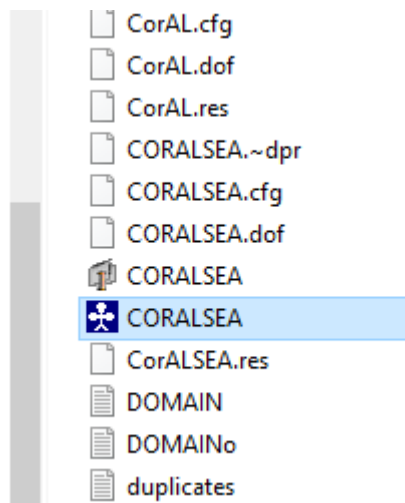


The version of the program developed in 2025 has some new features, which are demonstrated using a certain artificial example. All necessary files are contained in the CORALSEA-2025 folder. To use the new features, you can perform the steps described on your computer without changing anything in the method or working files. In the future, you can then use your files and select your CORALSEA method.

It is considered a situation where the user has a data set in a file "TotalSet290.txt". The new version of the program will remove duplicates, if any (mandatory). The Las Vegas algorithm can be used or not. All other methods and options of the program are the same as before.

#### 1. Run the program CORALSEA.exe



CORAL: Please load method or make a random split

**Don't show optimization**

MODEL Y↑

Load method Method.txt

Active Training set (A) EXPERIMENT

Preparation A-,P-,C-, and V-files

MODEL Y↑

Passive Training set (P) EXPERIMENT

Select Las Vegas algorithm parameters

MODEL Y↑

Calibration set (C) EXPERIMENT

Split Info

W%	N111	N110	N101	N100	Nall	Defect	Density
0	0	0	0	0	0	0	0

SMILES/quasi-SMILES on A,P,C sets

**CORALSEA-2025rd**

☐ GRAPH ☐ HSG ☐ HFG ☐ GAO

e0 e1 e2 e3 p2 p3 p4 s2 s3 nn

☐ R3 ☐ R4 ☐ R5 ☐ R6 ☐ R7

☐ SMILES ☐ Chaos ☐ s ☐ ss ☐ sss ☐ BOND ☐ NOSP ☐ HALO ☐ HARD

Atoms pairs proportions  
Cl Br N O S P = #

F. ☐ Cl. ☐ Br. ☐ N. ☐ O. ☐ S. ☐ P. ☐ =

☐ xyx ☐ xyxx ☐ xyzyx

☐ Classification model

☐ CLASSIC SCHEME

☒ CORRELATION BALANCE

☐ Index of Ideality of Correlation

☐ Correlation Intensity Index

☐ Conformism coefficient

dR weight ☐ IIC weight ☐ CII weight ☐ CCCP weight

D start ☐ d limit ☐ Nepoch ☐

