

**NOVASIGN PRE-RELEASE PROTOTYPE OF THE
HYBRID MODELING TOOLBOX
v2.3.0.8 (FREE VERSION)
SHORT USER MANUAL**



Last edited: September 11th 2019

In case of errors or questions, contact us via the Novasign portal

<https://portal.novasign.eu/>

Before you start:

This is a pre-release version of a powerful hybrid-modeling toolbox. This program only works on Windows. We do not support macOS at the moment. However, with windows emulators, you might overcome this issue.

The current functionalities of the toolbox aim to describe cell cultivation processes as fast as possible. Other unit operations are currently under development, but might already be modeled with this pre-release version.

This free pre-release version of the toolbox requires an internet connection and the registration of the user. The internet connection is required to periodically change the content of the advertisement. If you do not register, you cannot activate the toolbox.

If you require a version which does not need an internet connection or you have any questions please contact us at any time via <https://portal.novesign.eu/>

We also provide you data sets as well as a whole test-project which you can use to understand the toolbox. Download it for free from the *Documentation* section (<https://portal.novesign.eu/>).

Contents

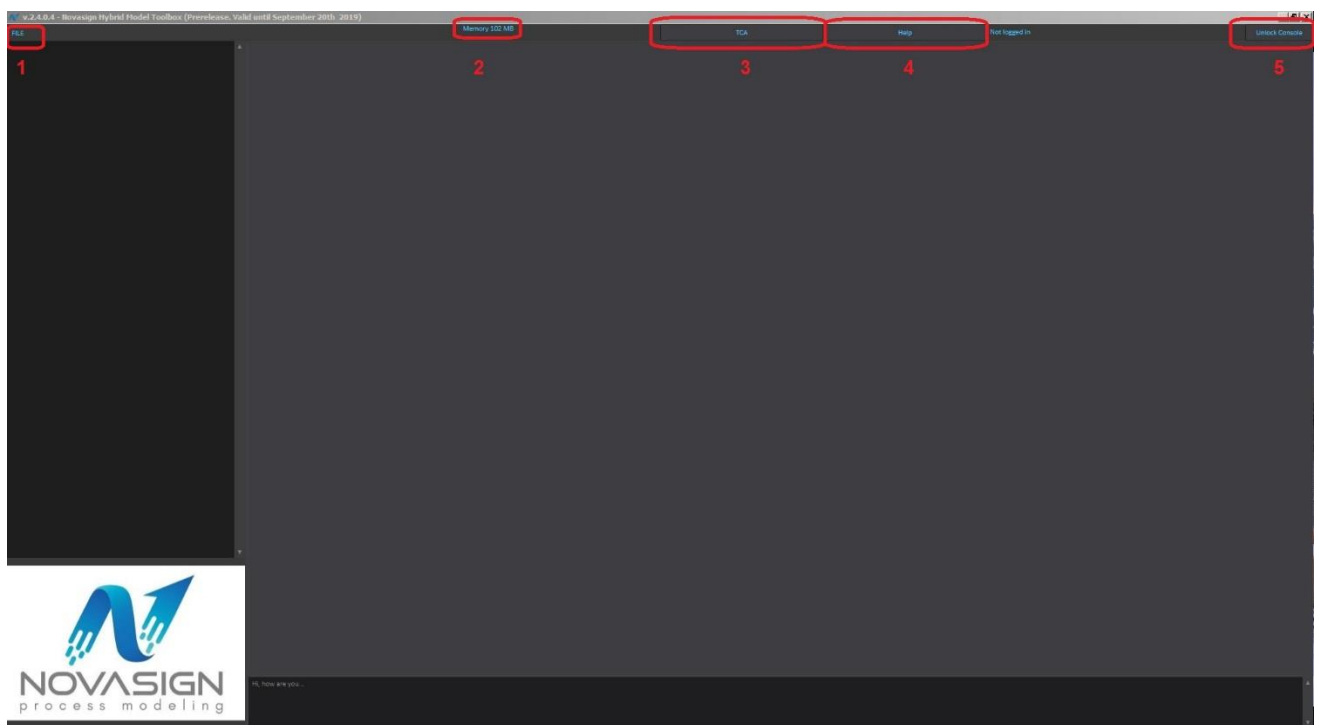
I. Download the Program	4
II. Starting the Program.....	4
III. Prepare the Data	7
IV. Create a new Project and add Data	9
V. Inspect and plot your Data	12
VI. Building a Hybrid Model	16
VII. Evaluate the Results.....	32

I. Download the Program

Download latest version of the Toolbox from the *Downloads* section (<https://portal.novasign.eu/>).

II. Starting the Program

- extract the folder from the downloaded .zip file
- enter the folder and open „IdentificationTool.exe“
- **NOTE:** This is an exportable desktop version, which you do not have to install. You agreed upon the terms and conditions upon registration. However, you can always get back to them using either the homepage or the documentation button inside the toolbox.
- open the Novasign Hybrid Model Toolbox to get into the home screen
- Initially, five objects are shown. The ‘FILE’ panel (1), the used PC memory (2), the terms and conditions (3), the help (4) and the ‘Unlock Console’ button (5). First, to work with the toolbox you have to login by clicking at the ‘Unlock Console’ button.

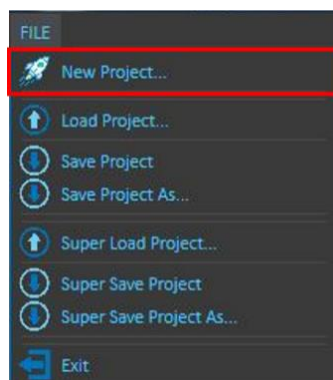


- at this panel, you have to login by entering your ‘User Name’ and the chosen ‘Password’

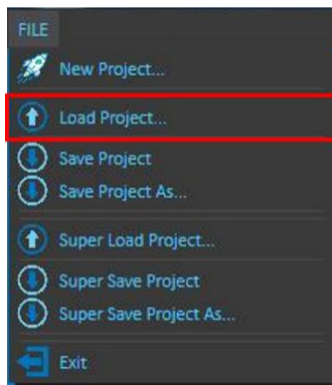
- if you don't have a user name, you have to sign up by clicking at the hyperlink in the marked panel. You will be forwarded to our registration form
- once you registered and unlocked the console with your user name and password, you can start using the toolbox
- you can lock the console by clicking on the button again and use 'Lock'
- by doing so, no changes can be done in the toolbox until you again unlock the console as before



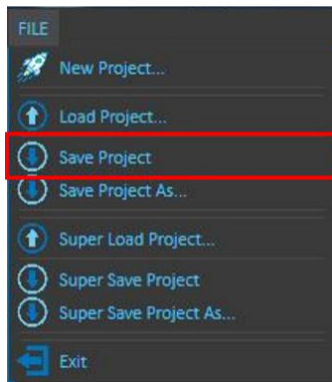
- choose the desired operation in the tab 'FILE'
- creates a new project



- loads an existing and saved project



- saves or super saves the current project. The difference is that when you use super save also all the graphics you stored are saved. This might require some time if you already created a project and built several hybrid models. The same applies to the load project buttons too. The super load might take longer because more data has to be loaded.



III. Prepare the Data

To ensure the proper functionality of the hybrid modeling toolbox, the data you want to import has to be structured properly. These detailed instructions are listed below. You can use the provided data set as the reference template and to better understand the required data structure.

