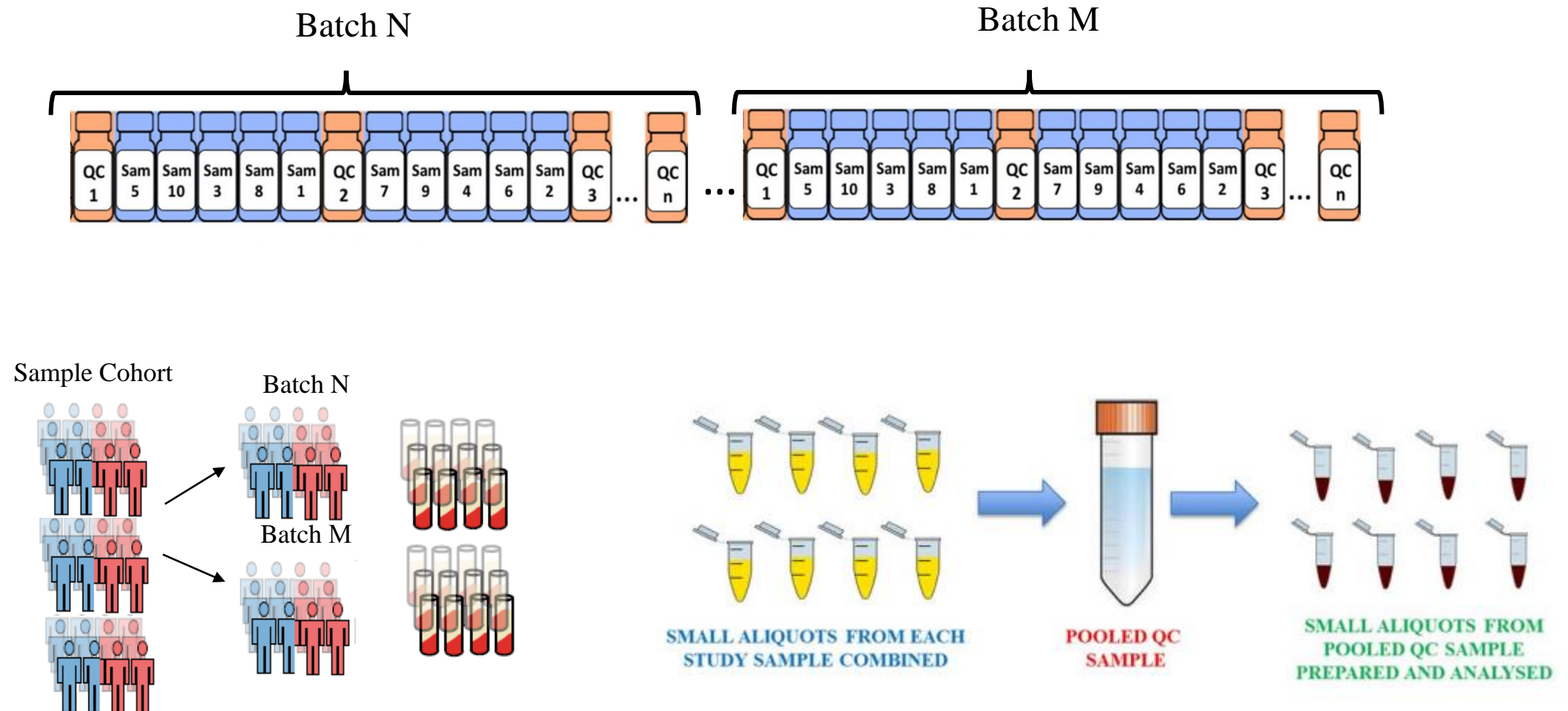
# Application of gradient boosting machine for signal processing in LC-MS metabolomics

I. V. Plyushchenko\*, I. A. Rodin

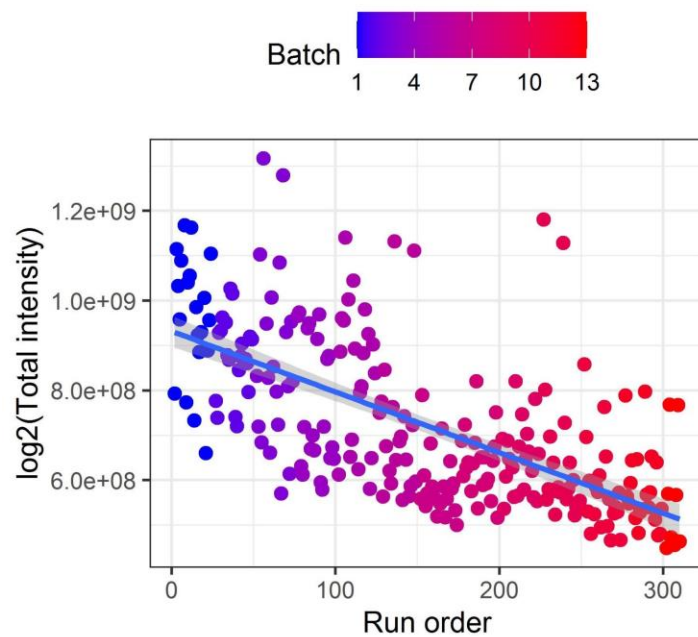
*Chemistry Department, Lomonosov Moscow State University, Moscow  
119991, Russia*