Threshold start ☐ Threshold maximal ☐ Number of probes ☐

Model Details.txt

Number of promoters increase/decrease ☐

STOP and SAVE current CWs ☐ DemoDCW ☐ EvolutionCorr

**EXIT**

2. Click of button indicated in the red frame

CORAL: Please load method or make a random split

**Don't show optimization**

MODEL Y↑

Load method Method.txt

Active Training set (A) EXPERIMENT

**Preparation A-,P-,C-, and V-files**

MODEL Y↑

Passive Training set (P) EXPERIMENT

Select Las Vegas algorithm parameters

MODEL Y↑

Calibration set (C) EXPERIMENT

Split Info

W%	N111	N110	N101	N100	Nall	Defect	Density
0	0	0	0	0	0	0	0

SMILES/quasi-SMILES on A,P,C sets

**CORALSEA-2025rd**

☐ GRAPH ☐ HSG ☐ HFG ☐ GAO

e0 e1 e2 e3 p2 p3 p4 s2 s3 nn

☐ R3 ☐ R4 ☐ R5 ☐ R6 ☐ R7

☐ SMILES ☐ Chaos ☐ s ☐ ss ☐ sss ☐ BOND ☐ NOSP ☐ HALO ☐ HARD

Atoms pairs proportions  
Cl Br N O S P = #

F. ☐ Cl. ☐ Br. ☐ N. ☐ O. ☐ S. ☐ P. ☐ =

☐ xyx ☐ xyxx ☐ xyzyx

☐ Classification model

☐ CLASSIC SCHEME

☒ CORRELATION BALANCE

☐ Index of Ideality of Correlation

☐ Correlation Intensity Index

☐ Conformism coefficient

dR weight ☐ IIC weight ☐ CII weight ☐ CCCP weight

D start ☐ d limit ☐ Nepoch ☐

Threshold start ☐ Threshold maximal ☐ Number of probes ☐

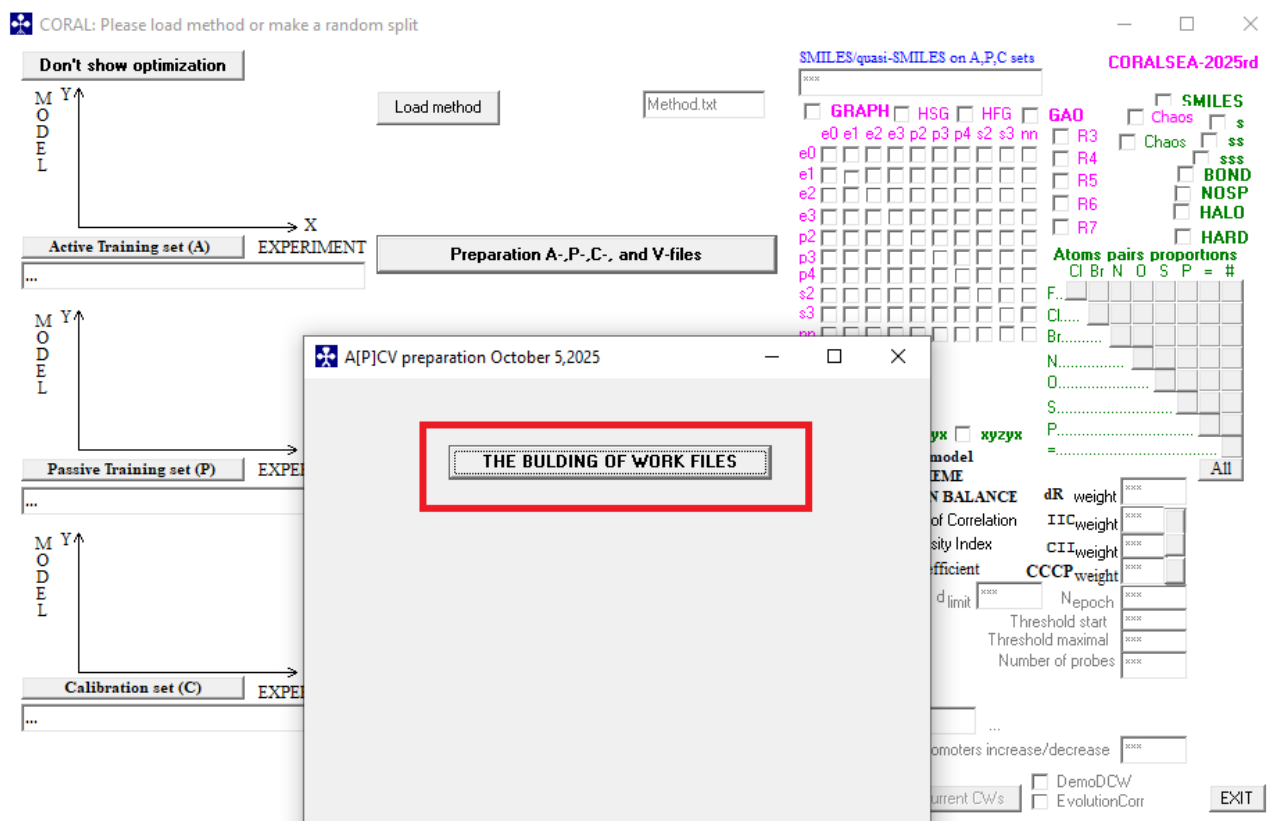
Model Details.txt

Number of promoters increase/decrease ☐

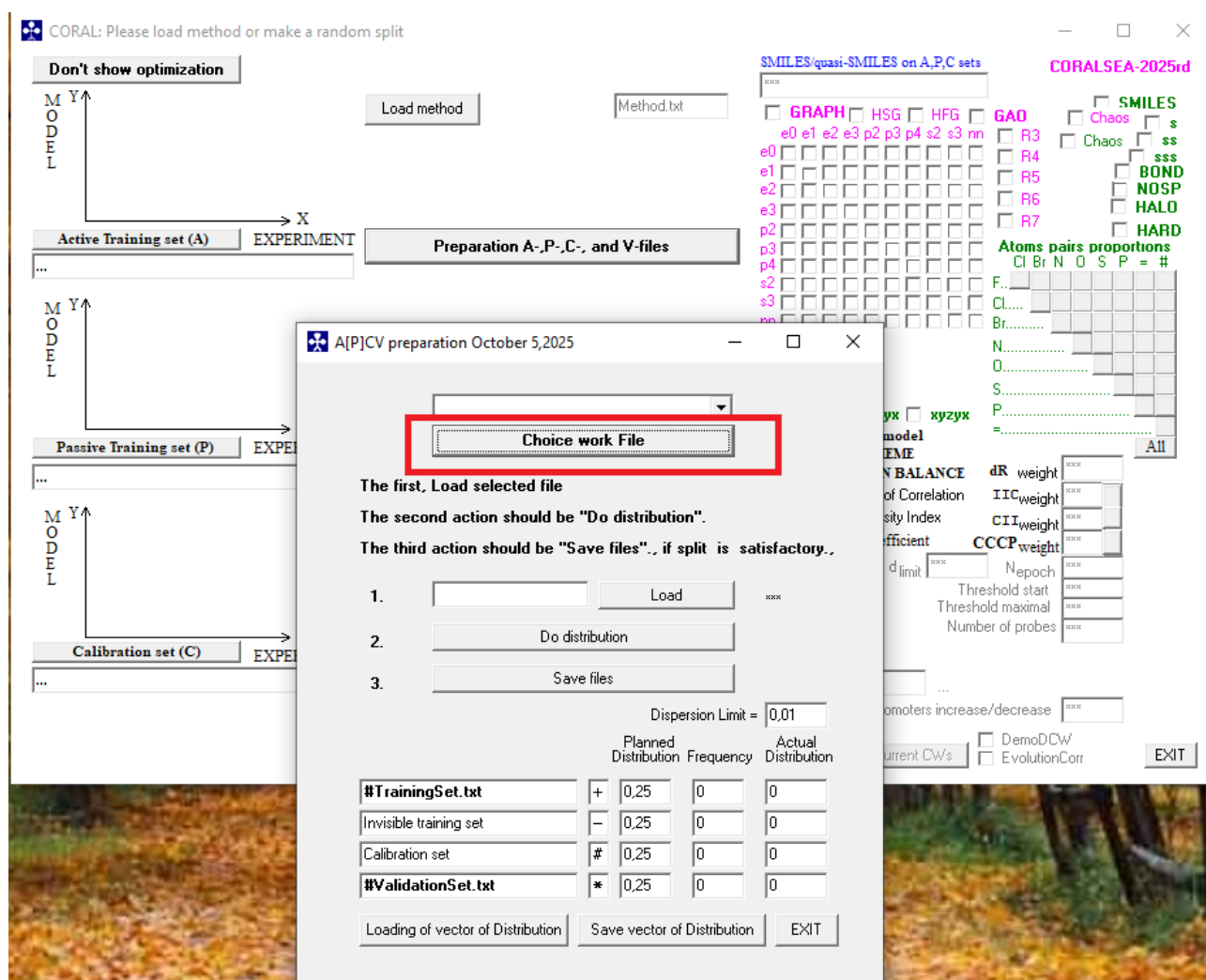
STOP and SAVE current CWs ☐ DemoDCW ☐ EvolutionCorr

**EXIT**

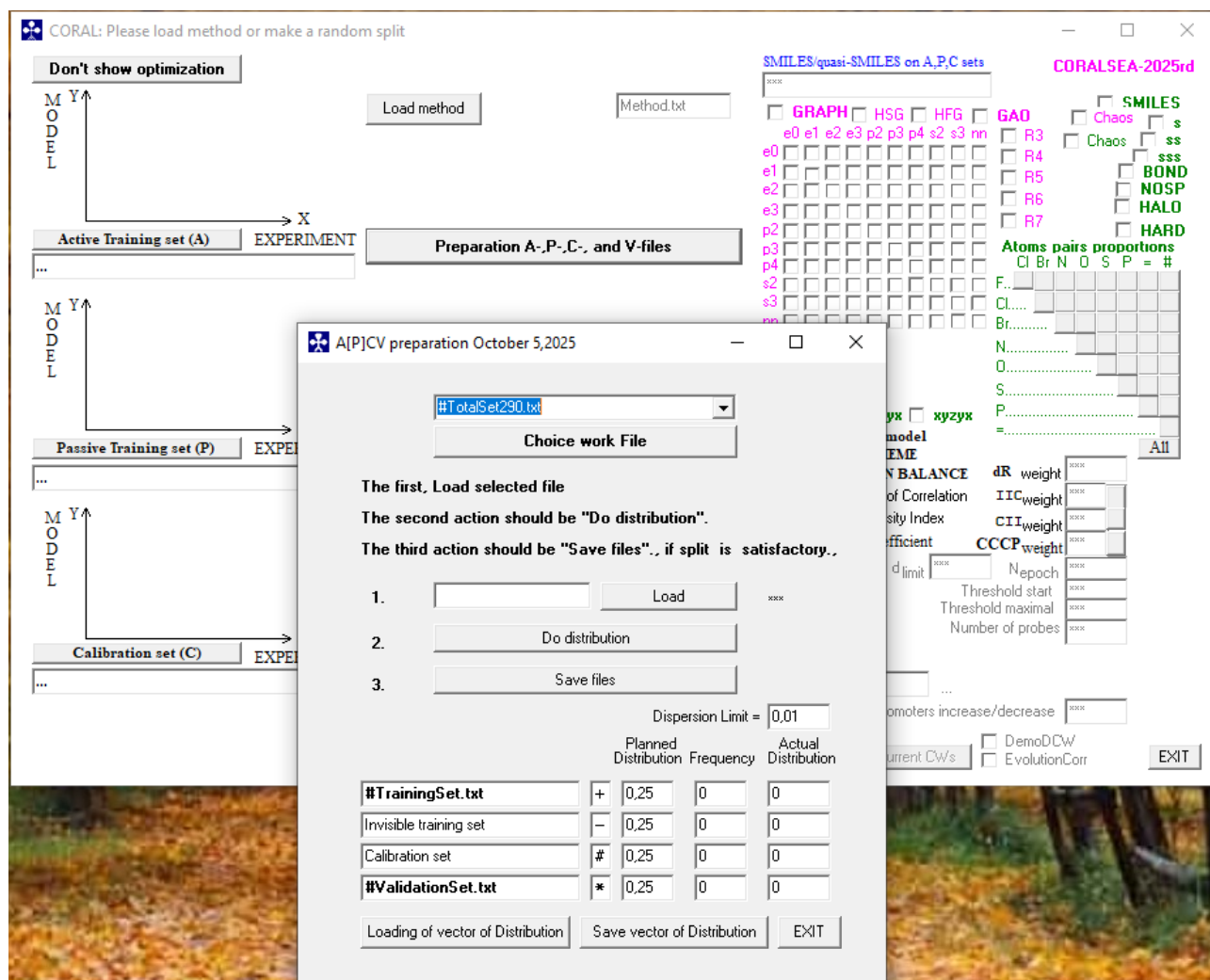
3. Again, click the button indicated in red



4. Again, click the button indicated in red



5. Select input file (here the file is #TotalSet290.txt), i.e., again click "Choose work file"



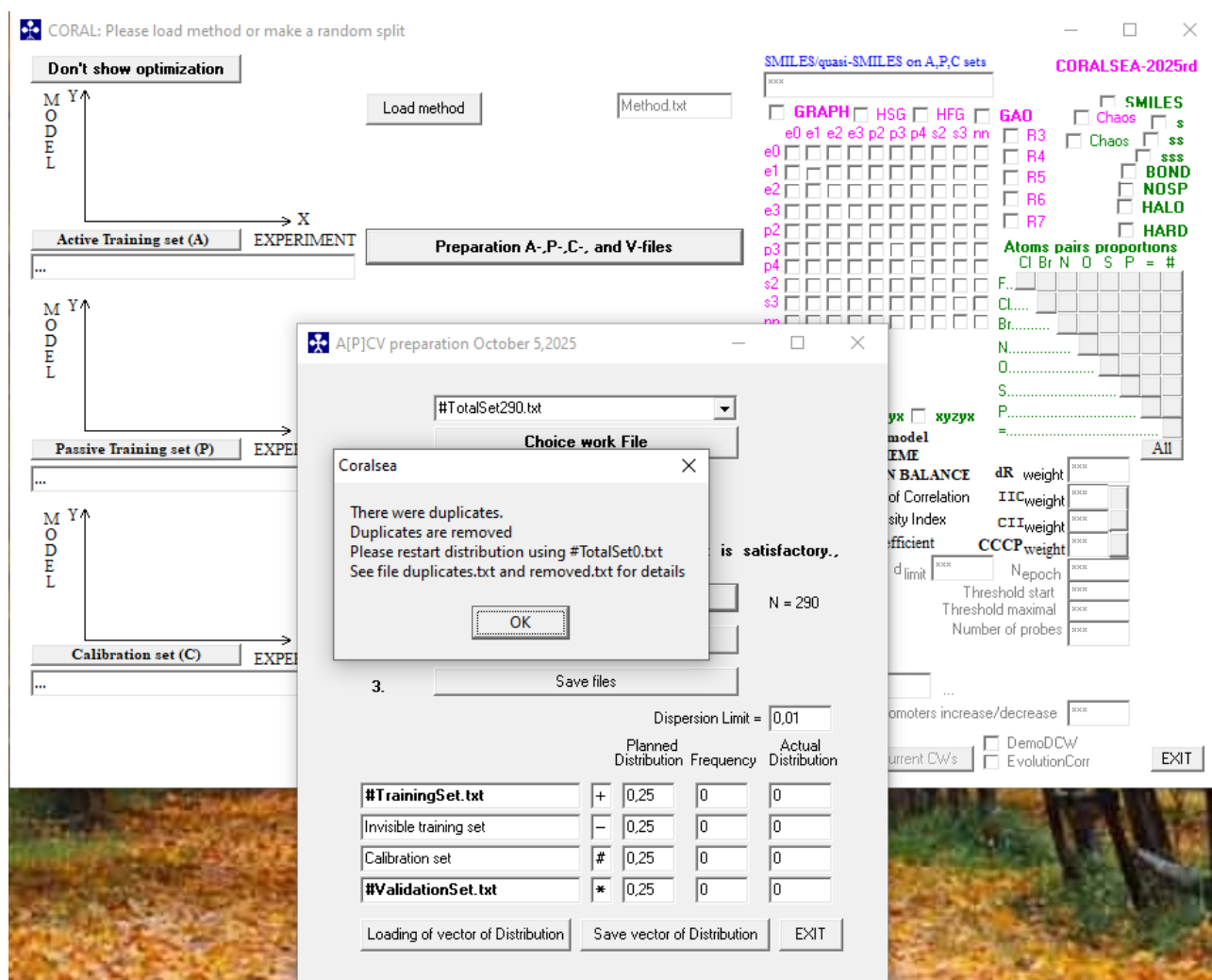
6. Click the button "Load"