Valid data formats to import: .csv, .xls, .dat, .mat

- arrange your data that each variable is represented by one column
- do not use letters in rows, apart from the first row; this will cause importing and calculation errors
- likewise, attach new runs after the last row of the last run
- put the time variable in the first column for each data set

time resolution: use an adequate time resolution for your online variables. If you model a microbial process do not use fewer than 3 minutes. Your model calculation will just take much longer without any benefits. For mammalian processes use 15 to 30 minutes

- assign the Run ID to the last column
- name the variables without special characters, e.g., square brackets and spaces. Spaces and square brackets will still allow you to import and plot your data but will crush the program when used as modeling inputs

Good variable name: biomass(g/L)

poor variable name biomass (g/L) or biomass [g/L]

- use bioprocess units for your parameters to be modeled. Otherwise, the mass balances are wrong, e.g., sampling and reactor volume (volume in L; masses in grams; concentrations in g/L)
- for interpolation, ensure all data series contain more than a certain number of values
 - for linear: min. 2 values are required
 - for cubic robust: min. 3 values are required
 - for Hermite: min. 5 are required
- define an invalid number for all empty fields in the data, e.g., the software standard is '-9.999'
- make sure to have a value (other than -9.999) for each first line of each variable you want to model! otherwise, the system will not work!
- split the data in a train/validation and a test set

you have to import at least two files; one for training and one for testing

you can also upload a file for each run and assemble the training runs together afterward

- example: typical data structure before import

feedtime(h)	CDw(g/L)	Produktsoll(g/L)	Temp(°C)	Feed(L)	InductorMASS(kg)	Inductor(yesno)	Base(L)	Reaktorvolumen(L)	SamplingVolume(L)	RunID
0	5.625	0	36.3	0	0	0	0.081	4.07332		42
0.05			36.3	0	0	0	0.081	4.07332		
0.1			36.2	0.002	0	0	0.081	4.07517		
0.15			36.4	0.005	0	0	0.081	4.07795		
0.2			36.4	0.008	0	0	0.081	4.08165		
0.25			36.8	0.011	0	0	0.081	4.08443		
27.8			37.1	7.147	0.00007818	1	0.638	11.95614		
27.85			37.1	7.185	0.00007856	1	0.642	11.93871		
27.9			37.1	7.223	0.00007896	1	0.648	12.04367		
27.95			37	7.261	0.00007936	1	0.65	12.08545		
28	17.7	1.724	37	7.3	0.00007974	1	0.655	12.12993	0.02	
0	5.625	0	36.8	0	0	0	0.04	4.03672		44
0.05			36.7	0	0	0	0.041	4.0373		
0.1			36.7	0	0	0	0.042	4.03839		

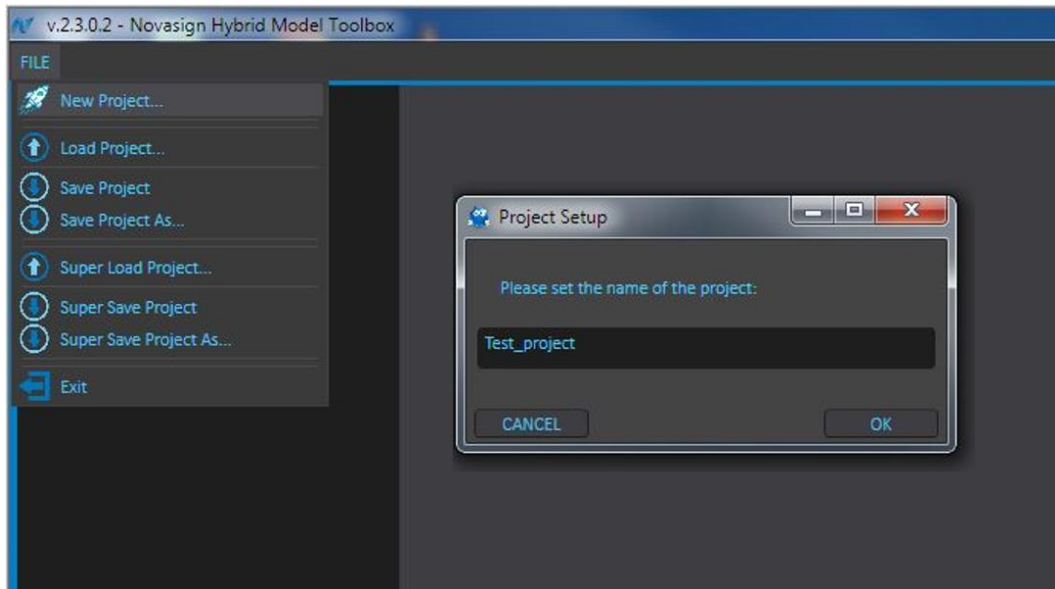
- example: correctly formatted data for modeling

feedtime(h)	CDw(g/L)	Produktsoll(g/L)	Temp(°C)	Feed(L)	InductorMASS(kg)	Inductor(yesno)	Base(L)	Reaktorvolumen(L)	SamplingVolume(L)	RunID
0	5.625	0	36.3	0	0	0	0.081	4.07332	-9.999	42
0.05	-9.999	-9.999	36.3	0	0	0	0.081	4.07332	-9.999	-9.999
0.1	-9.999	-9.999	36.2	0.002	0	0	0.081	4.07517	-9.999	-9.999
0.15	-9.999	-9.999	36.4	0.005	0	0	0.081	4.07795	-9.999	-9.999
0.2	-9.999	-9.999	36.4	0.008	0	0	0.081	4.08165	-9.999	-9.999
0.25	-9.999	-9.999	36.8	0.011	0	0	0.081	4.08443	-9.999	-9.999
27.8	-9.999	-9.999	37.1	7.147	0.00007818	1	0.638	11.95614	-9.999	-9.999
27.85	-9.999	-9.999	37.1	7.185	0.00007856	1	0.642	11.93871	-9.999	-9.999
27.9	-9.999	-9.999	37.1	7.223	0.00007896	1	0.648	12.04367	-9.999	-9.999
27.95	-9.999	-9.999	37	7.261	0.00007936	1	0.65	12.08545	-9.999	-9.999
28	17.7	1.724	37	7.3	0.00007974	1	0.655	12.12993	0.02	-9.999
0	5.625	0	36.8	0	0	0	0.04	4.03672	-9.999	44
0.05	-9.999	-9.999	36.7	0	0	0	0.041	4.0373	-9.999	-9.999
0.1	-9.999	-9.999	36.7	0	0	0	0.042	4.03839	-9.999	-9.999

