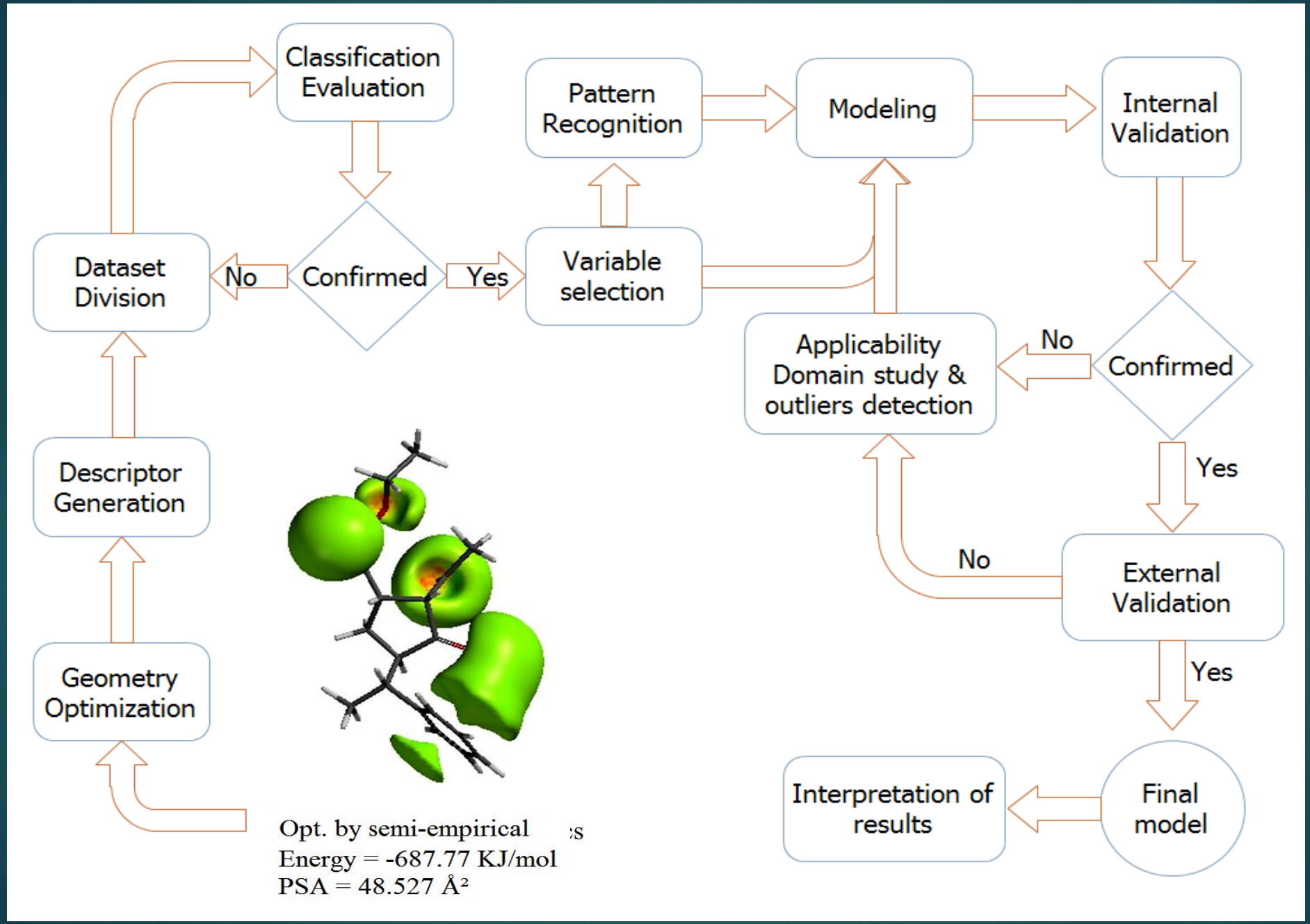
14th International Conference on
Environmental Science and Technology
CEST2015, 3-5 September 2015, Rhodes, Greece

Application of retention time prediction models for suspect and non-target HRMS screening of emerging contaminants in the aquatic environment