\*e-mail: [plyushchenko.ivan@gmail.com](mailto:plyushchenko.ivan@gmail.com)

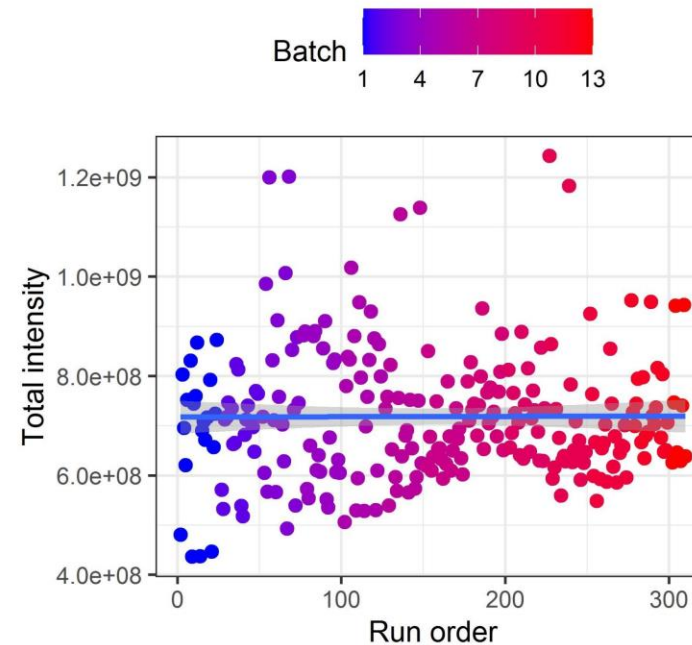
# Batch (Sequence) in LC-MS Untargeted Metabolomics



# Signal Drift



Raw data



After correction

# Signal Correction

$$M_i = f_{i,QC}(y_{i,QC} \sim x_{i,QC}) \quad (1)$$

$$F_i = M_i(x_i) \quad (2)$$

$$I'_i = \frac{I_i}{F_i} 1000 \quad (3)$$

where  $i$  is the index of metabolite,  $QC$  is the index of QC samples (all study samples are denoted without this index),

$M$  is a machine learning model, that is fitted by function  $f$ ,

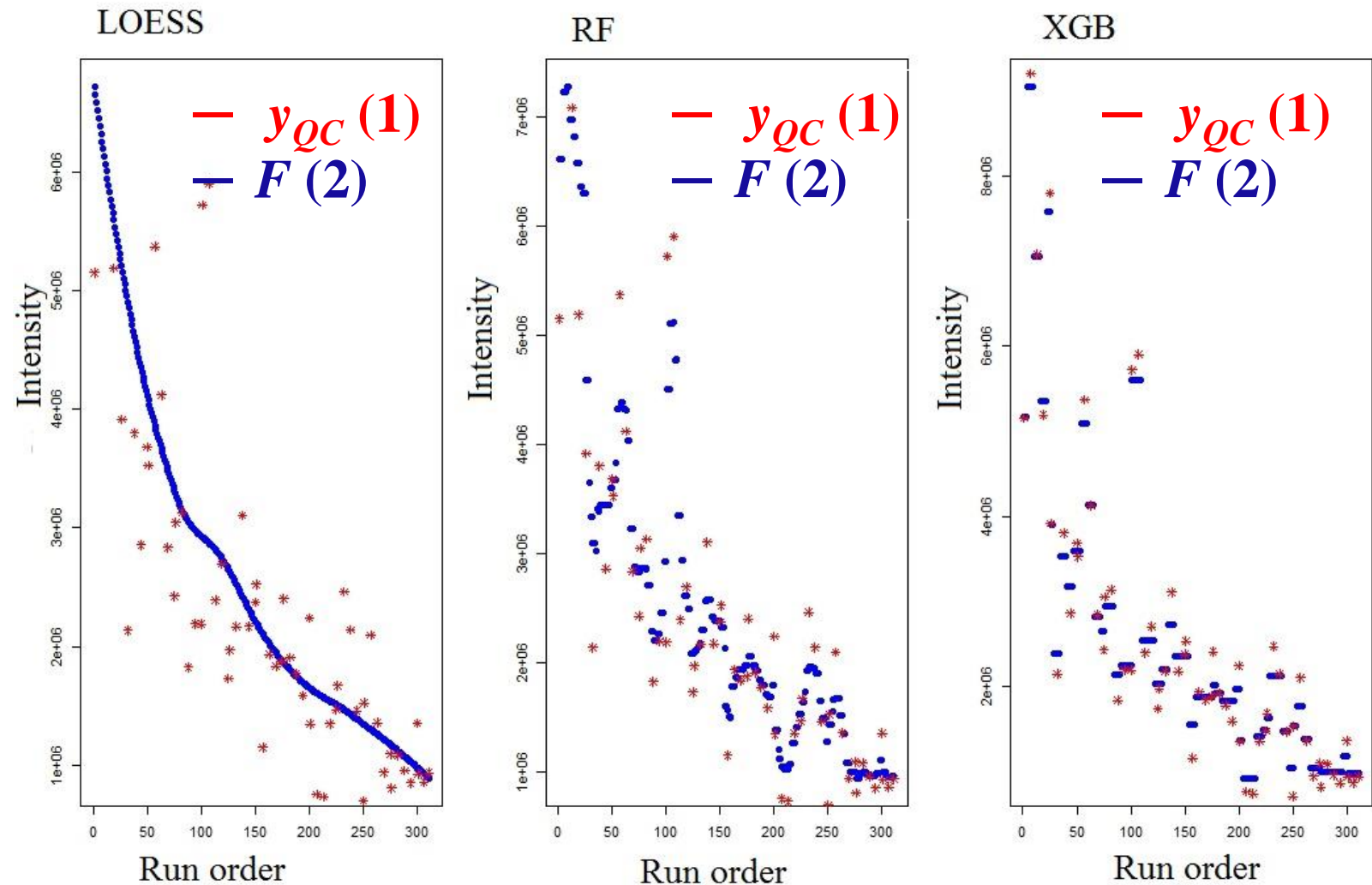
$y$  is an intensity vector (dependent variable),