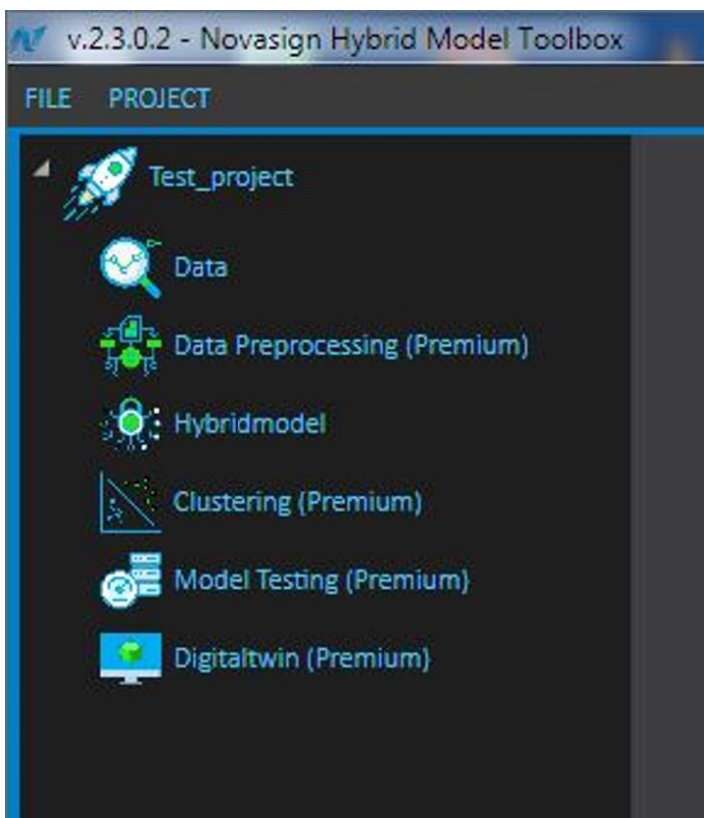
IV. Create a new Project and add Data

To start modeling, you first have to create a project.

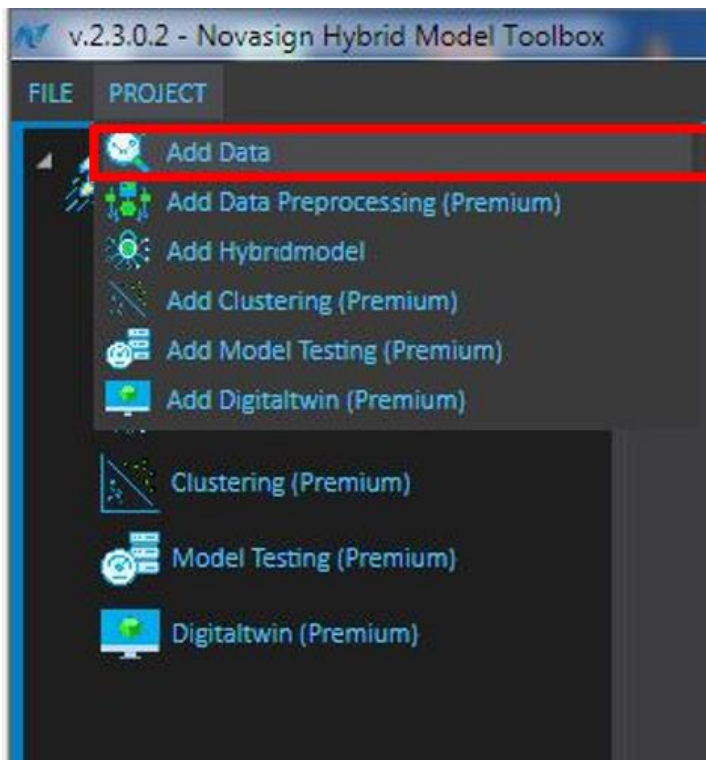
- go to the 'FILE' tab and click on 'New Project'
- name your project



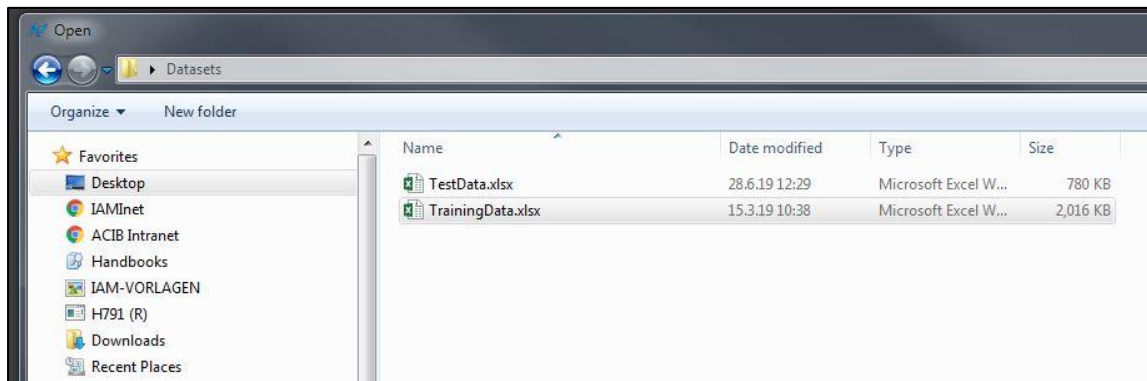
- the tab 'PROJECT' will appear and within the created project and all subpanels



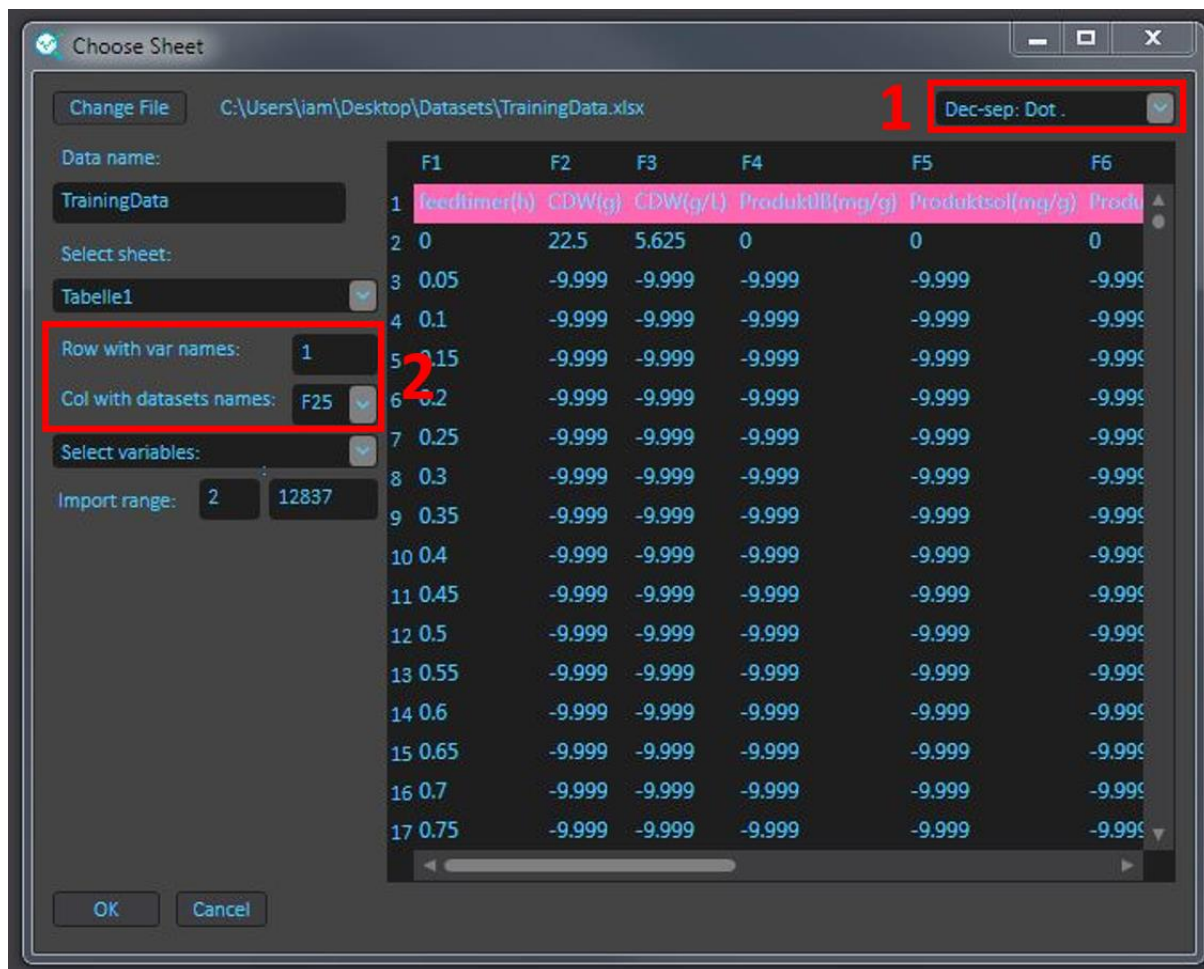
- in the next step, add data to the created project