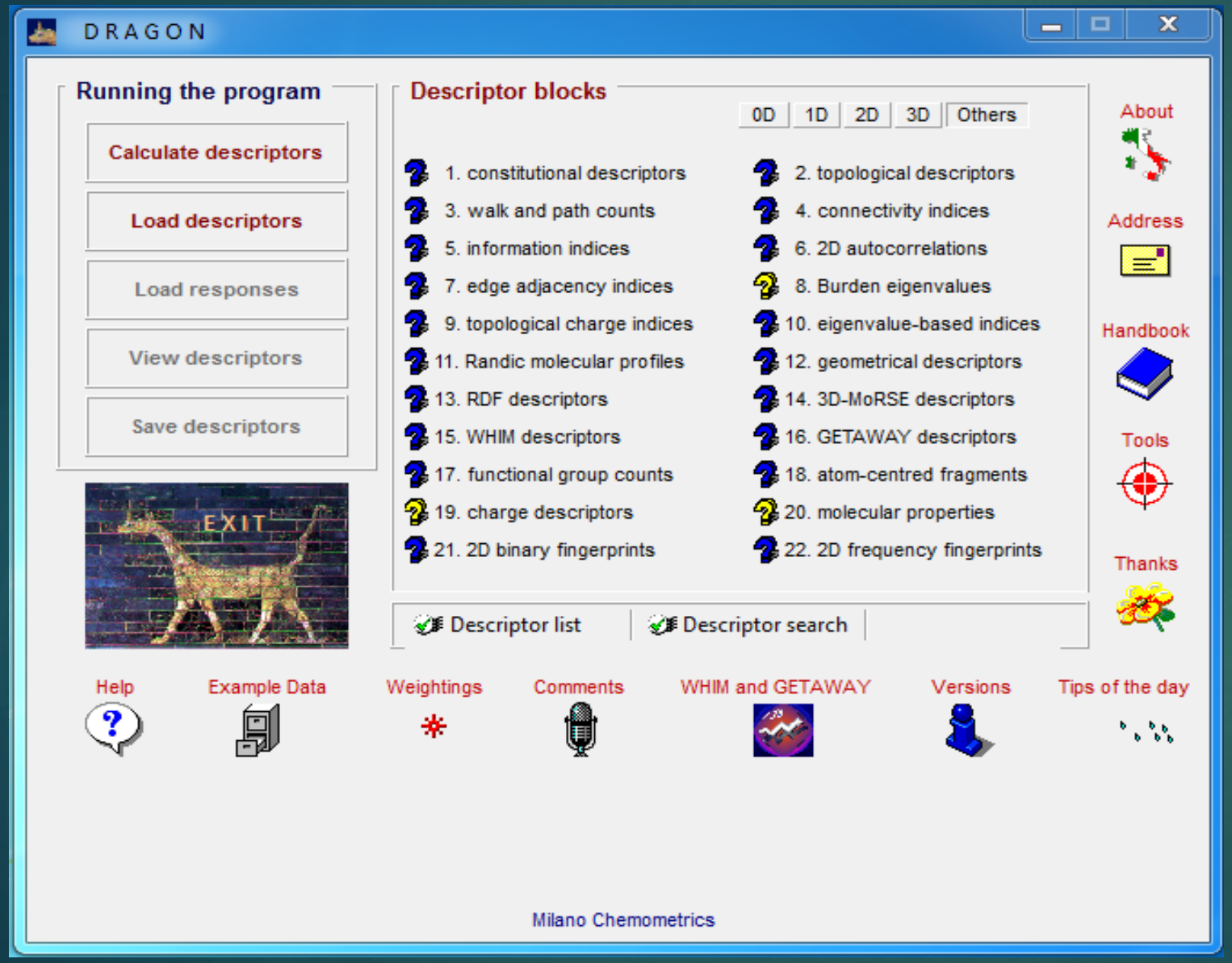
PRESENTER: REZA AALIZADEH

Laboratory of Analytical Chemistry,
Department of Chemistry
National and Kapodistrian
University of Athens