The screenshot displays the CORAL software interface. At the top, a menu bar includes "CORAL: Please load method or make a random split". Below it, there are several panels:

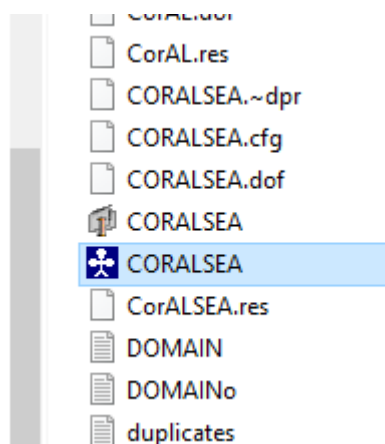
- Left Panel:** Contains three sections labeled "MODEL Y" vs "EXPERIMENT X":
  - "Don't show optimization"
  - "Active Training set (A)"
  - "Passive Training set (P)"
  - "Calibration set (C)"
- Top Center:** A "Load method" button and a "Method.txt" input field.
- Center:** A large window titled "A[P]CV preparation October 5, 2025". It contains:
  - A dropdown menu showing "#TotalSet290.txt".
  - A "Choice work File" button.
  - Instructions: "The first, Load selected file", "The second action should be 'Do distribution'.", and "The third action should be 'Save files'., if split is satisfactory..".
  - A numbered list (1-3) corresponding to the instructions, each with a button: "Load", "Do distribution", and "Save files".
  - A "Dispersion Limit = 0,01" field.
  - A table with columns: "Planned Distribution", "Frequency", and "Actual Distribution". The rows are:
 

	+	0,25	0	0
#TrainingSet.txt	+	0,25	0	0
Invisible training set	-	0,25	0	0
Calibration set	#	0,25	0	0
#ValidationSet.txt	*	0,25	0	0
  - Buttons at the bottom: "Loading of vector of Distribution", "Save vector of Distribution", and "EXIT".
- Right Panel:** Titled "SMILES/quasi-SMILES on A,P,C sets" and "CORALSEA-2025rd". It features a grid of checkboxes for various methods (GRAPH, HSG, HFG, GAO, R3, R4, R5, R6, R7) and a section for "Atoms pairs proportions" with a matrix for elements F, Cl, Br, N, O, S, P.

## 7. Click “OK”



## 8. Run CORALSEA.exe





**CORAL: Please load method or make a random split**

**Don't show optimization**

**MODEL** Y↑ X  
Active Training set (A) EXPERIMENT  
...  
**Preparation A-P-, C-, and V-files**

**MODEL** Y↑ X  
Passive Training set (P) EXPERIMENT  
...  
**Select Las Vegas algorithm parameters**

**MODEL** Y↑ X  
Calibration set (C) EXPERIMENT  
...  
W% N111 N110 N101 N100 Nall Defect Density  
Split Info [0] [0] [0] [0] [0] [0] [0] [0]

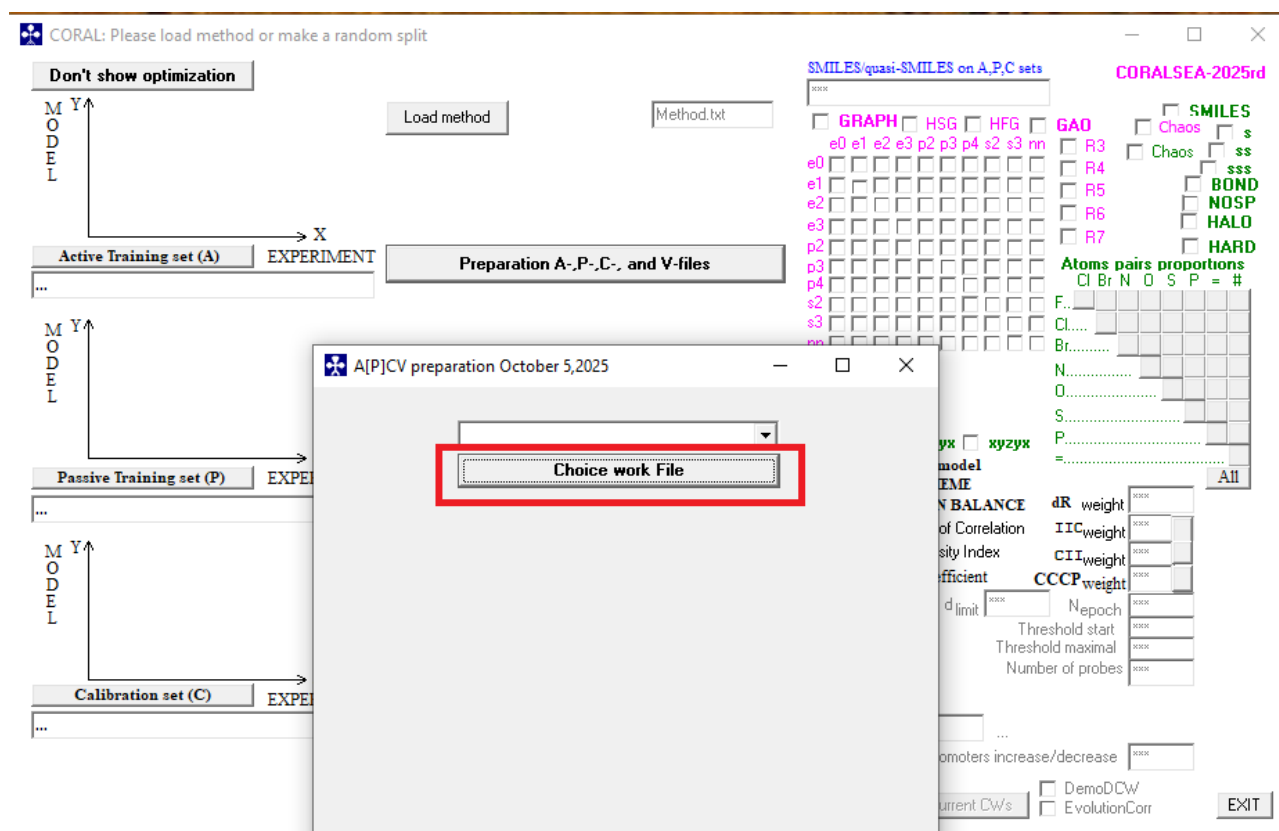
**SMILES/quasi-SMILES on A,P,C sets**  
xxxx  
☐ GRAPH ☐ HSG ☐ HFG ☐ GAO  
e0 ☐ e1 ☐ e2 ☐ p2 ☐ p3 ☐ p4 ☐ s2 ☐ s3 ☐ nn  
e1 ☐  
e2 ☐  
e3 ☐  
p2 ☐  
p3 ☐  
p4 ☐  
s2 ☐  
s3 ☐  
nn ☐

**M %**  
xxxx xxxx  
☐ xyx ☐ xyxx ☐ xyzyx  
☐ Classification model  
☒ CLASSIC SCHEME  
☒ CORRELATION BALANCE  
☐ Index of Ideality of Correlation  
☐ Correlation Intensity Index  
☐ Conformism coefficient  
D start xxxx d limit xxxx Epoch xxxx  
Threshold start xxxx Threshold maximal xxxx Number of probes xxxx