- select the data you want to import; take care that the data is prepared and structured as suggested previously

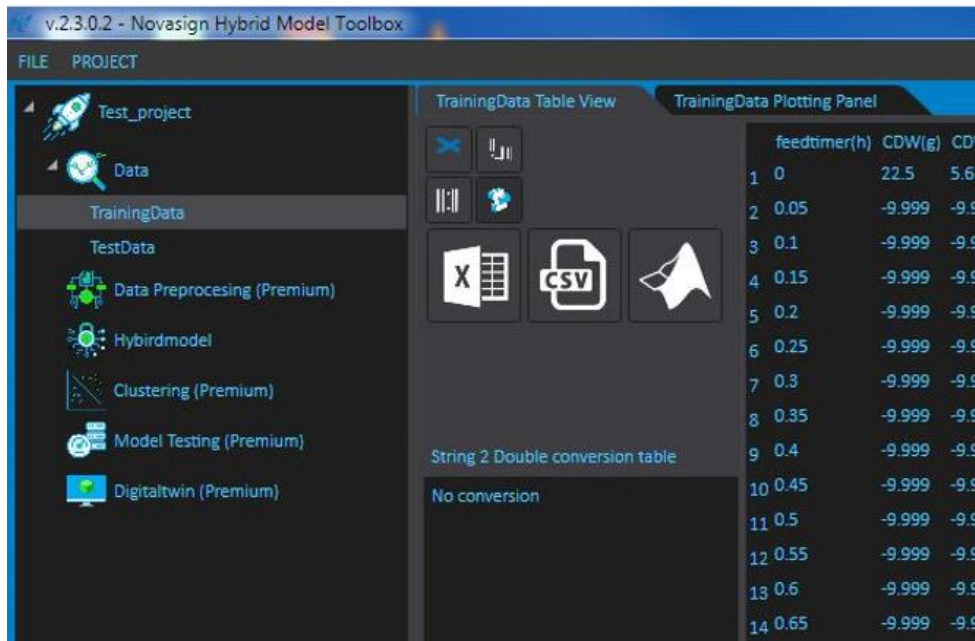


- once selected, a new window will open, where you have to specify the properties of your data sheet
- you have to define the character used as the decimal separator (1), eg., '.' or ','
- also, define the number of the row with your variable names and the column with the name of the data sets (run ID) (2). Make sure you have a number that identifies your run in the first row, followed by '-9.999' for the rest of each subset until you reach the next run (subset)
- in the case of the provided data sets, '.' is used as the decimal separator, the first row comprises the variable names and the run ID is in column 25.
- repeat these steps for all data files you want to import; take care that you choose the same properties for all data sets