Workflow for performing a QSPR modeling

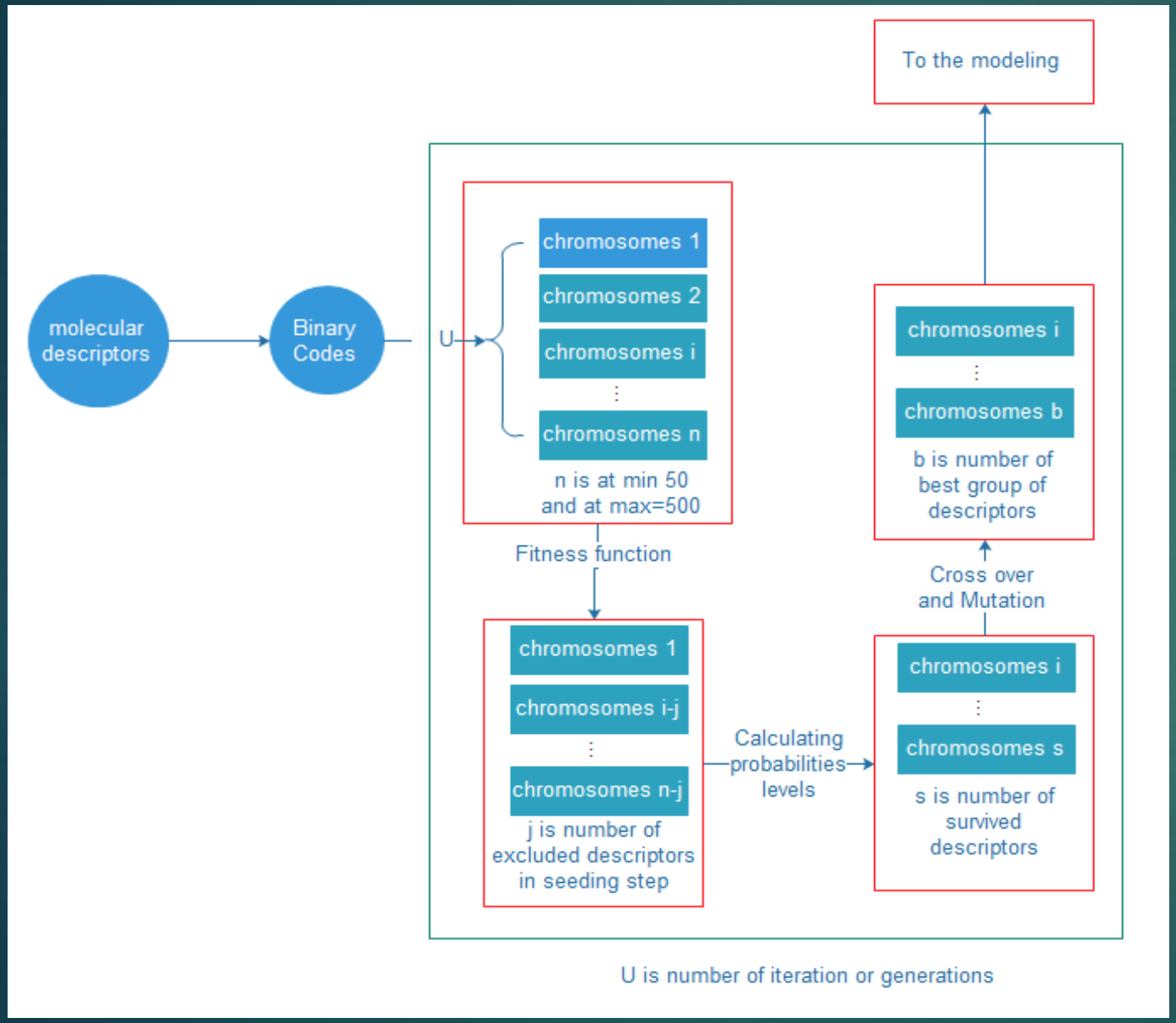


Descriptor calculation

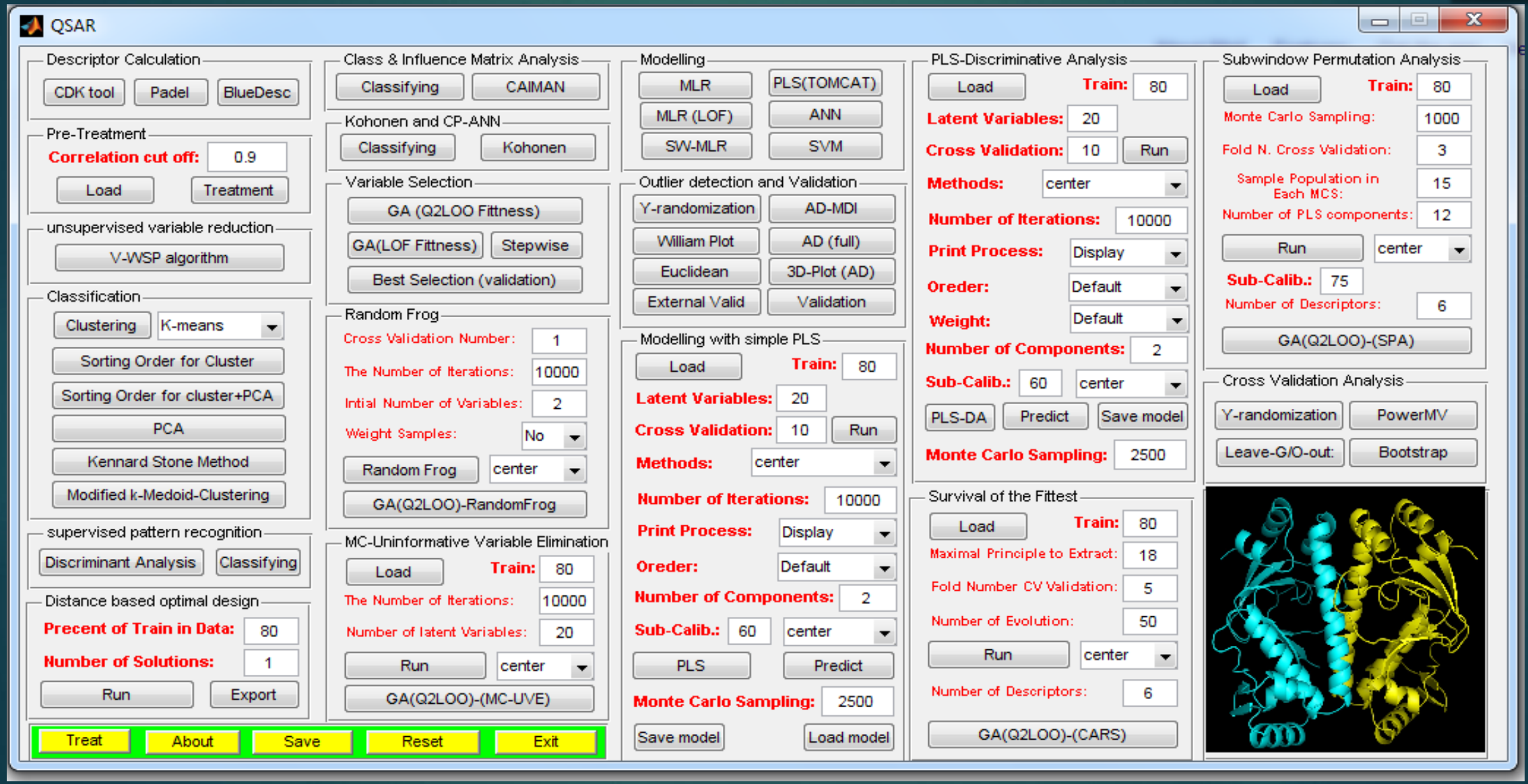


Dataset Division & Accuracy Assessment

Molecular Descriptor Selection



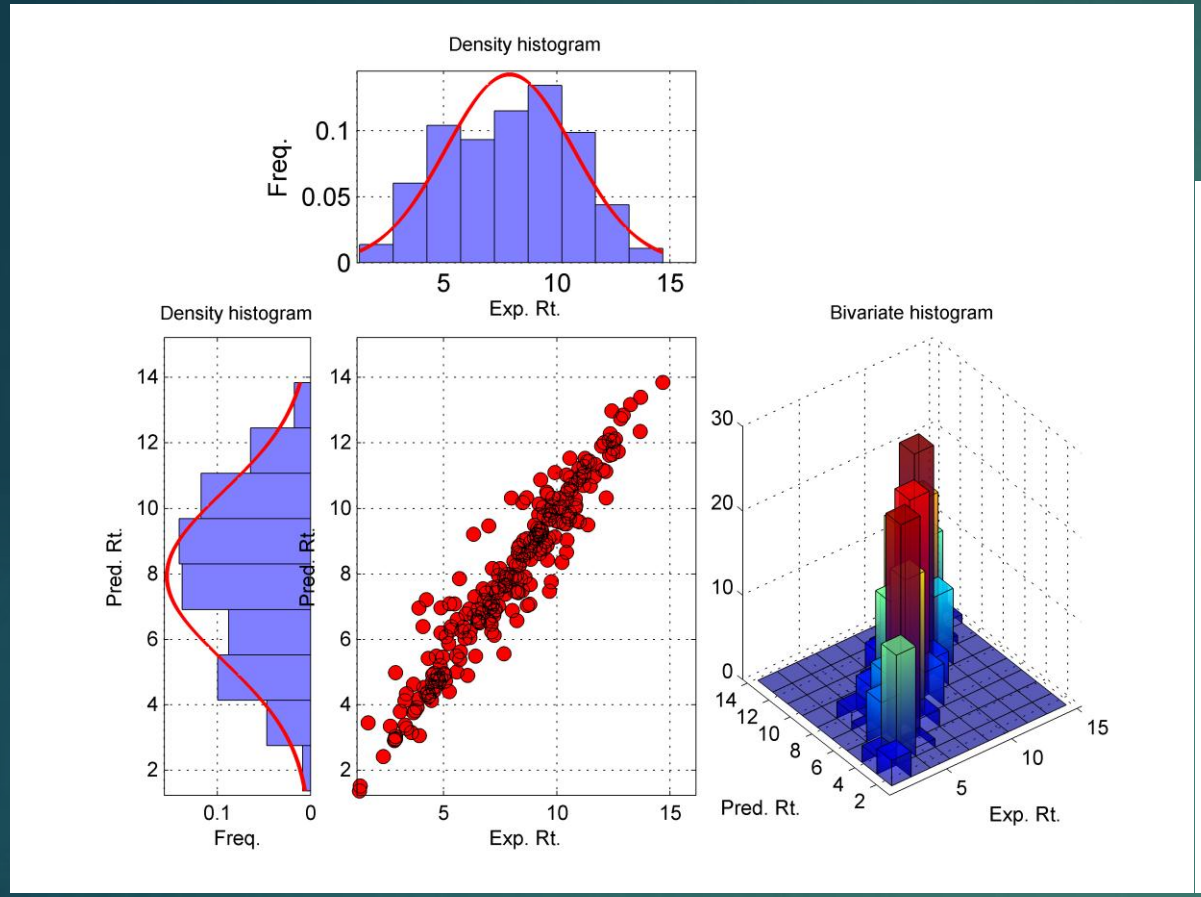
Modeling and whole procedure

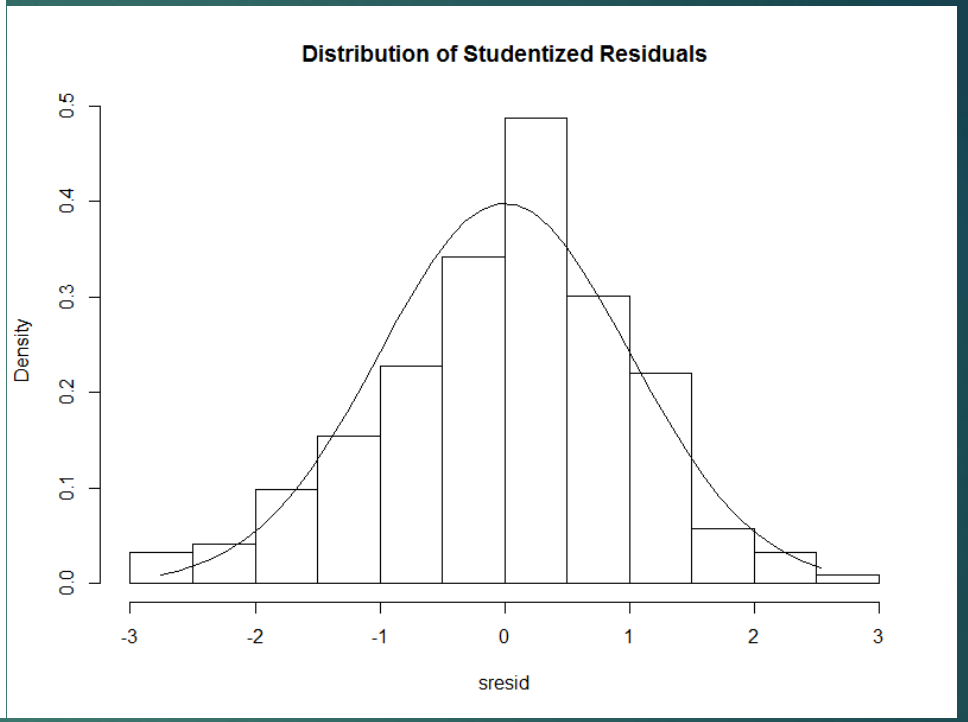


RetTrAMS


Menue
Prediction of Rt:
(Negative) ESI
(Positive) ESI
RP
HILIC
XlogP
LogD
Linear
Non-Linear
Applicability domain study:
Distribution
PCA
Dendrogram
Plot OTrAMS
OTrAMS
Save OTrAMS
Advanced outlier study:
Load
Mapping
Molecule:
Ex.: m25
Search
Searching Database:
Name:
Ex.: Amitrol
Search
Manual
Reset
Exit
This package is part of RetTrAMS program.
Drewing Structure:
Download Online
RetLogD
LogD Calculator
Assign pH (Negative)
Assign pH (Positive)
pH:
Ex.:3.6
LogD
LogP
load
Calculate
Cancel
LogD Calculator is part of RetTrams program.
LogD calculation is performed by ChemAxon program.
XLOGPcal
XlogP Calculator
SMILES:
Ex.: NC1C=CC(Cl)=CC=1C(=O)C1C=CC=CC=1
Calculate
Padel
Cancel
Install Packages
XlogP Calculator is part of RetTrams program.