$x$  is a run order vector (independent variable),

$F$  is a vector of predicted values by the model  $M$  (correction factor),

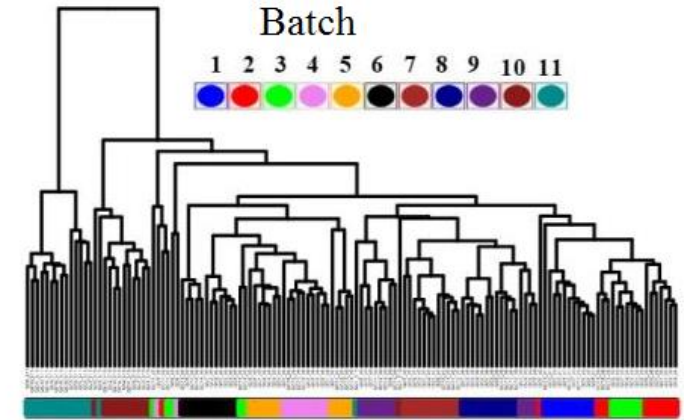
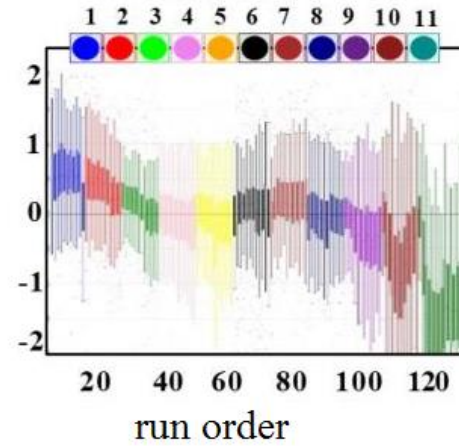
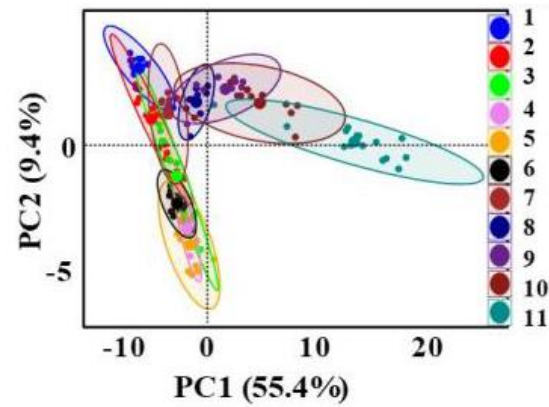
$I'$  is a vector of corrected intensity values,

$I$  is a vector of original intensity values.