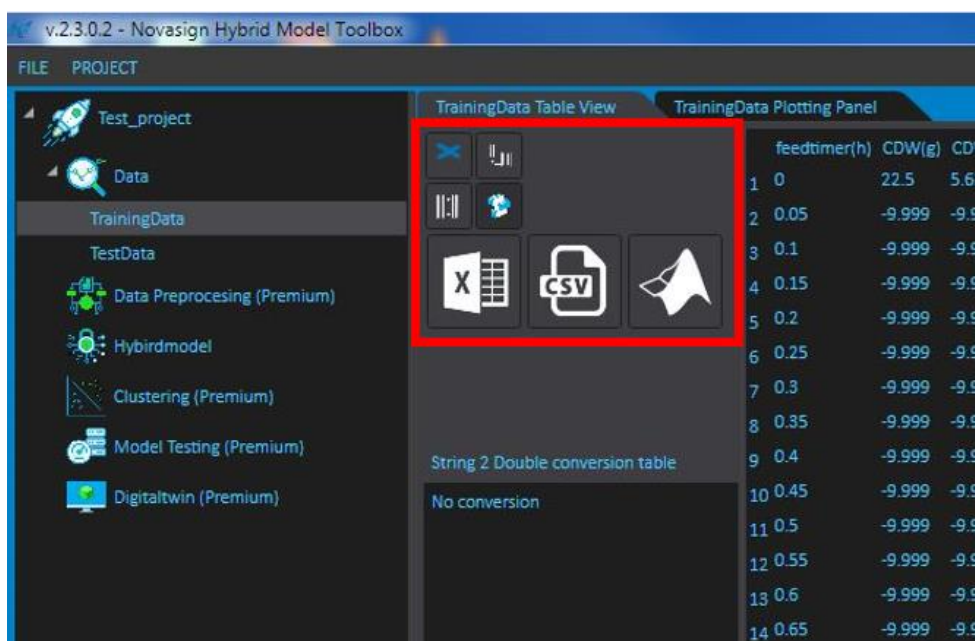


V. Inspect and plot your Data

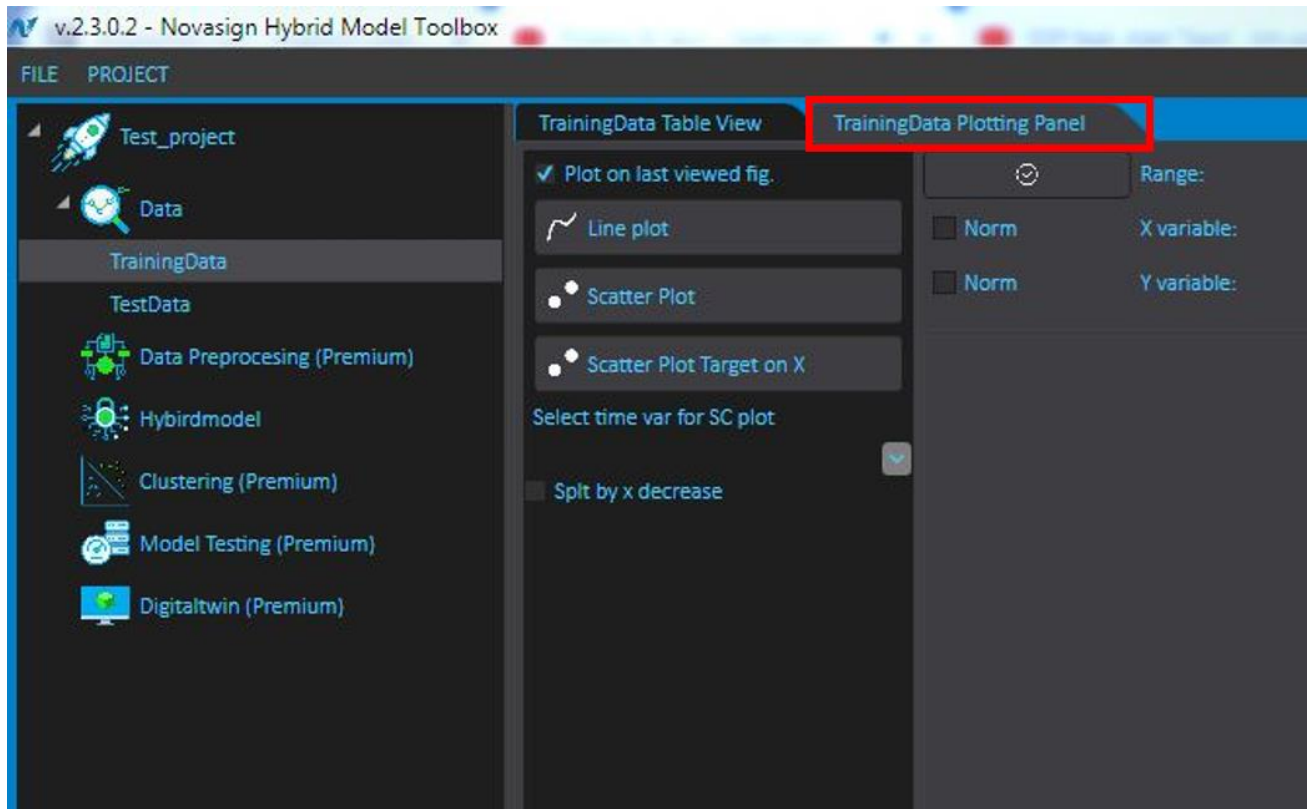
- the added data sets will appear in the subpanel 'Data'
- click on the respective data set to get into its table view tab. This has the same structure as your imported file before and additionally the properties you set while importing the data to the toolbox
- possible import errors are displayed in the lower-left panel



- you have different options how to deal with the respective data set, e.g., you can delete the respective dataset, append and merge datasets and export the data to different formats using the indicated buttons



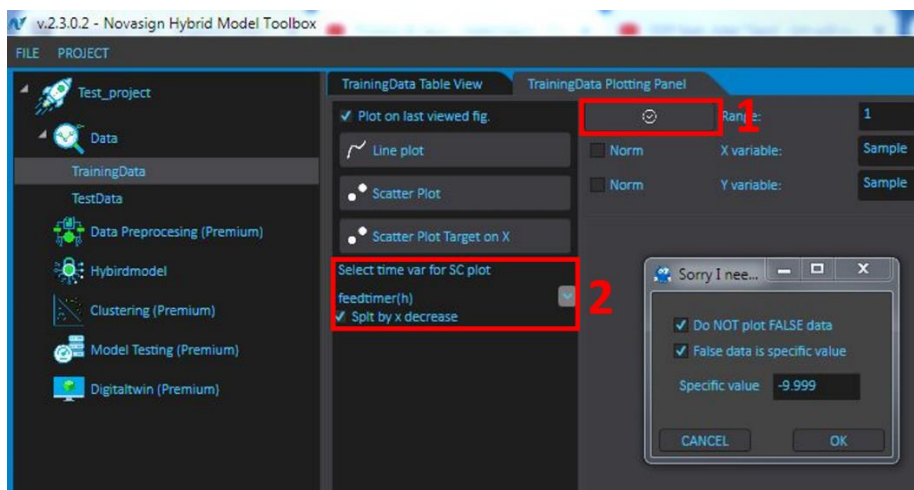
- in the second panel of the respective data set, you can visually inspect your data and plot it by switching to the plotting subpanel



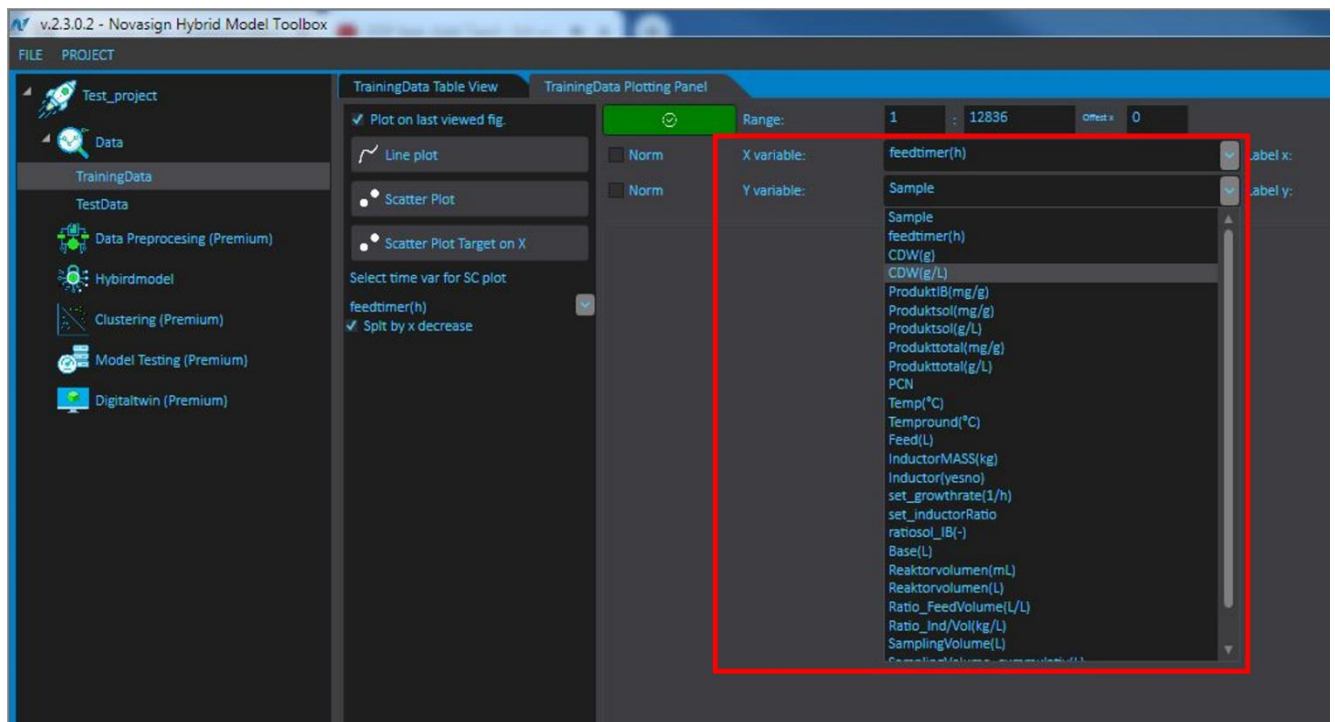
- in the beginning, select the invalid number you have chosen for your dataset by clicking on the button (1), i.e., activate the two boxes to not plot the false data and to assign the chosen value for non-defined data in your data set