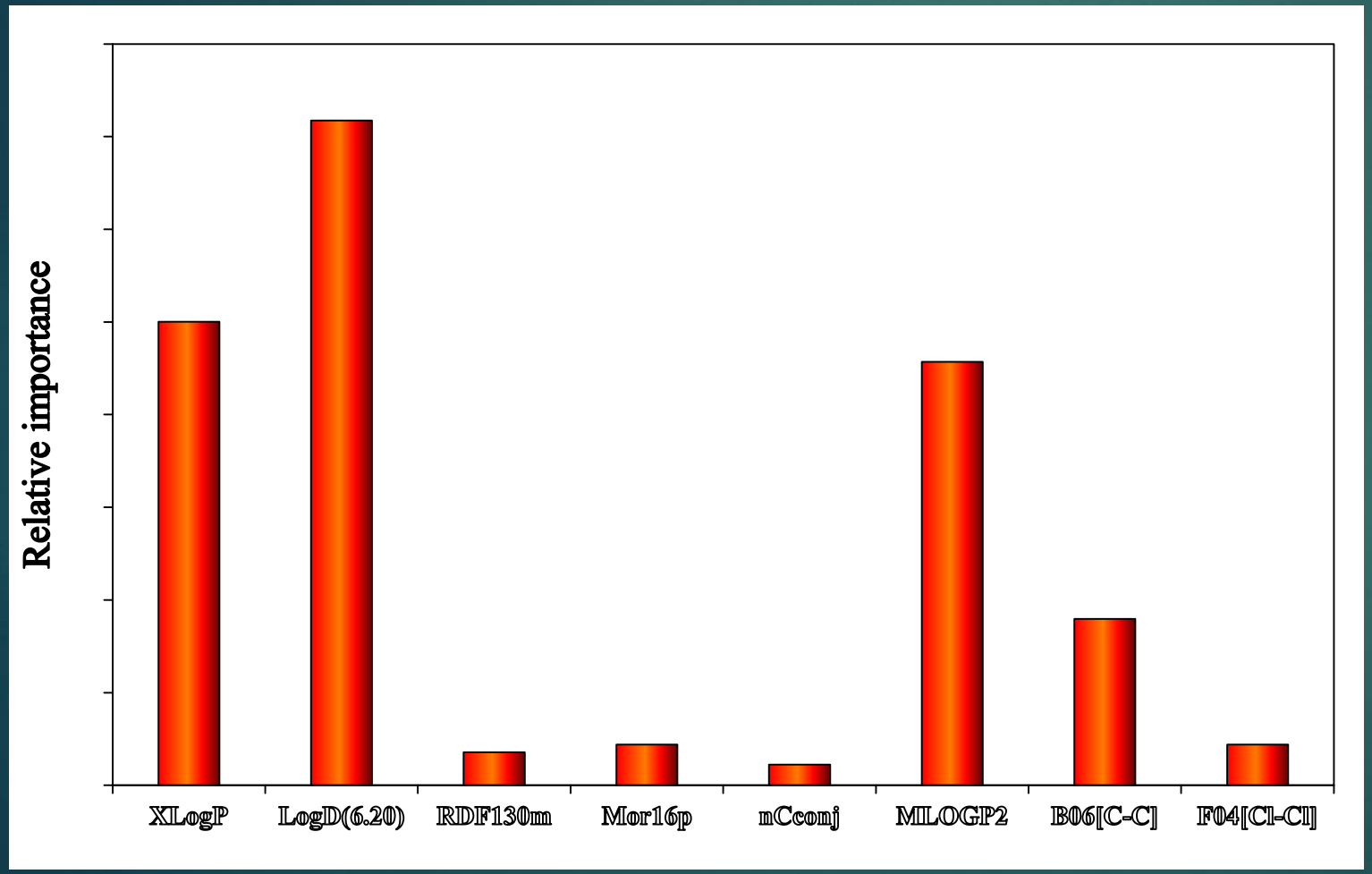
RP_(-)ESI





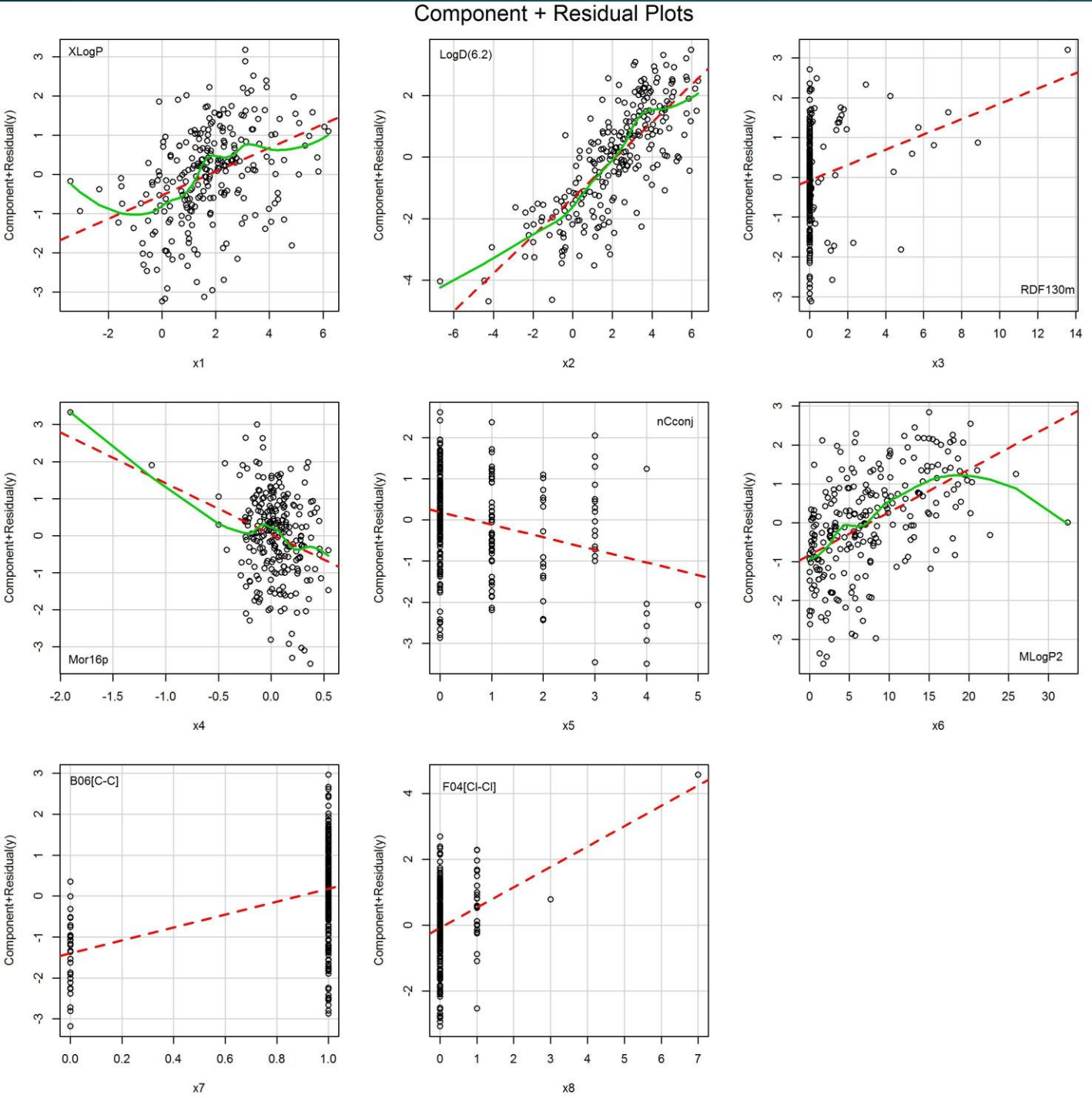
	Training			Test		
	R2	**RMSE**	**F**	**R2**	**RMSE**	**F**
MLR	0.841	1.116	156.651	0.844	1.103	36.012
SVM	0.911	0.842	278.615	0.857	1.052	35.009