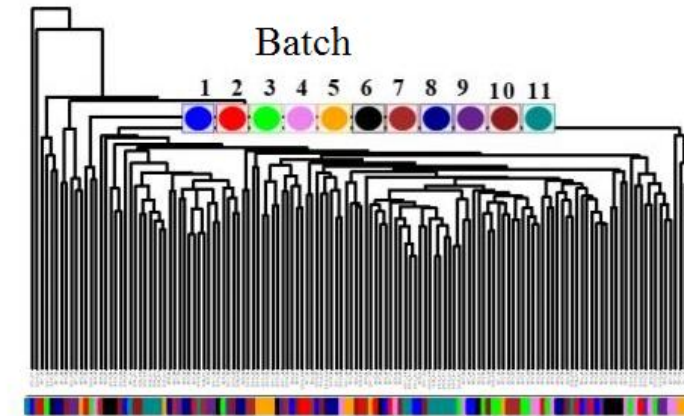
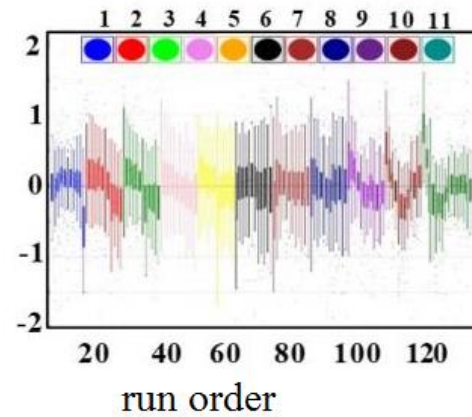
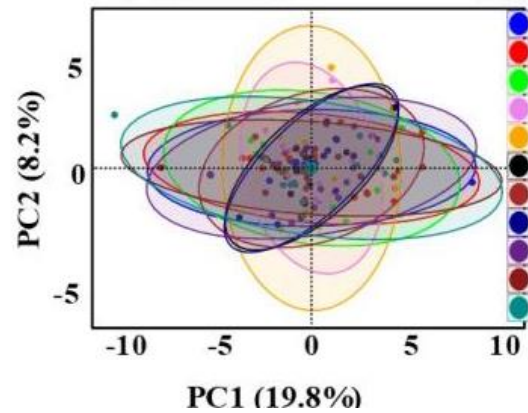


# Batch Effect

## Raw data



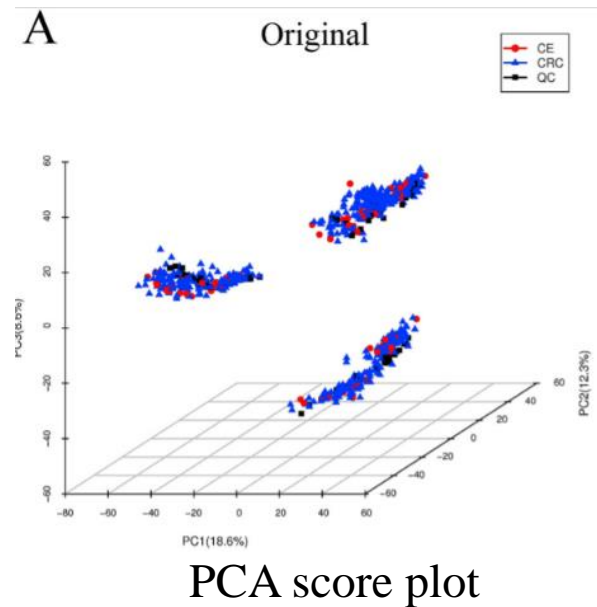
## After correction



# Test data

## Amide [1] metabolomics dataset

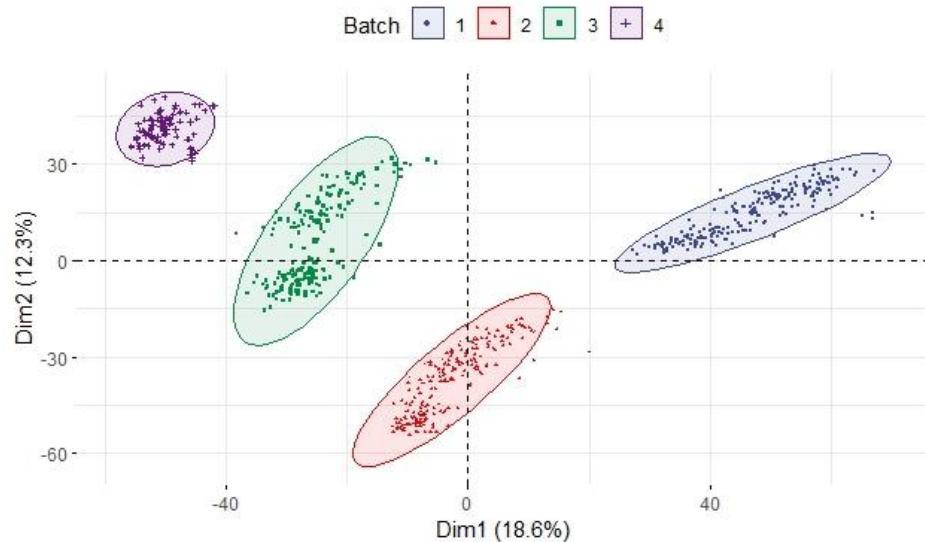
Batch	Number of samples	Number of features
4	729	6969