- this function is activated when the box is green

- afterward, select the time variable of your data set and set 'split by x decrease' to separate the different runs from your data set to display them separately (2), i.e., every time the value for the chosen x variable (herein, the time variable) decreases, compared to the row above, it is considered a new experiment

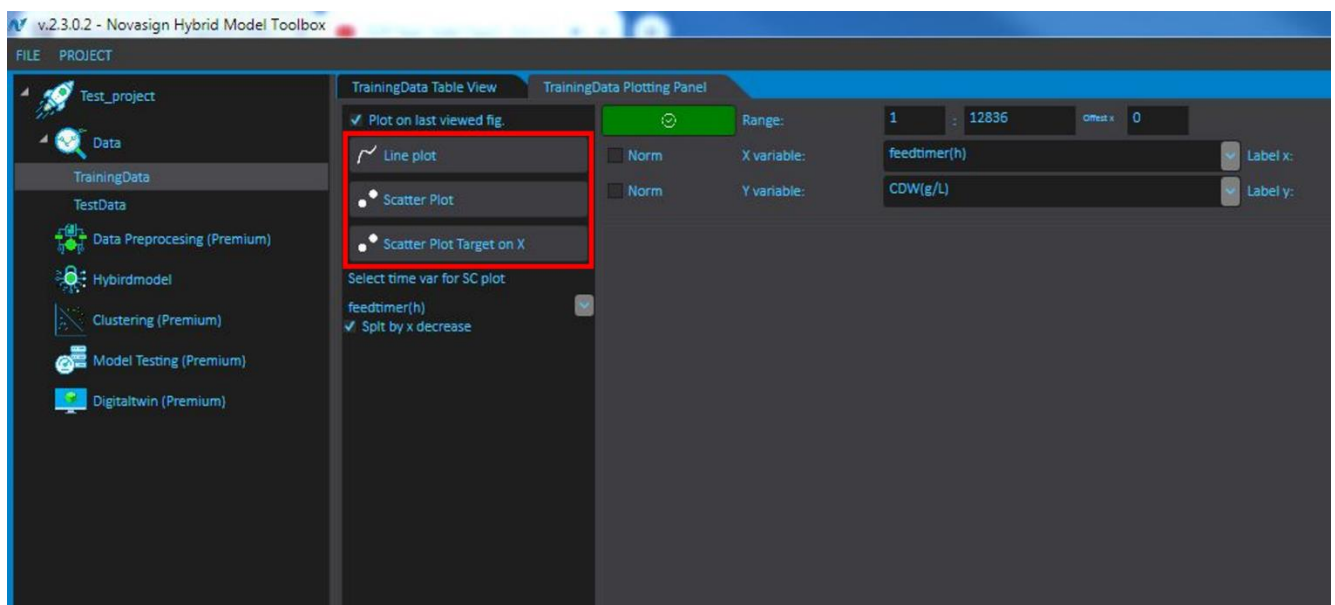


- in respect to the imported data set, select the x and y variables you want to plot, e.g., the feed time of the process as the x variable and the biomass concentration as the y variable

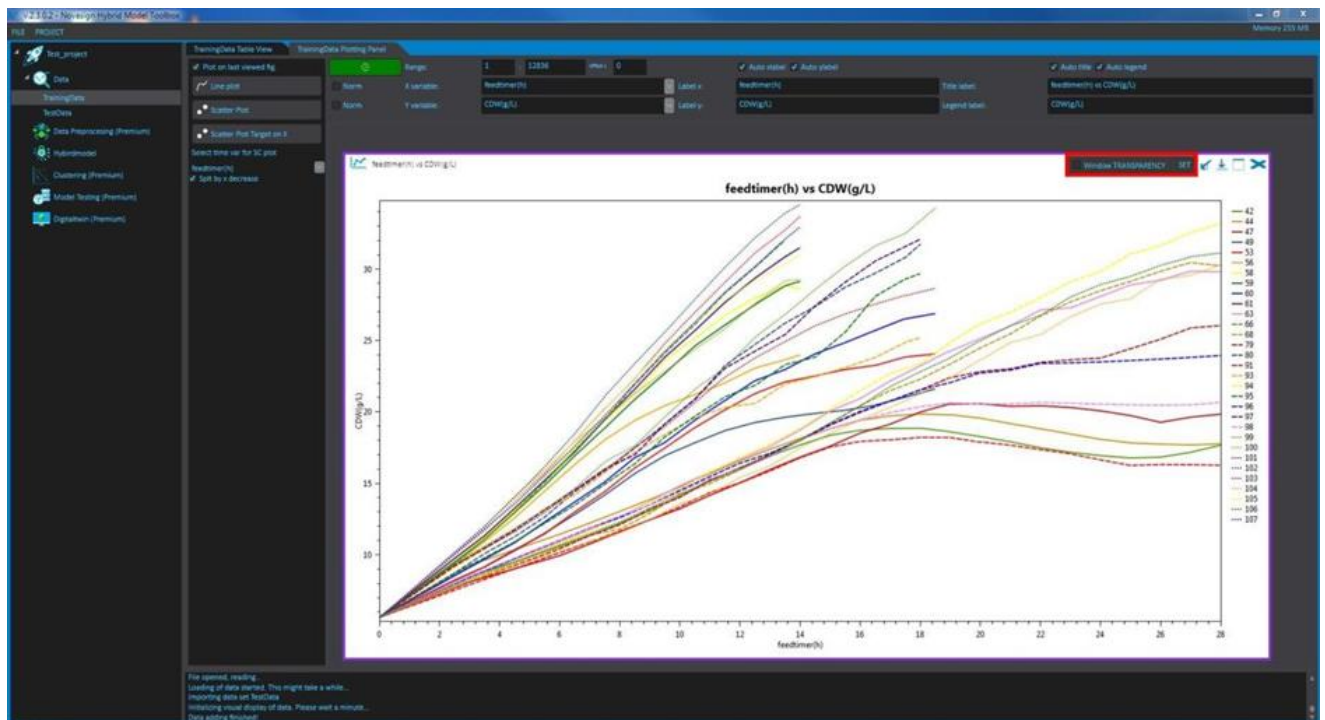


- if everything is set and the false data value is specified (green box), choose the plot you want to create; line plot, scatter plot or scatter plot with 1:1 line). You can also combine these plots by clicking on more than one plot option