RP_(-)ESI

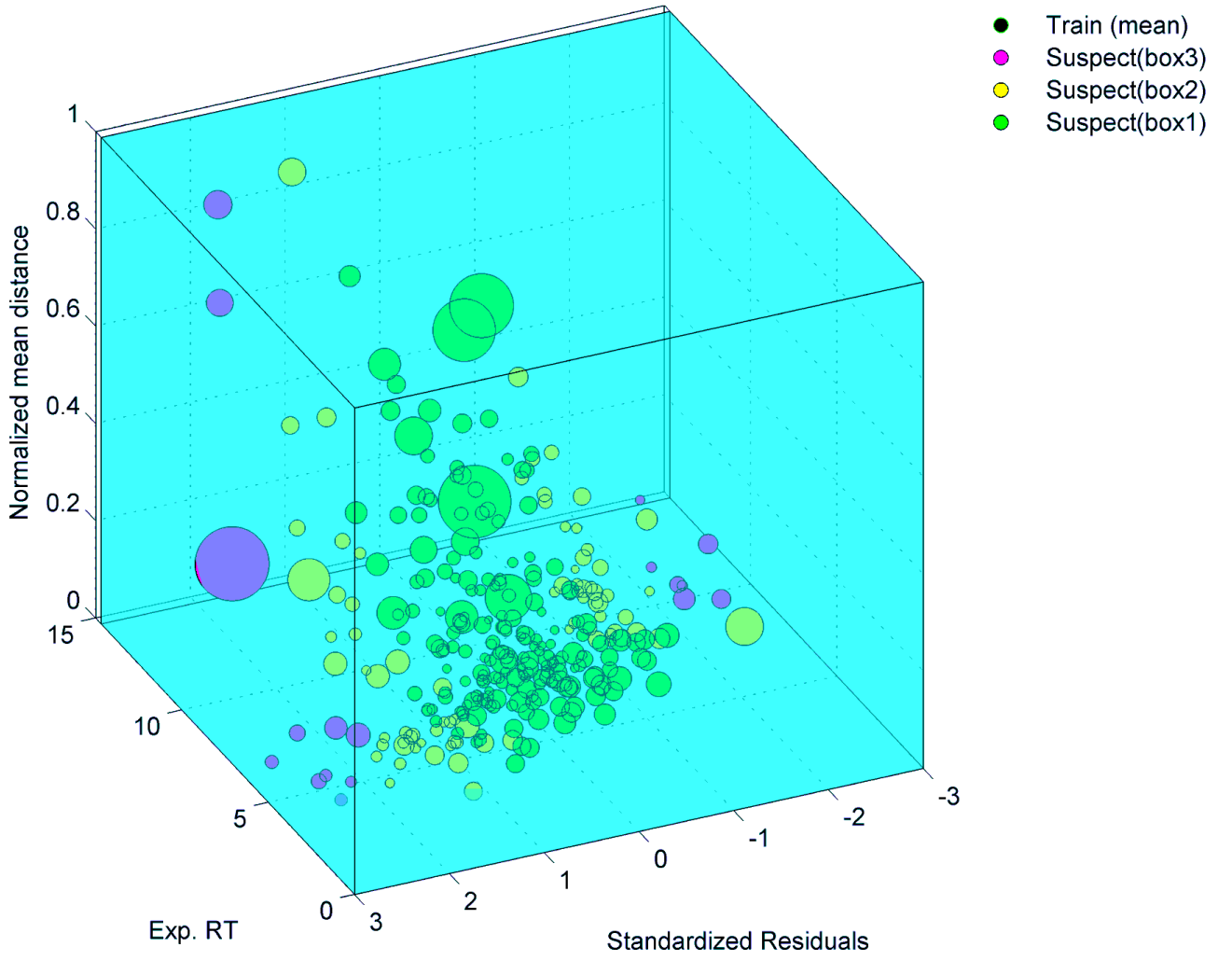


Negative Ionization	
XLogP	+
LogD (pH=6.2)	+
RDF130m	+
Mor16p	-
nCconj	-
MLOGP2	+
B06(C-C)	+
F04(Cl-Cl)	+

RP_(-)ESI

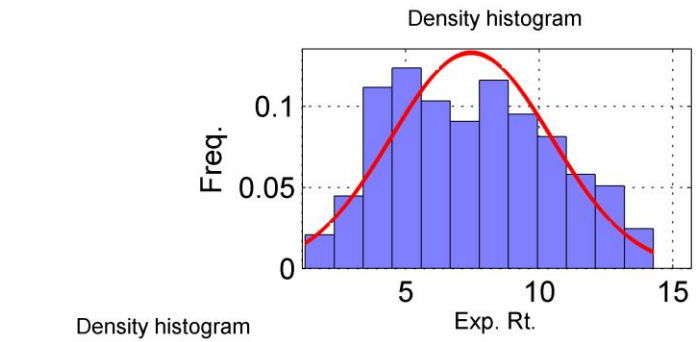

Component + Residual Plots
XLogP
LogD(6.2)
RDF130m
Mor16p
nCconj
MLogP2
B06[C-C]
F04[Cl-Cl]
Component+Residual(y)
x1
x2
x3
x4
x5
x6
x7
x8

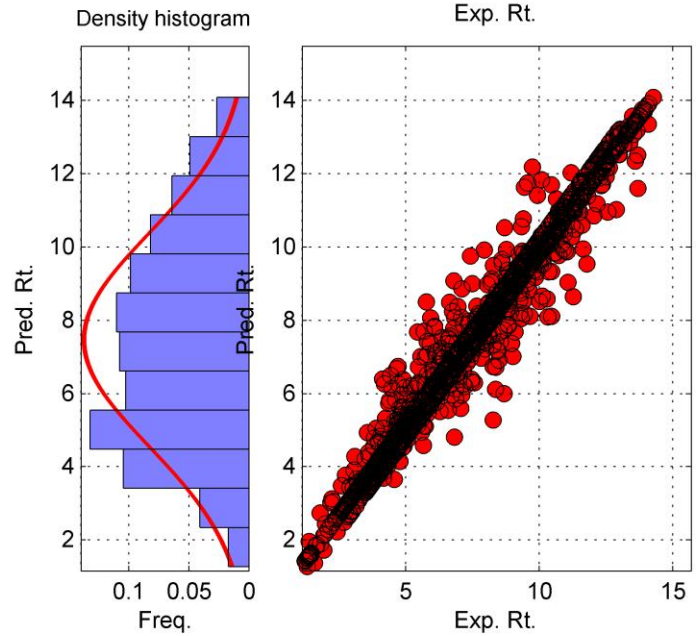
RP_(-)ESI

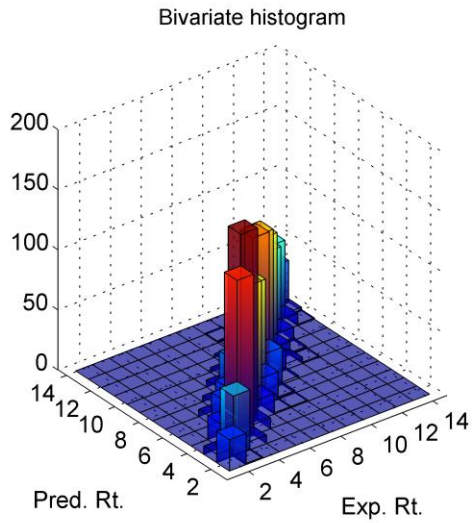


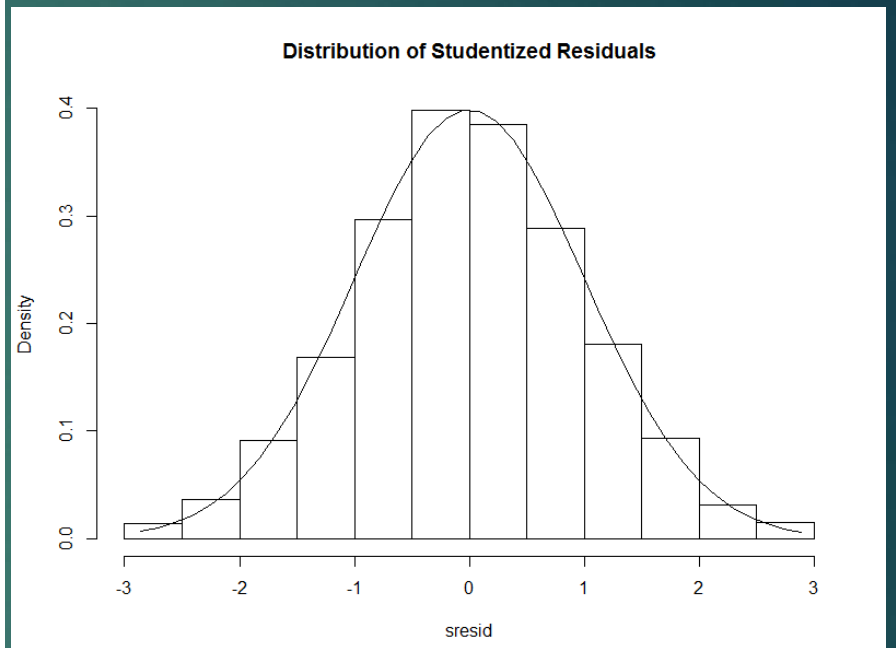
$$\text{Standardized Residuals} = \frac{r(i)}{\left[\text{sqrt}\left(\frac{r' * r}{n-p} * (1-h(i))\right)\right]}$$