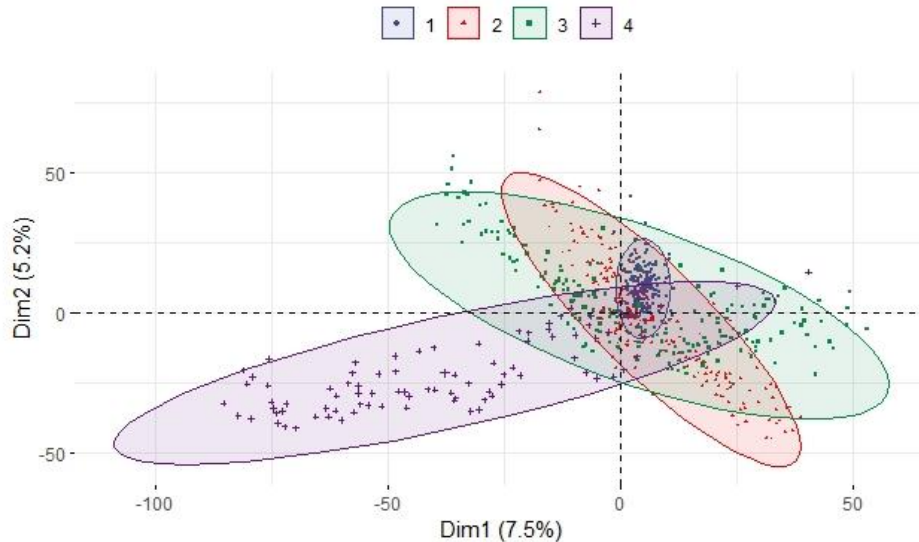


# Batch Effect Elimination

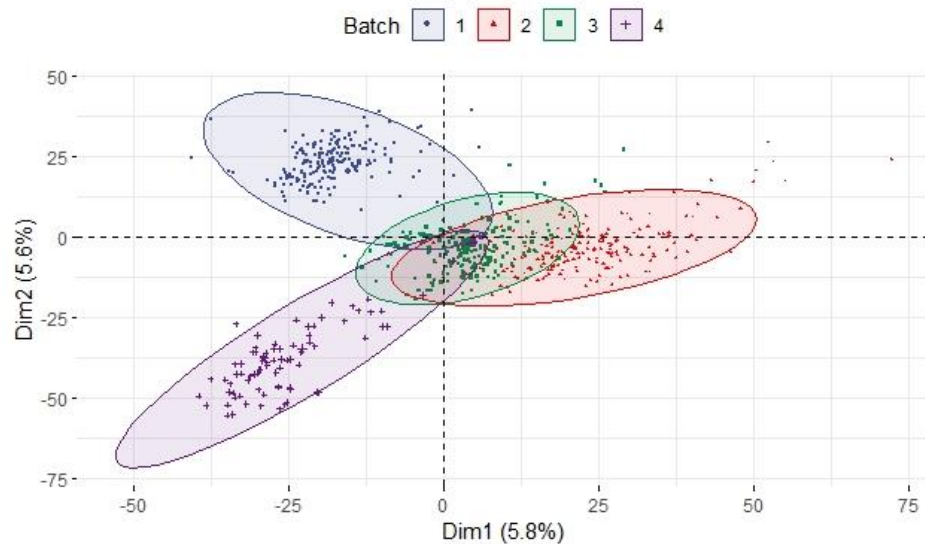
RAW



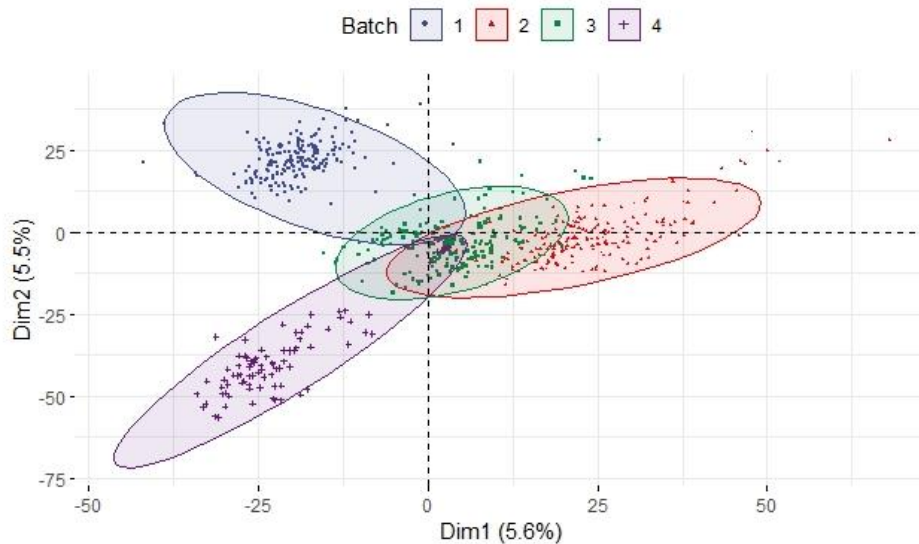
LOESS



RF

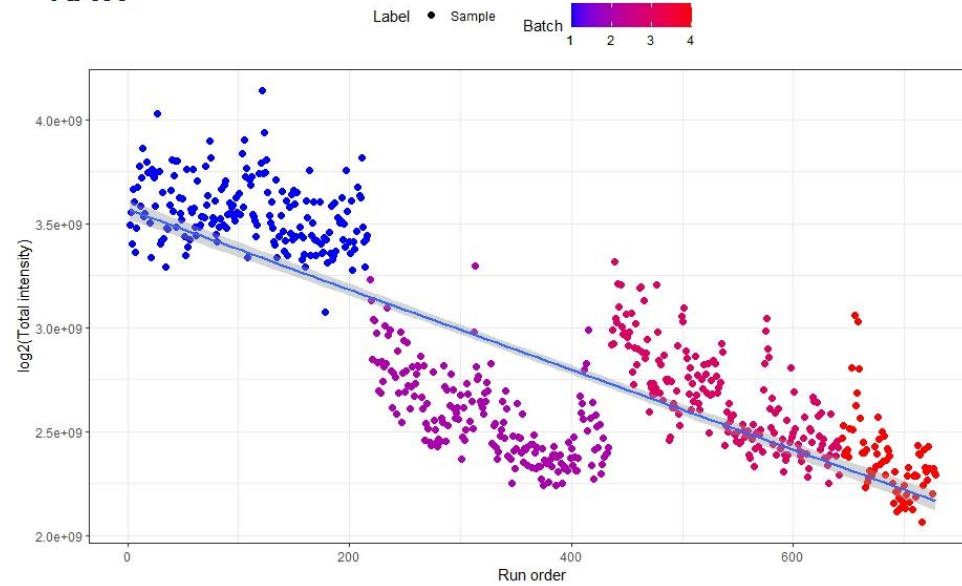


XGB

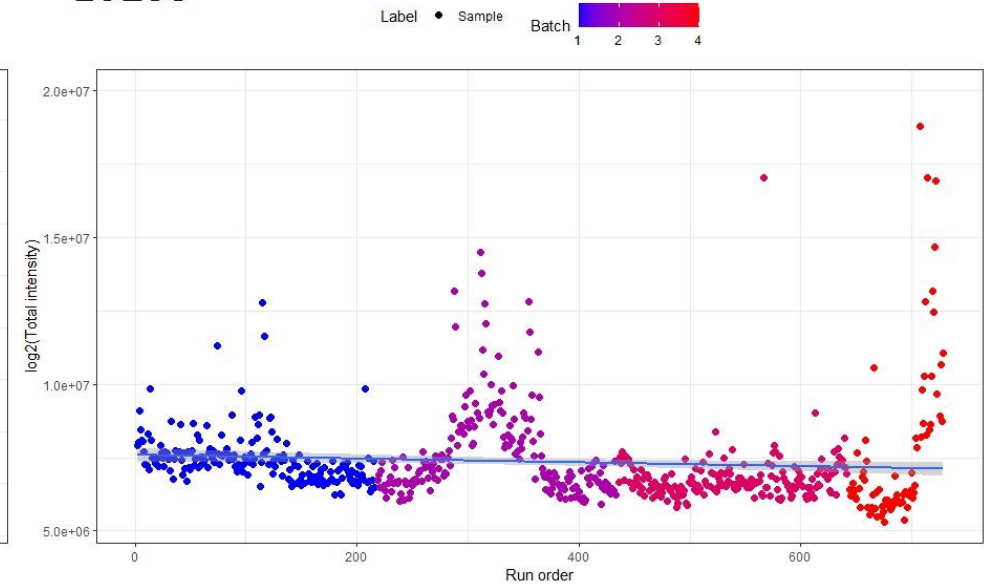