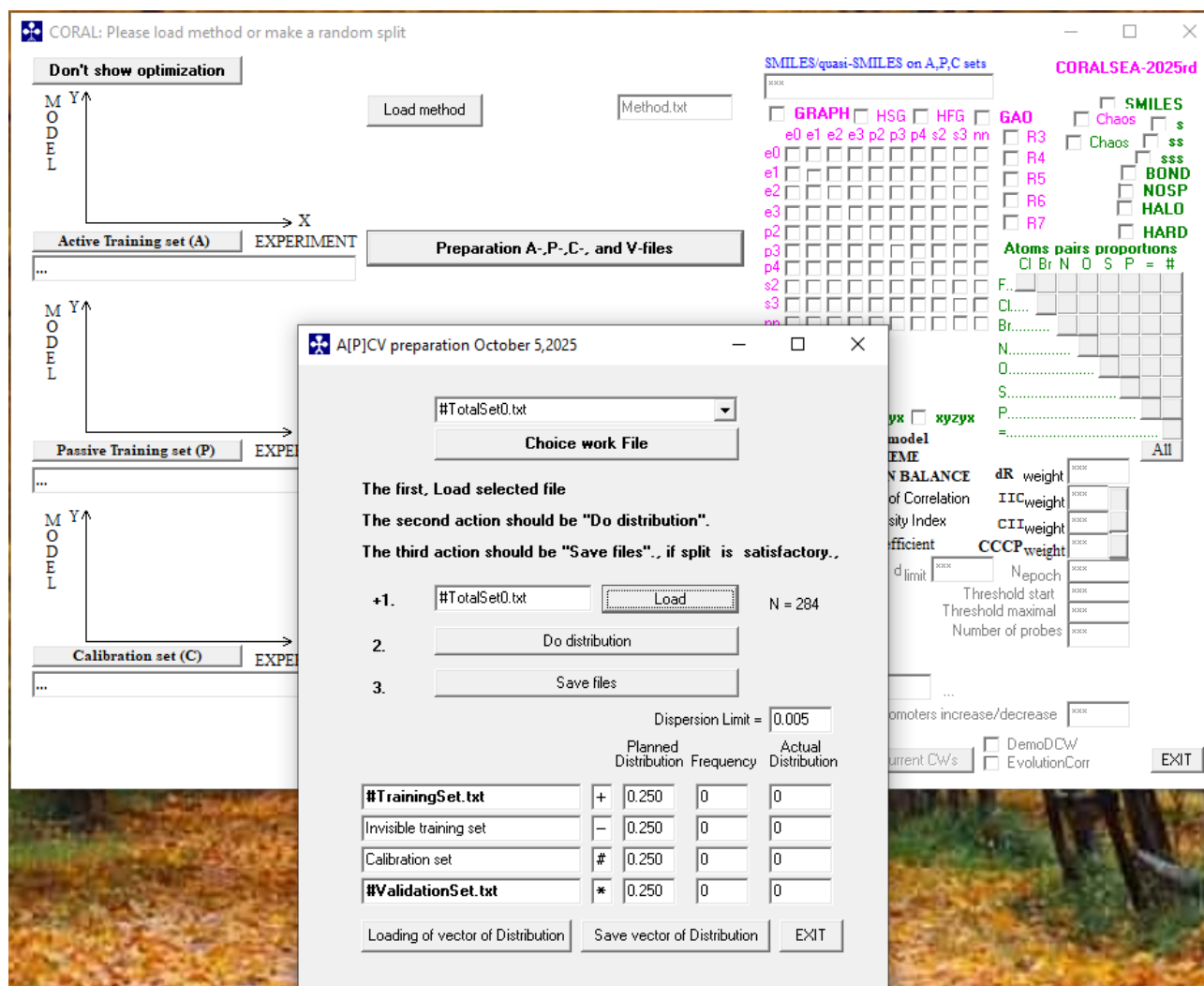
**CORALSEA-2025rd**  
**SMILES**  
☐ Chaos ☐ s  
☐ Chaos ☐ ss  
☐ BOND  
☐ NOSP  
☐ HALO  
☐ HARD  
**Atoms pairs proportions**  
Cl Br N O S P = #  
F. ....  
Cl....  
Br.....  
N.....  
O.....  
S.....  
P.....  
=..... All


**dR weight** xxxx  
**IIC weight** xxxx  
**CII weight** xxxx  
**CCCP weight** xxxx  
Model Details.txt ...  
Number of promoters increase/decrease xxxx  
☐ DemoDCW/  
☐ EvolutionCorr  
**STOP and SAVE current CW's** **EXIT**

10. Click the button in red



11. Select file "#TotalSet0.txt" (the file does not contain duplicates);
12. Click the button "Load";
13. Click the button "Save files";
14. Click the button "Exit"



 A[P]CV preparation October 5,2025

#TotalSet0.txt

Choice work File

The first, Load selected file

The second action should be "Do distribution".

The third action should be "Save files"., if split is satisfactory.,

+1.

#TotalSet0.txt

Load

N = 284

+2.

Do distribution

3.

Save files

Dispersion Limit =

0.005

		Planned Distribution	Frequency	Actual Distribution
#TrainingSet.txt	+	0.250	70	0.246
Invisible training set	-	0.250	71	0.250
Calibration set	#	0.250	72	0.254
#ValidationSet.txt	*	0.250	71	0.250

Loading of vector of Distribution

Save vector of Distribution

EXIT

#TotalSet0.txt

Choice work File

The first, Load selected file

The second action should be "Do distribution".

The third action should be "Save files"., if split is satisfactory.,

+1. #TotalSet0.txt Load N = 284

+2. Do distribution

+3. Save files

Dispersion Limit = 0.005

		Planned Distribution	Frequency	Actual Distribution
#TrainingSet.txt	+	0.250	70	0.246
Invisible training set	—	0.250	71	0.250
Calibration set	#	0.250	72	0.254
#ValidationSet.txt	*	0.250	71	0.250

Loading of vector of Distribution

Save vector of Distribution

EXIT

15. Click the button “Load method”

16. Click button “Select Las Vegas algorithm parameters”: M = number of probes; % is part of compounds for which random exchange of status, e.g., it was “A” becomes “P”; correspondingly, compounds in status “P” become “A”.

**CORAL: Please load method or make a random split**

Don't show optimization

Load method

Method.txt

Active Training set (A) EXPERIMENT

Preparation A-,P-,C-, and V-files

Passive Training set (P) EXPERIMENT

Select Las Vegas algorithm parameters

Calibration set (C) EXPERIMENT

W% N111 N110 N101 N100 Nall Defect Density

Split Info 0 0 0 0 0 0 0

SMILES/quasi-SMILES on A,P,C sets

CORALSEA-2025rd

GRAPH HSG HFG GAO

e0 e1 e2 e3 p2 p3 p4 s2 s3 nm

R3 R4 R5 R6 R7

SMILES

Chaos s

Chaos ss

BOND

NOSP

HALO

HARD

Atoms pairs proportions

Cl Br N O S P = #

F. 0 0 0 0 0 0 0

Cl. 0 0 0 0 0 0 0

Br. 0 0 0 0 0 0 0

N. 0 0 0 0 0 0 0

O. 0 0 0 0 0 0 0

S. 0 0 0 0 0 0 0

P. 0 0 0 0 0 0 0

= 0 0 0 0 0 0 0

All

Classification model

CLASSIC SCHEME

CORRELATION BALANCE

Index of Ideality of Correlation

Correlation Intensity Index

Conformism coefficient

dR weight

IIC weight

CII weight

CCC weight

D start 0.1 d limit 0.1 Nepoch 15

Threshold start 3

Threshold maximal 3

Number of probes 1

Model Details.txt

Number of promoters increase/decrease

STOP and SAVE current CW's

DemoDCW

EvolutionCorr

EXIT

**CORAL: select Phase (1 or 2); or change and save method; or import current model**

Show optimization DotSize-1

Load method

Save method

Method.txt

Phase 1: Search for preferable model (T\*,N\*)

Preparation A-,P-,C-, and V-files

Phase 2: Building up preferable model (T\*,N\*)

Select Las Vegas algorithm parameters

Calibration set (C) EXPERIMENT

W% N111 N110 N101 N100 Nall Defect Density

Split Info 0 0 0 0 0 0 0

SMILES/quasi-SMILES on A,P,C sets

#TrainingSet.txt

GRAPH HSG HFG GAO

e0 e1 e2 e3 p2 p3 p4 s2 s3 nm

R3 R4 R5 R6 R7

SMILES

Chaos s

Chaos ss

BOND

NOSP

HALO

HARD

Atoms pairs proportions

Cl Br N O S P = #

F. 0 0 0 0 0 0 0

Cl. 0 0 0 0 0 0 0

Br. 0 0 0 0 0 0 0

N. 0 0 0 0 0 0 0

O. 0 0 0 0 0 0 0

S. 0 0 0 0 0 0 0

P. 0 0 0 0 0 0 0

= 0 0 0 0 0 0 0

All

Classification model

CLASSIC SCHEME

CORRELATION BALANCE

Index of Ideality of Correlation

Correlation Intensity Index

Conformism coefficient

dR weight 0.1

IIC weight 0.3

CII weight 0

CCC weight

D start 0.1 d limit 0.1 Nepoch 15

Threshold start 3

Threshold maximal 3

Number of probes 1

Model Details.txt

Number of promoters increase/decrease 15

STOP and SAVE current CW's

DemoDCW

EvolutionCorr

EXIT

## 17. Click the button “Start Las Vegas”

CORAL: select Phase (1 or 2); or change and save method; or import current model

**Show optimization** DotSize=1

MODEL Y → X

Active Training set (A) EXPERIMENT

Preparation A-.P-.C-. and V-files

MODEL Y → X

Passive Training set (P) EXPERIMENT

MODEL Y → X

Calibration set (C) EXPERIMENT

Import of current model

W%	N111	N110	N101	N100	Nall	Defect	Density
Split Info	0	0	0	0	0	0	0

SMILES/quasi-SMILES on A,P,C sets

#TotalSet.txt

☐ GRAPH ☐ HSG ☐ HFG ☐ GAO ☐ SMILES ☒ Chaos ☒ s ☒ ss ☒ sss ☒ BOND ☒ NOSP ☒ HALO ☒ HARD

Atoms pairs proportions

	Cl	Br	N	O	S	P	=	#
F	0	0	0	0	0	0	0	0
Cl	0	0	0	0	0	0	0	0
Br	0	0	0	0	0	0	0	0
N	0	0	0	0	0	0	0	0
O	0	0	0	0	0	0	0	0
S	0	0	0	0	0	0	0	0
P	0	0	0	0	0	0	0	0
=	0	0	0	0	0	0	0	0

M %

10 0.1

☐ xyx ☐ xyxx ☐ xyzyx

☐ Classification model

☐ CLASSIC SCHEME

☒ CORRELATION BALANCE

☐ Index of Ideality of Correlation

☒ Correlation Intensity Index

☐ Conformism coefficient

d start 0.1 d limit 0.1 Nepoch 15

Threshold start 3

Threshold maximal 3

Number of probes 10

Model Details.txt

Number of promoters increase/decrease 15

☒ DemoDCW

☒ EvolutionCorr

**Start Las Vegas: define M and %**

The process is accompanied by information about the status of calculations

CORAL: Wait please...