RP_(+)ESI



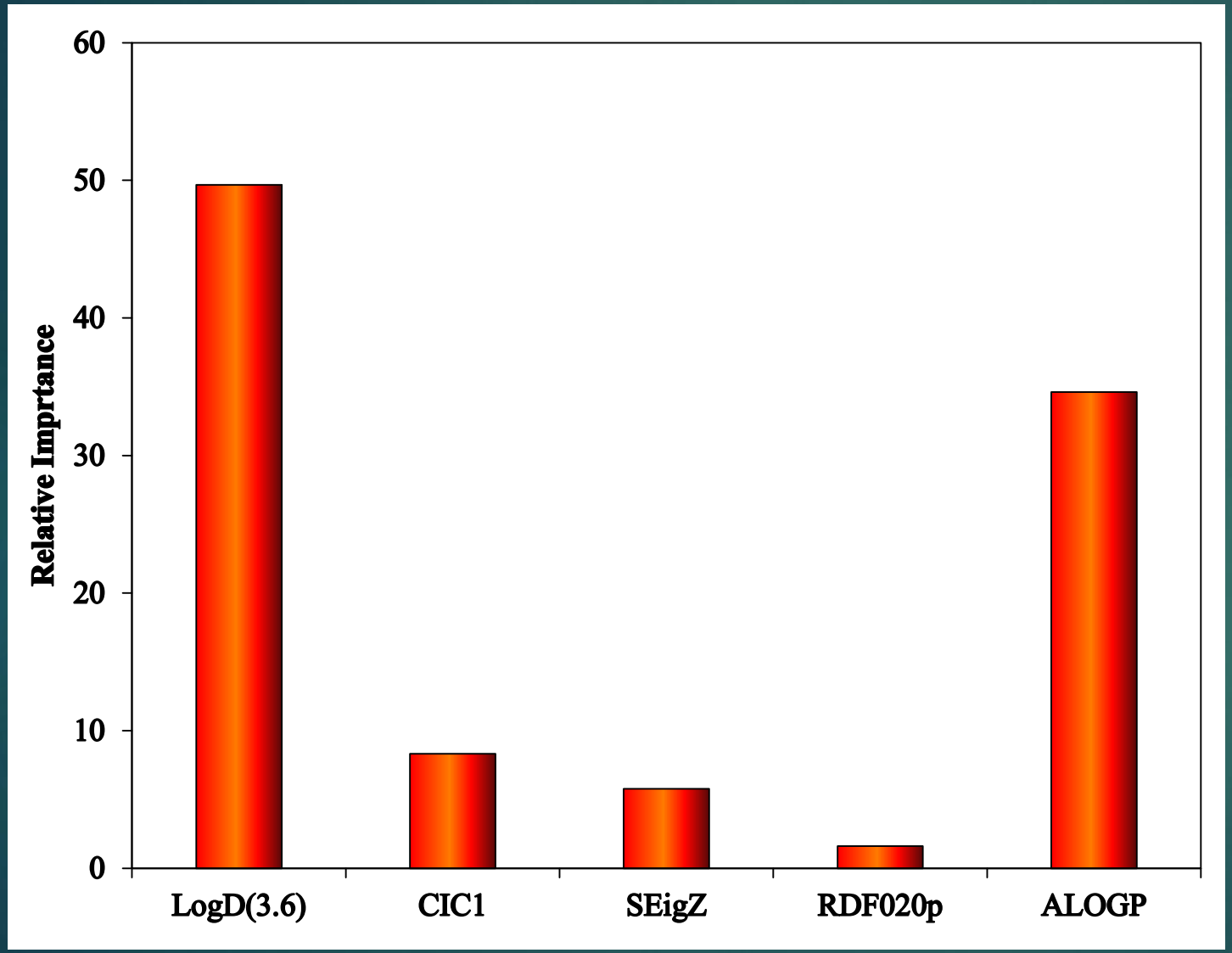




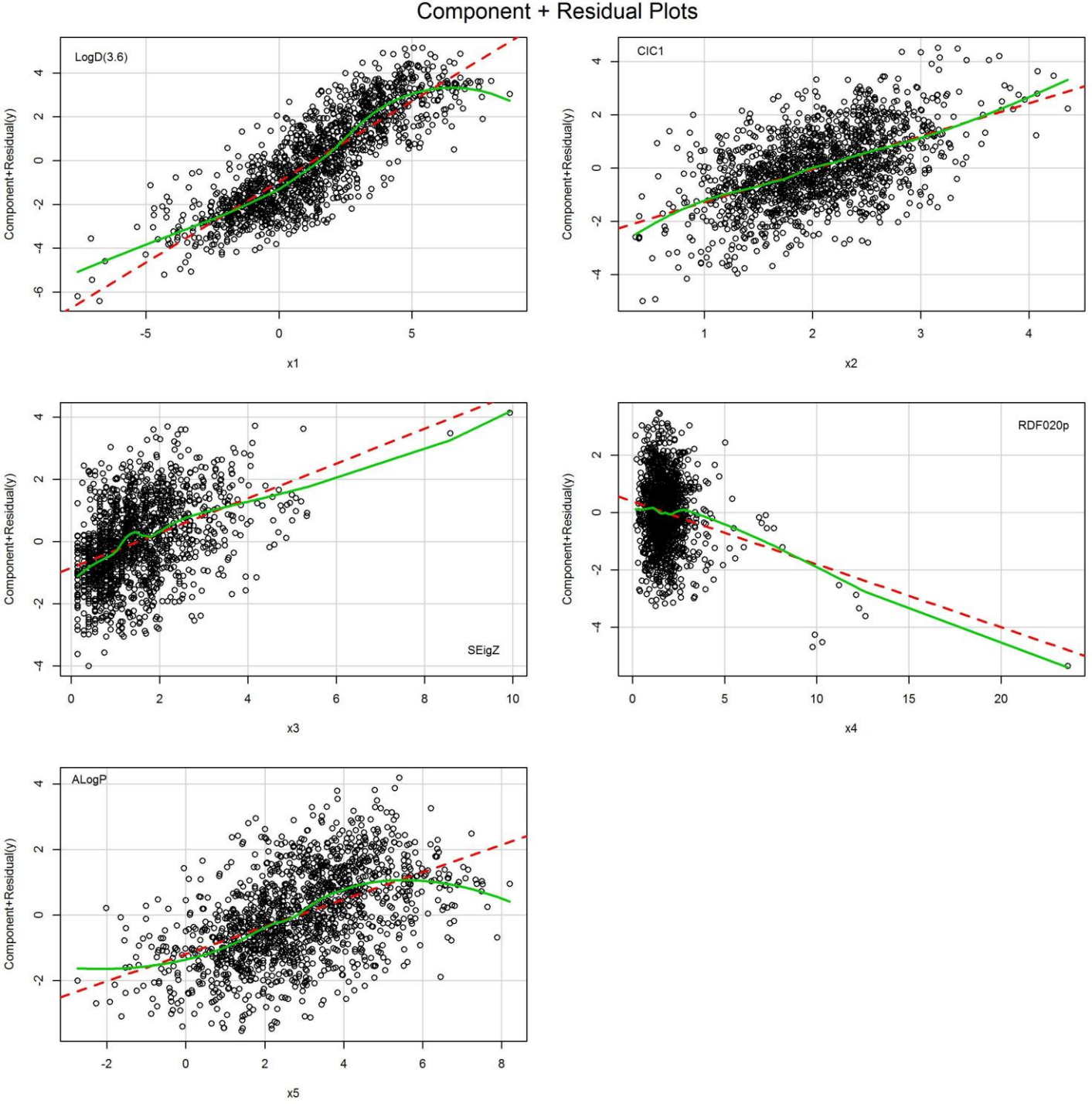


	Training			Test		
	R2	RMSE	F	R2	RMSE	F
MLR	0.844	1.187	1569.088	0.848	1.179	433.091
SVM	0.971	0.514	9281.657	0.881	1.042	574.810

RP_(+)ESI

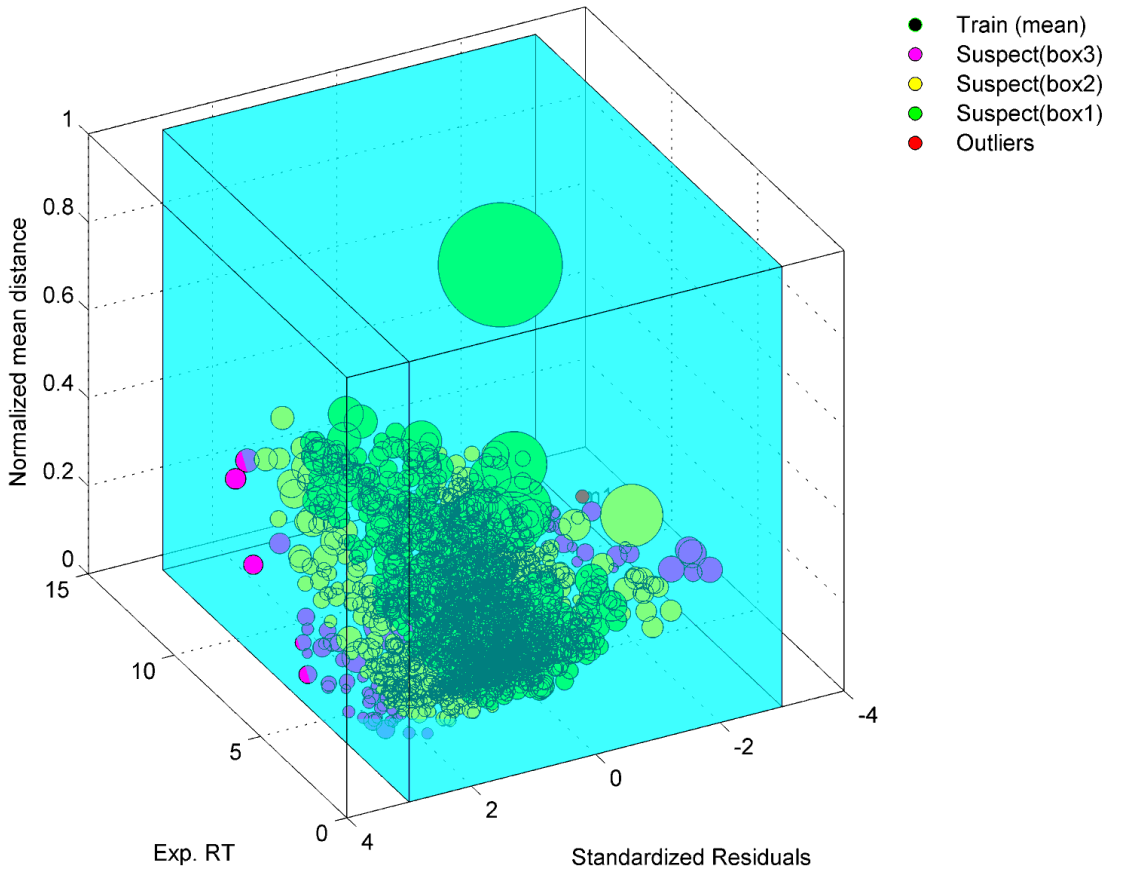


Positive Ionization	
LogD (pH=3.6)	+
CIC1	+
SEigZ	+
RDF020p	-
AlogP	+


Component + Residual Plots
LogD(3.6)
CIC1
SEigZ
RDF020p
ALogP
Component+Residual(y)
x1
x2
x3
x4
x5

RP_(+)ESI

RP_(+)ESI



	Number of compounds inside each box	Percent of compounds inside each box
box1	1259	69
box2	483	26
box3	87	5
box4	1	0

Protocols

- To accept or reject a suspect structure, perform RetTrAMS and OTrAMS → locate the points in boxes
- If the suspect compound locates in box 1 and box 2 → the suspect structure is accepted.
- If the suspect compound locates in box 3 → further validation should be done.

- If the suspect compound locates in box 4 → the suspect structure is rejected.

Acknowledgments

Nikolaos S. Thomaidis **Anna Bletsou**

Pablo Gago Ferrero

Thanks for your attention!

This research has been co-financed by the European Union and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) – ARISTEIA 624 (TREMEPOL project).