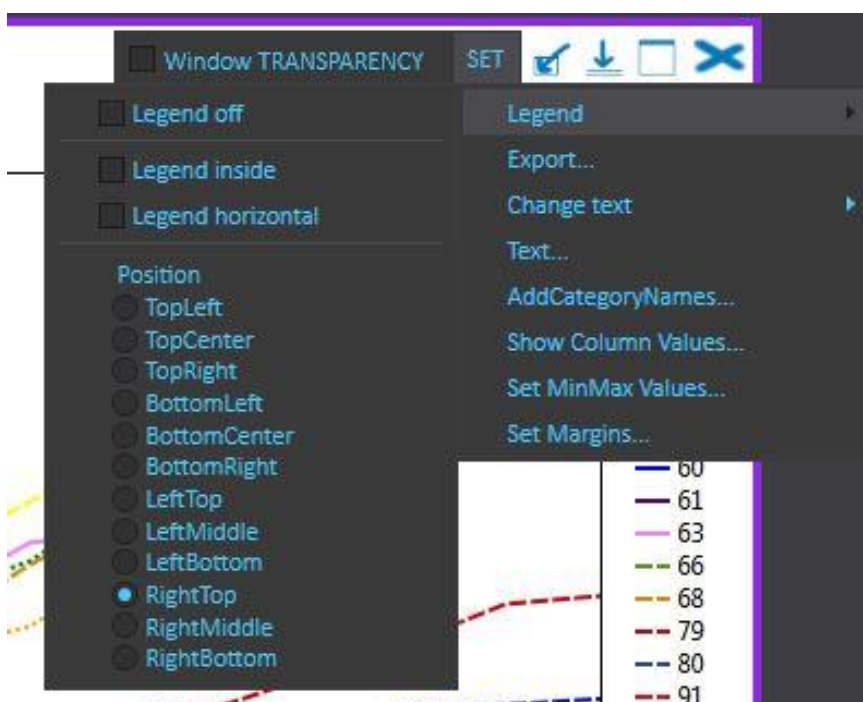
- additional options for your plot are available, next to the field where you define your x and y variables, you can also change the range of the displayed data set, e.g., if you only want to display certain experiments or rename the labels of your axes



- once the plot (in this example, a line plot) is created, it is displayed according to your set properties
- additional options for the plot can be chosen in the upper right button of the plot, highlighted in red



- there are various additional options to personalize your plot, aside from pining, minimizing, maximizing, and closing the plot. The tab 'SET' contains multiple options to modify and add features to your plot. You are able to change the location of your legend, export your plot and change the labeling for the title, the legend and the axes.



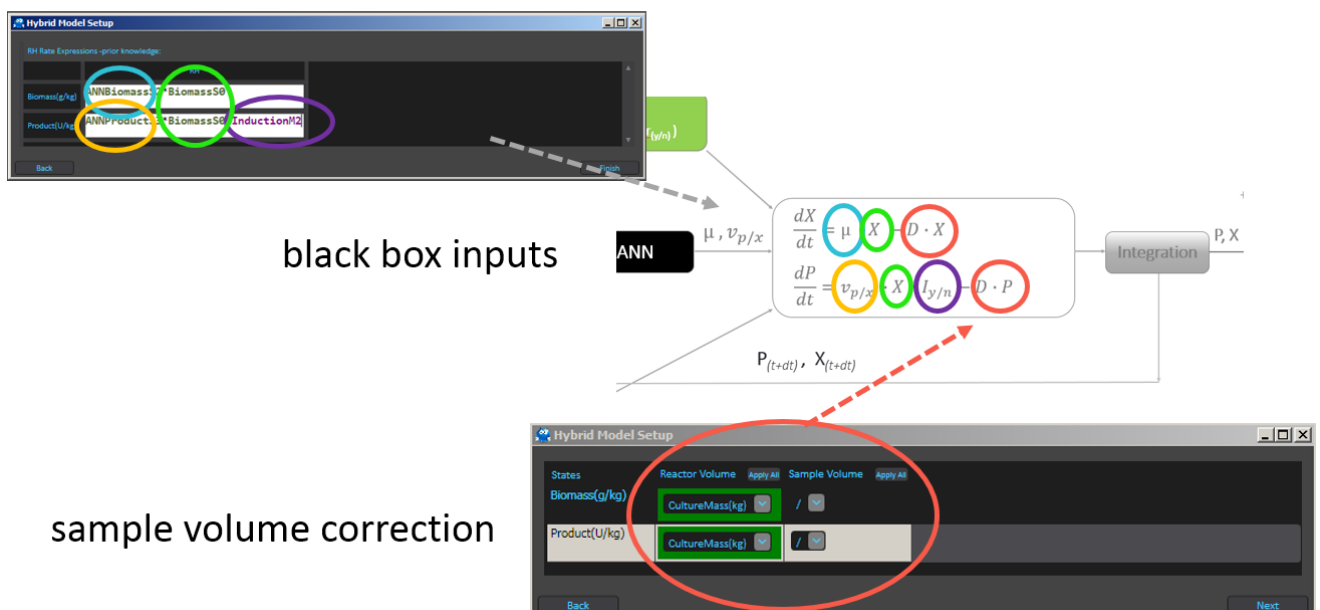
VI. Building a Hybrid Model

There are two general ways of modeling, either called black box or white box modeling. Black box models refer to purely data-driven approaches, for which no further process knowledge is needed. This kind of modeling is easy to approach and has good interpolation abilities. However, once the model is applied out of the performed design space, it lacks the ability to appropriately extrapolate and the model leads to poor modeling results. Otherwise, white-box models are based on first principles and therefore are a suitable match for extrapolation. Since they solely assume a mechanistic trend and do not account for empiric phenomena, also their predictability often ends up inaccurately.

One way to exploit the advantages of both modeling approaches is called hybrid modeling. Hereby, it is possible to benefit from the positive aspects of both modeling approaches, black box and white box, and make up for their respective drawbacks. This hybrid modeling approach establishes a more robust model, compared to one that only uses one of these techniques, to deal with more complex problems. In contrast to simple black or white box models, hybrid models are able to describe the whole process and not only the endpoints. As a result, the dynamics and deviations of a process can be incorporated. A graphical example of how a hybrid model structure may look like and how the different parts are integrated is given below.

The black box inputs to your data-driven model (ANN) are the measurements from your experiments, with which you want to model your target variables, e.g., for the provided exemplary data set, these are the biomass concentration and the product concentration. The outcome of the black box part will consecutively be used in the white box part of your hybrid model, in which they are incorporated into the first principles. Herein, the kinetic terms and the correction for the sample volume (if necessary) are taken into account for robust modeling of your target variable.

white box kinetics

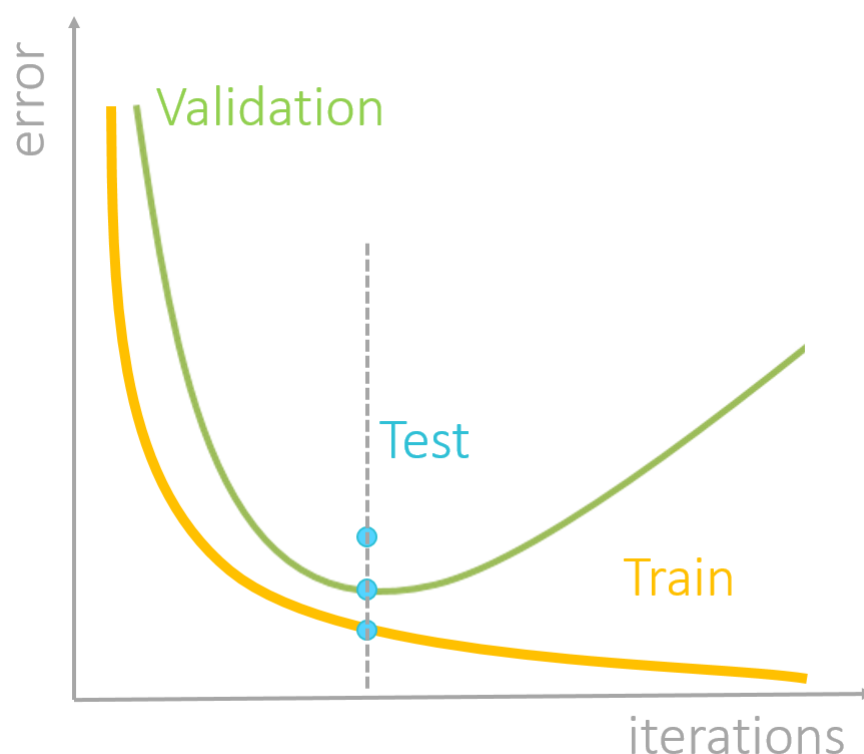


Your goal is to develop a hybrid model that performs well on new data and predicts its target variables accurately. To achieve this you have to take care of several aspects, e.g., to choose more iterations will only increase the model performance to a certain point until overfitting occurs. Hereby, the training error will decrease further and further, while the models' performance on new data will get worse. To avoid this issue, you should split your data into a training/validation set and a test set (external validation). This will ensure that the model will stay valid when applied to new data and avoids a wrong picture about the performance of your model due to overfitting.