# Signal Drift Elimination

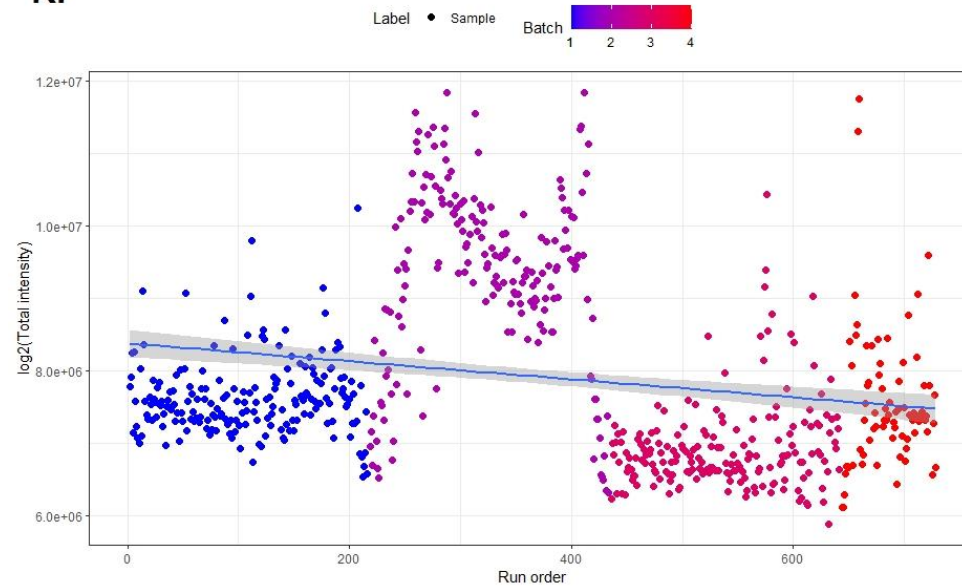
RAW



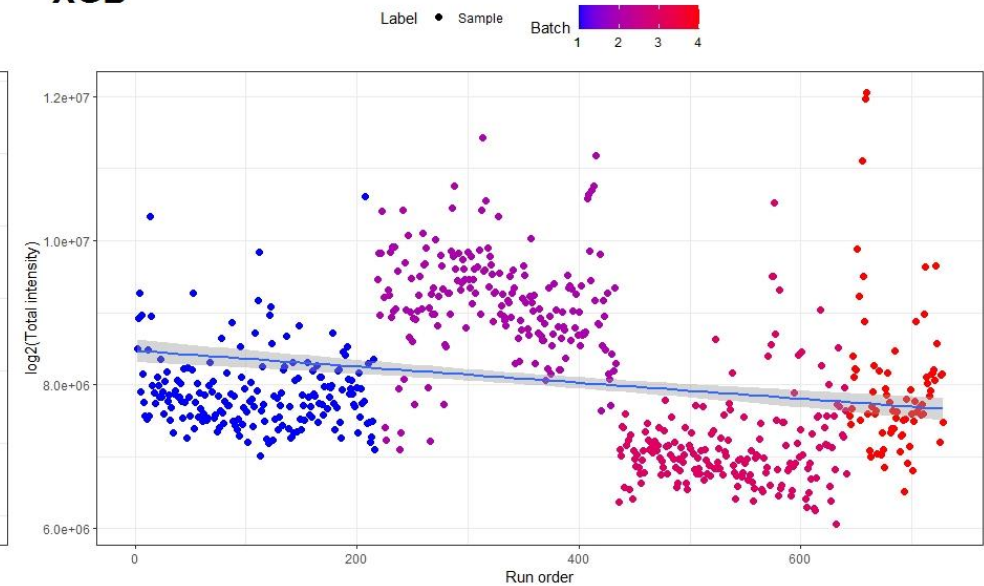
LOESS



RF

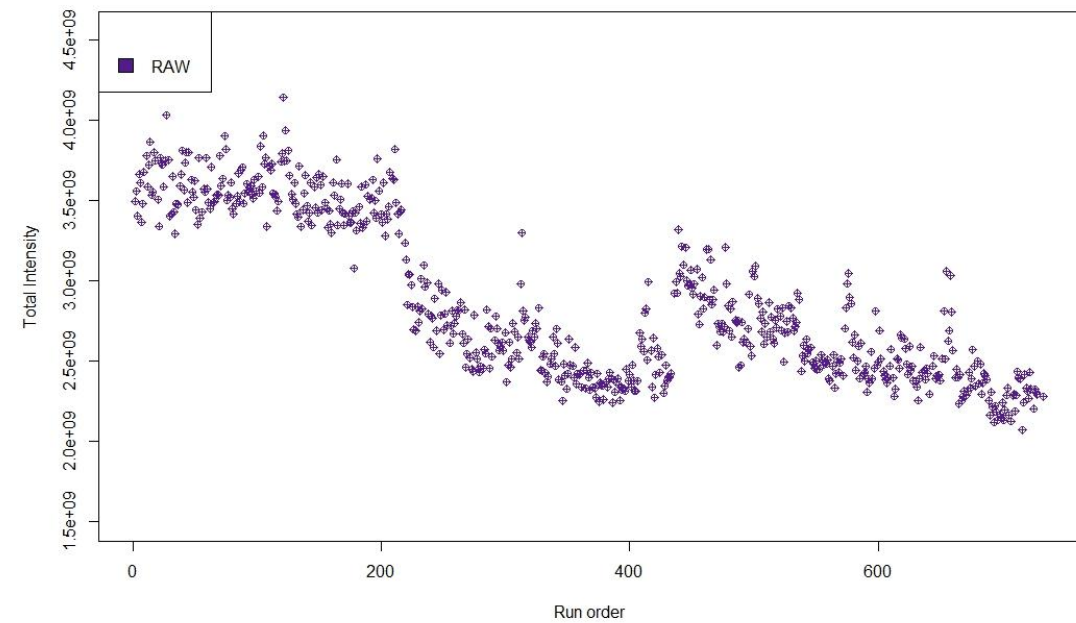
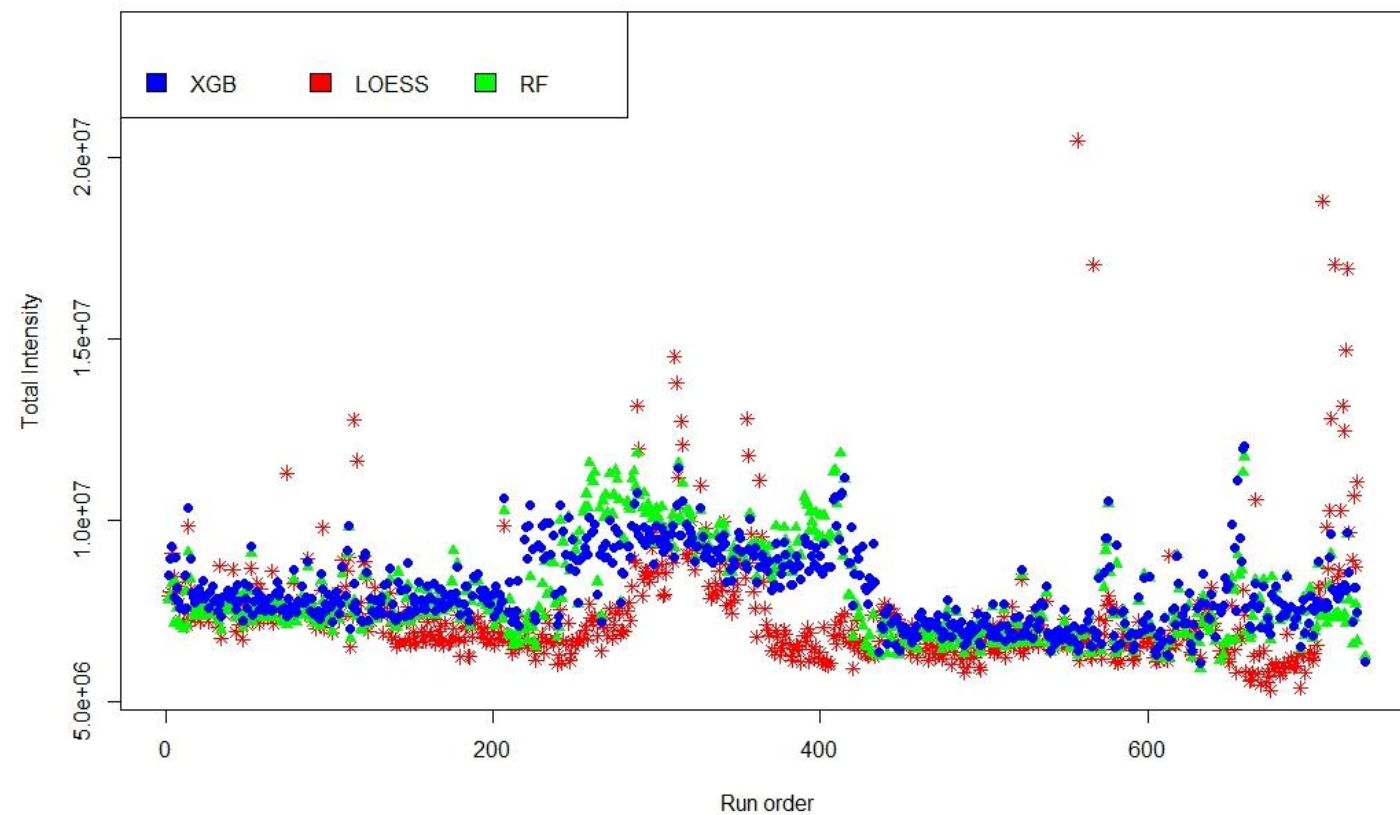


XGB





# Signal Drift Elimination



# Reference, code availability

*Journal of Proteome Research (2021)*

## **Omics Untargeted Key Script: R-Based Software Toolbox for Untargeted Metabolomics with Bladder Cancer Biomarkers Discovery Case Study**

*Plyushchenko, Ivan V.; Fedorova, Elizaveta S.; Potoldykova, Natalia V.; Polyakovskiy, Konstantin A.; Glukhov, Alexander I.; Rodin, Igor A.*



### **QC-XGB:**

[https://github.com/plyush1993/OUKS/blob/5140dd7af725e90abde166312d197409bb210239/Scripts%20\(R\)/4.%20Correction.R#L1005-L1272](https://github.com/plyush1993/OUKS/blob/5140dd7af725e90abde166312d197409bb210239/Scripts%20(R)/4.%20Correction.R#L1005-L1272)