**Show optimization** DotSize=1

MODEL Y → X

Active Training set (A) EXPERIMENT

Preparation A-.P-.C-. and V-files

MODEL Y → X

Passive Training set (P) EXPERIMENT

MODEL Y → X

Calibration set (C) EXPERIMENT

Import of current model

W%	N111	N110	N101	N100	Nall	Defect	Density
Split Info	100	40	0	0	0	40	61.14 0.396

SMILES/quasi-SMILES on A,P,C sets

#TotalSet.txt

☐ GRAPH ☐ HSG ☐ HFG ☐ GAO ☐ SMILES ☒ Chaos ☒ s ☒ ss ☒ sss ☒ BOND ☒ NOSP ☒ HALO ☒ HARD

Atoms pairs proportions

	Cl	Br	N	O	S	P	=	#
F	0	0	0	0	0	0	0	0
Cl	0	0	0	0	0	0	0	0
Br	0	0	0	0	0	0	0	0
N	0	0	0	0	0	0	0	0
O	0	0	0	0	0	0	0	0
S	0	0	0	0	0	0	0	0
P	0	0	0	0	0	0	0	0
=	0	0	0	0	0	0	0	0

M %

10 0.1

☐ xyx ☐ xyxx ☐ xyzyx

☐ Classification model

☐ CLASSIC SCHEME

☒ CORRELATION BALANCE

☐ Index of Ideality of Correlation

☒ Correlation Intensity Index

☐ Conformism coefficient

d start 0.1 d limit 0.1 Nepoch 15

Threshold start 3

Threshold maximal 3

Number of probes 10

Model Details.txt

Number of promoters increase/decrease 15

☒ DemoDCW

☒ EvolutionCorr

**Las Vegas Calculations in process**

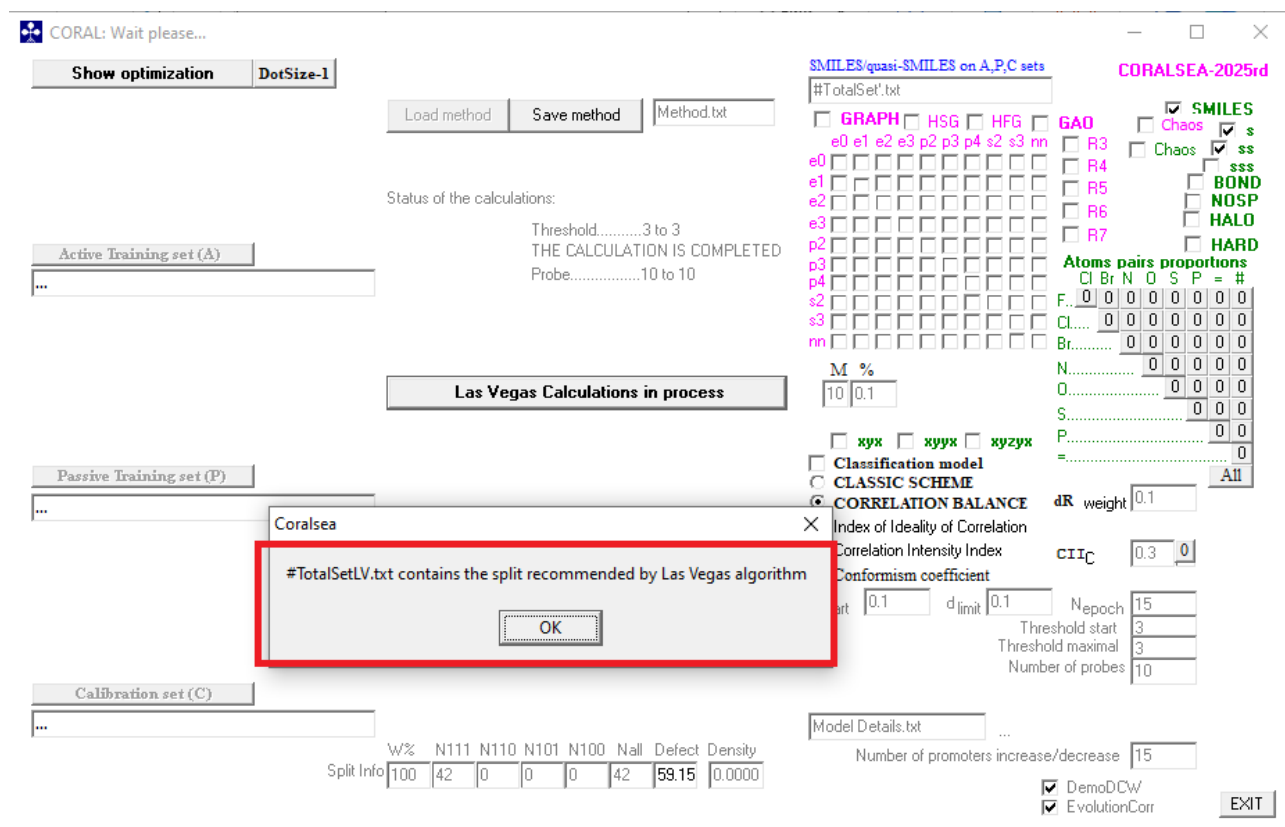
Status of the calculations:

Threshold.....3 to 3

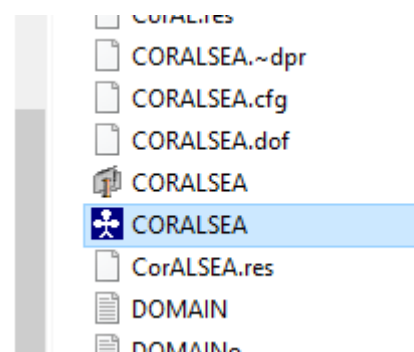
Epoch.....16 to 15

Probe.....1 to 10

18. When appears message “#TotalSetLV.txt contains the split recommended by the Las Vegas algorithm” appears, please click “OK”

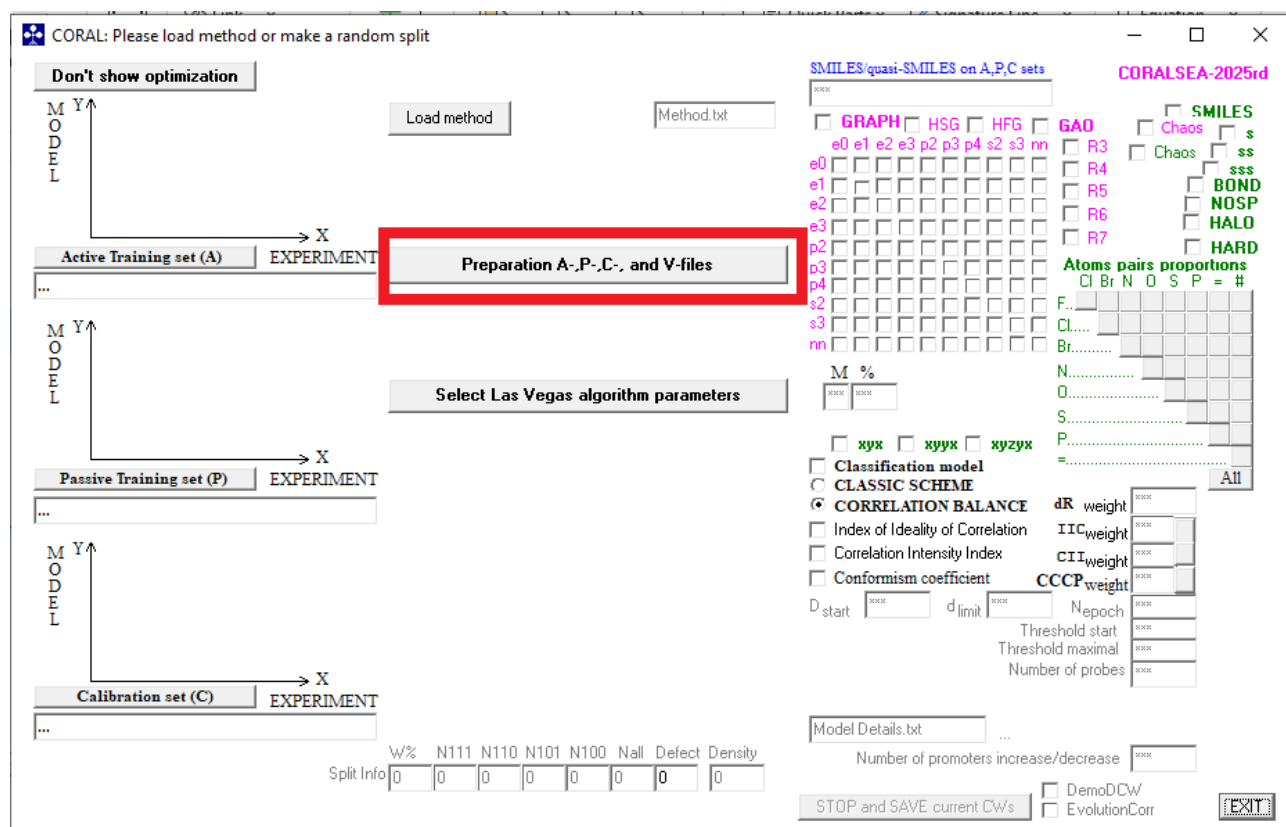


19. Run CORALSEA.txt





20. Click the button in the red frame



21. Click the button “THE BUILDING OF WORK FILES”



22. Select #TotalSetLV.txt
23. Load
24. Save files
25. Load method

A[P]CV preparation October 5,2025

#TotalSetLV.txt

Choice work File

The first, Load selected file

The second action should be "Do distribution".

The third action should be "Save files"., if split is satisfactory.,

+1. #TotalSetLV.txt Load N = 284

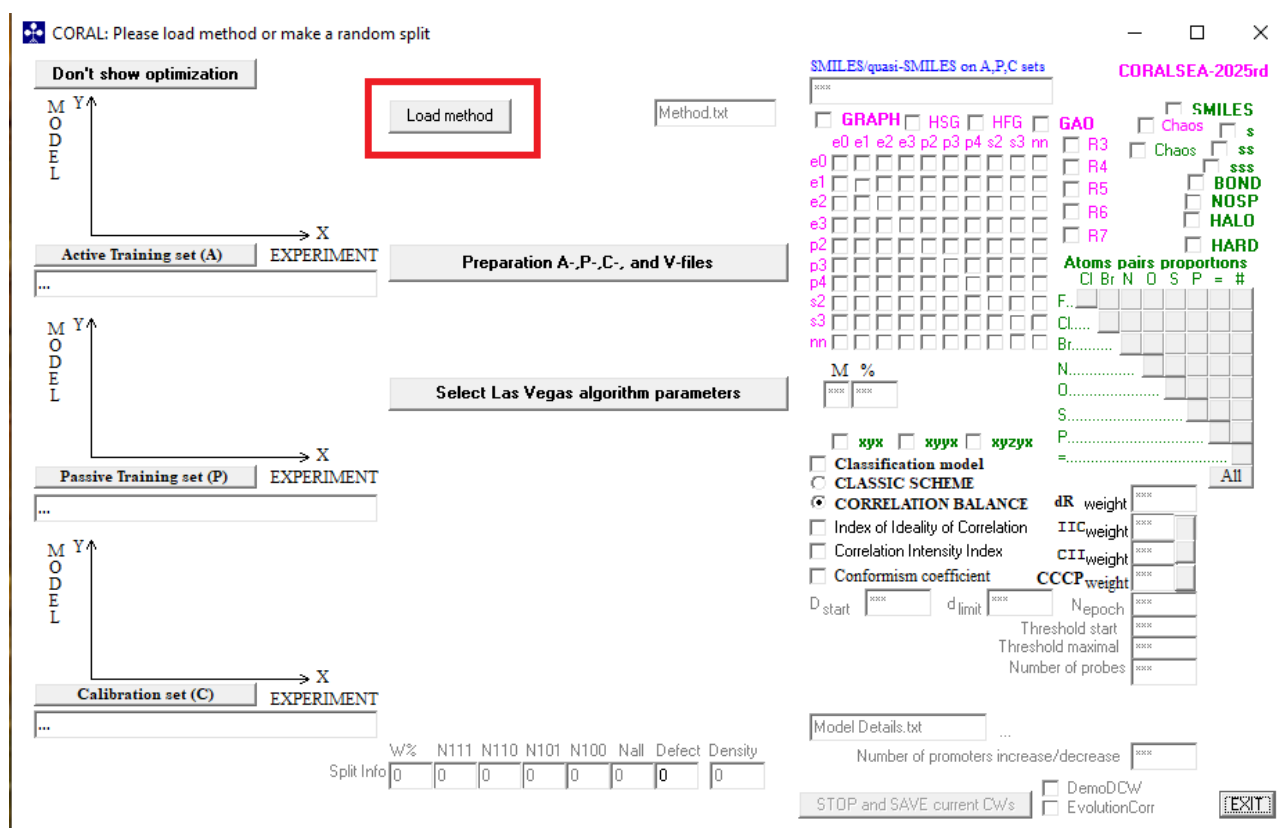
2.

+3. Save files

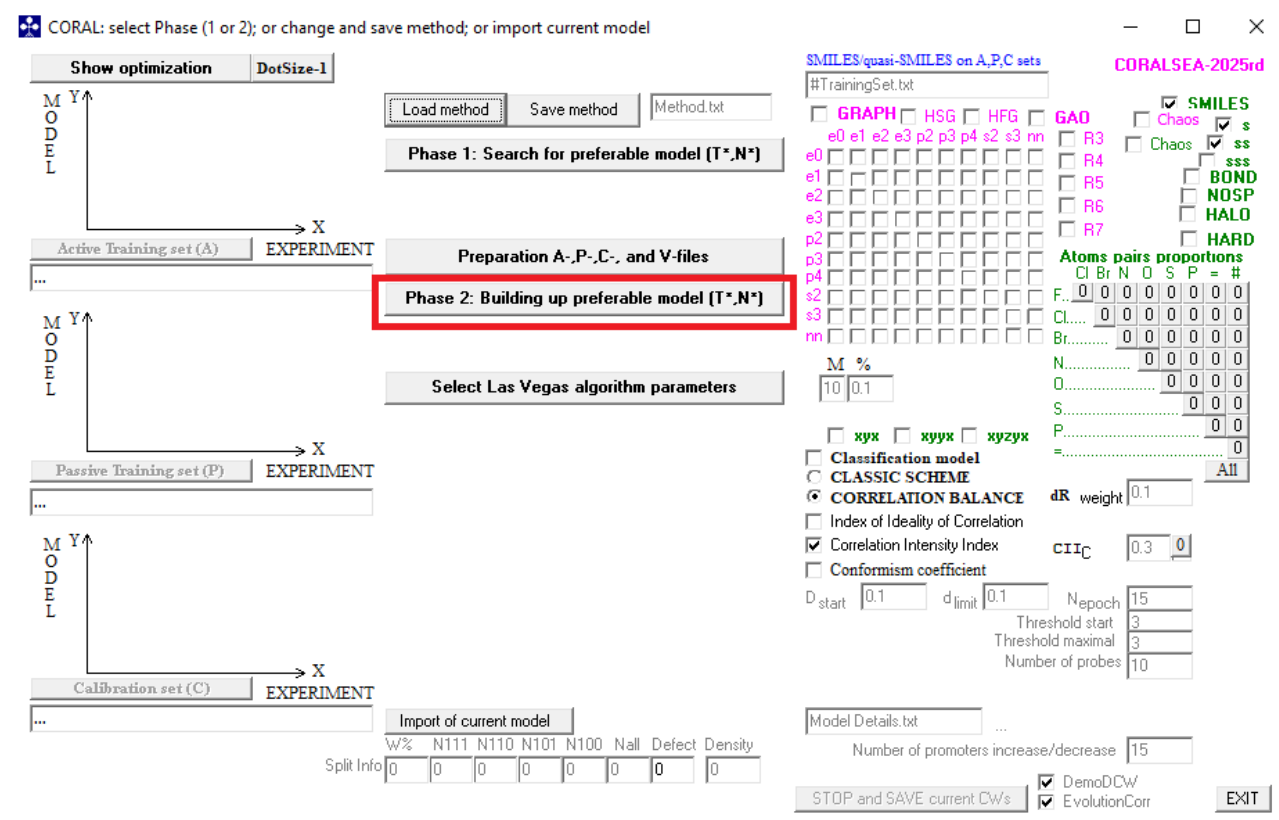
Dispersion Limit = 0.005

		Planned Distribution	Frequency	Actual Distribution
#TrainingSet.txt	+	0.250	70	0.246
Invisible training set	-	0.250	71	0.250
Calibration set	#	0.250	72	0.254
#ValidationSet.txt	*	0.250	71	0.250

Loading of vector of Distribution Save vector of Distribution EXIT



26. Click the button in the red frame



**CORAL: Building up preferable model**

**Show optimization** **DotSize=1**

**Active Training set (A)** **EXPERIMENT**

**Passive Training set (P)** **EXPERIMENT**

**Calibration set (C)** **EXPERIMENT**

**Phase 2: Building up preferable model (T\*, N\*)**

Define threshold and press Continue

**Continue**

C0 =  C1 =

**Import of current model**

W% N111 N110 N101 N100 Nall Defect Density

Split Info

**SMILES/quasi-SMILES on A,P,C sets**

#TrainingSet.txt

☐ **GRAPH** ☐ HSG ☐ HFG ☒ **GAO**

e0 ☐ R3 ☐ ☒ **Chaos** ☒ **SMILES**

e1 ☐ R4 ☐ ☒ **Chaos** ☒ **s**

e2 ☐ R5 ☐ ☒ **Chaos** ☒ **sss**

e3 ☐ R6 ☐ ☒ **Chaos** ☒ **BOND**

p2 ☐ R7 ☐ ☒ **Chaos** ☒ **NOSP**

p3 ☐ ☒ **Chaos** ☒ **HALO**

p4 ☐ ☒ **Chaos** ☒ **HARD**

s2 ☐ ☒ **Chaos** ☒ **HALO**

s3 ☐ ☒ **Chaos** ☒ **HALO**

nn ☐ ☒ **Chaos** ☒ **HALO**

**Atoms pairs proportions**

Cl Br N O S P = #

F..... 0 0 0 0 0 0 0

Cl..... 0 0 0 0 0 0 0

Br..... 0 0 0 0 0 0 0

N..... 0 0 0 0 0 0 0

O..... 0 0 0 0 0 0 0

S..... 0 0 0 0 0 0 0

P..... 0 0 0 0 0 0 0

=..... 0 0 0 0 0 0 0

**CLASSIC SCHEME**

☒ **CORRELATION BALANCE**

☐ Index of Ideality of Correlation

☒ Correlation Intensity Index

☐ Conformism coefficient

dR weight

CIIC

D\_start  d\_limit  Nepoch

Threshold start

Threshold maximal

Number of probes

**Model Details.txt**

Number of promoters increase/decrease

**STOP and SAVE current CW's** ☒ DemoDCW ☒ EvolutionCorr

**EXIT**

21

CORAL: you should save model now

Show optimization DotSize-1

Active Training set (A)

Passive Training set (P)

Calibration set (C)

Selected threshold is 3  
THE CALCULATION IS COMPLETED

Phase 2: Building up preferable model (T\*,N\*)

Define threshold and press Continue 3 Continue

C0 = 0.8207640 C1 = 0.3998357

Insert a SMILES for calculation of DCW' and EndPoint

Demo of calculation of DCW' and endpoint will be in file

Start of DCW' and Endpoint calculation for SMILES from file

Save Model

W% N111 N110 N101 N100 Nail Detect Density

Split Info 100 40 0 0 0 40 57.14 0.412

SMILES/quasi-SMILES on A,P,C sets

#TrainingSet.txt

GRAPH HSG HFG GAO

e0 e1 e2 e3 p2 p3 p4 s2 s3 nn

R3 R4 R5 R6 R7

SMILES

Chaos s

Chaos ss

BOND

NO SP

HALO

HARD

Atoms pairs proportions

Cl Br N O S P = #

F... 0 0 0 0 0 0 0 0

Cl... 0 0 0 0 0 0 0 0

Br... 0 0 0 0 0 0 0 0

N... 0 0 0 0 0 0 0 0

O... 0 0 0 0 0 0 0 0

S... 0 0 0 0 0 0 0 0

P... 0 0 0 0 0 0 0 0

=... 0 0 0 0 0 0 0 0

All

xyx xyxy xzyx

Classification model

CLASSIC SCHEME

CORRELATION BALANCE

Index of Ideality of Correlation

Correlation Intensity Index

Conformism coefficient

dR weight 0.1

CIIC 0.3 0

D\_start 0.1 d\_limit 0.1 Nepoch 15

DemoDCW'.txt

Threshold start 3

Threshold maximal 3

Number of probes 1

#ValidationSet.txt

#ModelForValidationSet.txt

Model Details.txt

Number of promoters increase/decrease 15

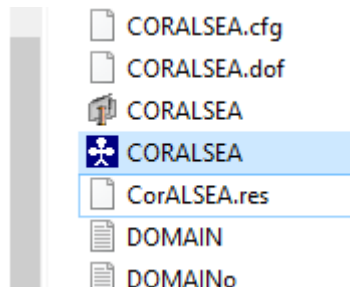
STOP and SAVE current CW's

DemoDCW

EvolutionCorr

EXIT

## 29. Run CORALSEA,exe



## 30. Click "Load method"

**CORAL: select Phase (1 or 2); or change and save method; or import current model**

**Show optimization** **DotSize=1**

**MODEL** **Y** **X**

**Active Training set (A)** **EXPERIMENT**

**MODEL** **Y** **X**

**Passive Training set (P)** **EXPERIMENT**

**MODEL** **Y** **X**

**Calibration set (C)** **EXPERIMENT**

**Load method** **Save method** **Method.txt**

**Phase 1: Search for preferable model (T\*.N\*)**

**Preparation A-,P-,C-, and V-files**

**Phase 2: Building up preferable model (T\*.N\*)**

**Select Las Vegas algorithm parameters**

**Import of current model**

**W%** **N111** **N110** **N101** **N100** **Nall** **Defect** **Density**

**Split Info** **0** **0** **0** **0** **0** **0** **0**

**SMILES/quasi-SMILES on A,P,C sets**

**#TrainingSet.txt**

**GRAPH** **HSG** **HFG** **GAD** **Chaos** **ss** **BOND** **NOSP** **HALO** **HARD**

**Atoms pairs proportions**

**Cl** **Br** **N** **O** **S** **P** **=** **#**

**0** **0** **0** **0** **0** **0** **0** **0**

**Cl** **0** **0** **0** **0** **0** **0** **0**

**Br** **0** **0** **0** **0** **0** **0** **0**

**N** **0** **0** **0** **0** **0** **0** **0**

**O** **0** **0** **0** **0** **0** **0** **0**

**S** **0** **0** **0** **0** **0** **0** **0**

**P** **0** **0** **0** **0** **0** **0** **0**

**=** **0** **0** **0** **0** **0** **0** **0**

**All**

**Classification model**

**CLASSIC SCHEME**

**CORRELATION BALANCE**

**Index of Ideality of Correlation**

**Correlation Intensity Index**

**Conformism coefficient**

**d<sub>start</sub>** **0.1** **d<sub>limit</sub>** **0.1** **Nepoch** **15**

**Threshold start** **3**

**Threshold maximal** **3**

**Number of probes** **1**

**Model Details.txt** **...**

**Number of promoters increase/decrease** **15**

**STOP and SAVE current CWs** **DemoDCW** **EvolutionCorr** **EXIT**

31. Click “Import current model”

**CORAL: select Phase (1 or 2); or change and save method; or import current model**

**Show optimization** DotSize=1

**MODEL** Y↑

Active Training set (A) → X EXPERIMENT

Passive Training set (P) → X EXPERIMENT

Calibration set (C) → X EXPERIMENT

**Phase 1: Search for preferable model (T\*,N\*)**

**Phase 2: Building up preferable model (T\*,N\*)**

**Select Las Vegas algorithm parameters**

**Import of current model**

W% NT11 NT10 NT01 NT00 Nall Defect Density

Split Info 0 0 0 0 0 0 0 0

**SMILES/quasi-SMILES on A,P,C sets**

#TrainingSet.txt

☐ GRAPH ☐ HSG ☐ HFG ☐ GAO

e0 e1 e2 e3 p2 p3 p4 s2 s3 nn

☐ R3 ☐ R4 ☐ R5 ☐ R6 ☐ R7

**CORALSEA-2025rd**

☒ SMILES ☐ Chaos ☒ s ☒ ss

☐ BOND ☐ NOSP ☐ HALO ☐ HARD

**Atoms pairs proportions**

	Cl	Br	N	O	S	P	=	#
F	0	0	0	0	0	0	0	0
Cl	0	0	0	0	0	0	0	0
Br	0	0	0	0	0	0	0	0
N	0	0	0	0	0	0	0	0
O	0	0	0	0	0	0	0	0
S	0	0	0	0	0	0	0	0
P	0	0	0	0	0	0	0	0
=	0	0	0	0	0	0	0	0
#	0	0	0	0	0	0	0	0

M % 10 0.1

☐ xyx ☐ xyxx ☐ xyzyx

☐ Classification model

☐ CLASSIC SCHEME

☒ CORRELATION BALANCE

☐ Index of Ideality of Correlation

☒ Correlation Intensity Index

☐ Conformism coefficient

D\_start 0.1 d\_limit 0.1 Nepoch 15

Threshold start 3

Threshold maximal 3

Number of probes 1

Model Details.txt

Number of promoters increase/decrease 15

☐ DemoDCW/ ☒ EvolutionCorr

**STOP and SAVE current CWs** **EXIT**

32. In order to consider the model for the validation set, click the button in the red frame



**CORAL: Calculation of model for external substances**

Hide DotSize-1

Phase 1: Search for preferable model [T\*,N\*]

Preparation A-,P-,C-, and V-files

Phase 2: Building up preferable model [T\*,N\*]

Calculation model for a list of SMILES

list.txt ListModel.txt

C0 = 0.8207640 C1 = 0.3998357

Insert a SMILES for calculation of DCW and EndPoint

Demo of calculation of DCW and endpoint will be in file

DCW(3,15)= EndPoint =

Start of DCW and Endpoint calculation for SMILES from file

Import of current model

W% N111 N110 N101 N100 Nall Defect Density

Split Info 100 40 0 0 0 40 57.14 0.412

Continue optimization

STOP and SAVE current CW's

Model Details.txt

Number of promoters increase/decrease 15

SMILES/quasi-SMILES on A,P,C sets

#TrainingSet.txt

GRAPH HSG HFG GAO

e0 e1 e2 e3 p2 p3 p4 s2 s3 nn

R3 R4 R5 R6 R7

CORALSEA-2025rd

SMILES

Chaos s

Chaos ss

BOND

NOSP

HALO

HARD

Atoms pairs proportions

Cl Br N O S P = #

F... 0 0 0 0 0 0 0

Cl... 0 0 0 0 0 0 0

Br... 0 0 0 0 0 0 0

N... 0 0 0 0 0 0 0

O... 0 0 0 0 0 0 0

S... 0 0 0 0 0 0 0

P... 0 0 0 0 0 0 0

= 0

All

Classification model

CLASSIC SCHEME

CORRELATION BALANCE

Index of Ideality of Correlation

Correlation Intensity Index

Conformism coefficient

dR weight 0.1

CIIC 0.3 0

D start 0.1 d limit 0.1 Nepoch 15

DemoDCW.txt Threshold start 3

Threshold maximal 3

Number of probes 1

#ValidationSet.txt #ModelForValidationSet.txt

Model Details.txt

Number of promoters increase/decrease 15

DemoDCW

EvolutionCorr

EXIT

**CORAL: Calculation of model for external substances**

Hide DotSize-1

Phase 1: Search for preferable model [T\*,N\*]

Preparation A-,P-,C-, and V-files

Phase 2: Building up preferable model [T\*,N\*]

Calculation model for a list of SMILES

list.txt ListModel.txt

C0 = 0.8207640 C1 = 0.3998357

Insert a SMILES for calculation of DCW and EndPoint

Demo of calculation of DCW and endpoint will be in file

DCW(3,15)= EndPoint =

Start of DCW and Endpoint calculation for SMILES from file

Import of current model

W% N111 N110 N101 N100 Nall Defect Density

Split Info 100 40 0 0 0 40 57.14 0.412

Continue optimization

STOP and SAVE current CW's

Model Details.txt

Number of promoters increase/decrease 15

SMILES/quasi-SMILES on A,P,C sets

#TrainingSet.txt

GRAPH HSG HFG GAO

e0 e1 e2 e3 p2 p3 p4 s2 s3 nn

R3 R4 R5 R6 R7

CORALSEA-2025rd

SMILES

Chaos s

Chaos ss

BOND

NOSP

HALO

HARD

Atoms pairs proportions

Cl Br N O S P = #

F... 0 0 0 0 0 0 0

Cl... 0 0 0 0 0 0 0

Br... 0 0 0 0 0 0 0

N... 0 0 0 0 0 0 0

O... 0 0 0 0 0 0 0

S... 0 0 0 0 0 0 0

P... 0 0 0 0 0 0 0

= 0

All

Classification model

CLASSIC SCHEME

CORRELATION BALANCE

Index of Ideality of Correlation

Correlation Intensity Index

Conformism coefficient

dR weight 0.1

CIIC 0.3 0

D start 0.1 d limit 0.1 Nepoch 15

DemoDCW.txt Threshold start 3

Threshold maximal 3

Number of probes 1

#ValidationSet.txt #ModelForValidationSet.txt

Model Details.txt

Number of promoters increase/decrease 15

DemoDCW

EvolutionCorr

EXIT

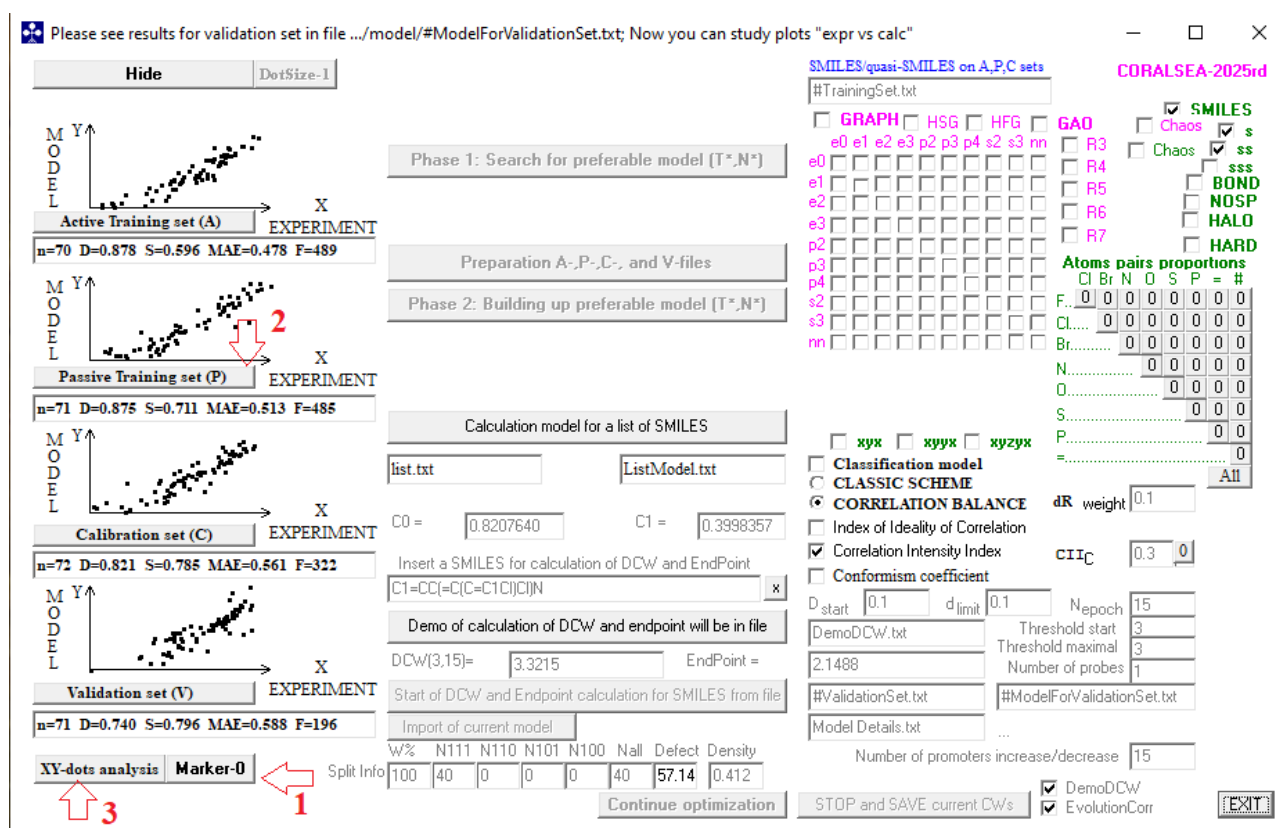
DCW/EndPoint calculation for data from #ValidationSet.txt is completed.

The numerical data is saved in model/#ModelForValidationSet.txt.

Next picture will be graphical representation of the model for Active training set; Passive training set; Calibration set; and Validation set.

File "DOMAIN.txt" contains the Domain of applicability, i.e. data without outliers which are defined according to the statistical defects calculated for SMILES or quasi-SMILES applied to build up the given model; file "DOMAIN0.txt" contains outliers

OK



Click "1": Marker-0 → Marker-1 → Marker-2 → Marker - 0 ... ;

Click "2" selection of a set active training set or passive training set, or calibration set;

Click "3" shows for selected set clustering red/green for calculation IIC (marker-1) or clustering for CII (marker-2)

Similarly, one can consider other details of the resulting model (see below).

Please see results for validation set in file .../model/#ModelForValidationSet.txt; Now you can study plots "expr vs calc"

Hide DotSize=1

MODEL Y

Active Training set (A) X EXPERIMENT

n=70 D=0.878 S=0.596 MAE=0.478 F=489

MODEL Y

Passive Training set (P) X EXPERIMENT

n=71 D=0.875 S=0.711 MAE=0.513 F=485

MODEL Y

Calibration set (C) X EXPERIMENT

n=72 D=0.821 S=0.785 MAE=0.561 F=322

MODEL Y

Validation set (V) X EXPERIMENT

n=71 D=0.740 S=0.796 MAE=0.588 F=196

XY-dots analysis Marker-1

Phase 1: Search for preferable model (T\*,N\*)

Preparation A-,P-,C-, and V-files

Phase 2: Building up preferable model (T\*,N\*)

Calculation model for a list of SMILES

list.txt ListModel.txt

Coralsea

The image of two correlation clusters

OK

Start of DCW and Endpoint calculation for SMILES from file

Import of current model

W% N111 N110 N101 N100 Nall Defect Density

Split Info 100 40 0 0 0 40 57.14 0.412

Continue optimization

SMILES/quasi-SMILES on A,P,C sets

#TrainingSet.txt

☐ GRAPH ☐ HSG ☐ HFG ☐ GAO ☐ Chaos ☒ SMILES

e0 e1 e2 e3 p2 p3 p4 s2 s3 nm ☐ R3 ☐ R4 ☐ R5 ☐ R6 ☐ R7

Atoms pairs proportions

	Cl	Br	N	O	S	P	#
F	0	0	0	0	0	0	0
Cl	0	0	0	0	0	0	0
Br	0	0	0	0	0	0	0
N	0	0	0	0	0	0	0
O	0	0	0	0	0	0	0
S	0	0	0	0	0	0	0
P	0	0	0	0	0	0	0
=	0	0	0	0	0	0	0

☐ xyx ☐ xyxx ☐ xyzyx

☐ Classification model

☐ CLASSIC SCHEME

☒ CORRELATION BALANCE

☐ Index of Ideality of Correlation

☒ Correlation Intensity Index

☐ Conformism coefficient

D start 0.1 d limit 0.1 Nepoch 15

DemoDCW.txt Threshold start 3

2.1488 Threshold maximal 3

#ValidationSet.txt Number of probes 1

#ModelForValidationSet.txt

Model Details.txt

Number of promoters increase/decrease 15

☒ DemoDCW

☒ EvolutionCorr

EXIT