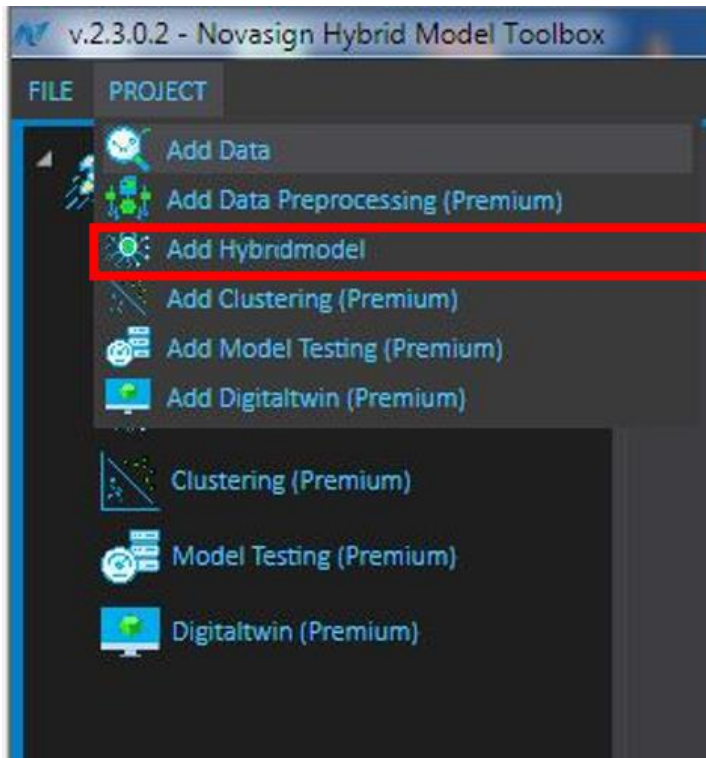
The model will be built using the previously defined training/validation set. This set is randomly split, with the train/val ratio you set, into a training and a validation part. While the model is built using the training part, its performance is determined by applying it to the validation part. This is called internal validation and used to test the models' predictability on new data (not used in the model-building process).

By doing so, the optimal number of iterations is found. Based on the performance of the internal validation, the best models can be chosen. These models are also applied to the previously defined test set to test their performance on new data (external validation).

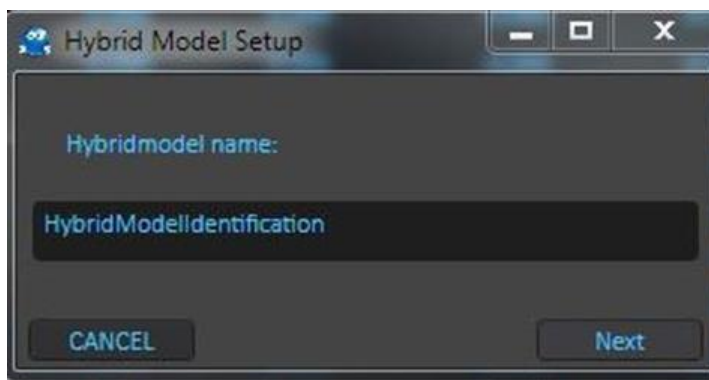
An example of how the number of iterations affects the model performance, with respect to the training, validation and test data, is given below.



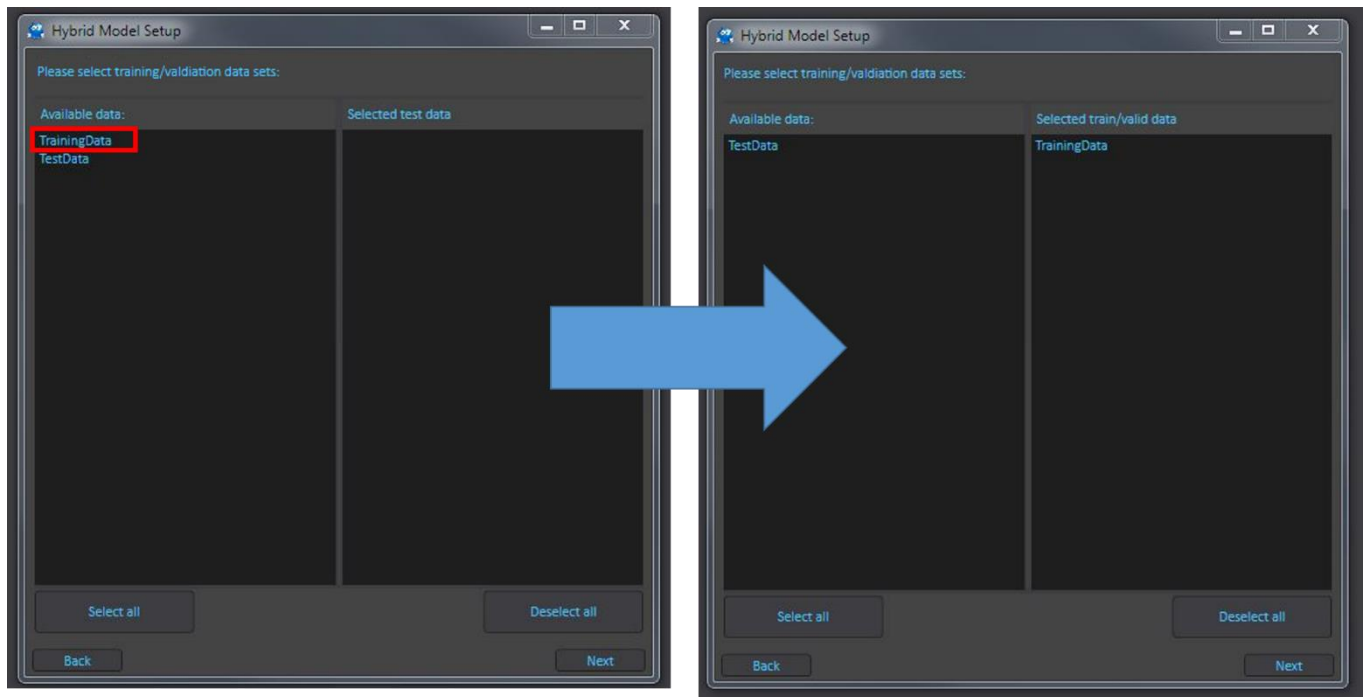
- to create a hybrid model with your added data sets, click on the tab 'PROJECT' and choose 'Add Hybridmodel'



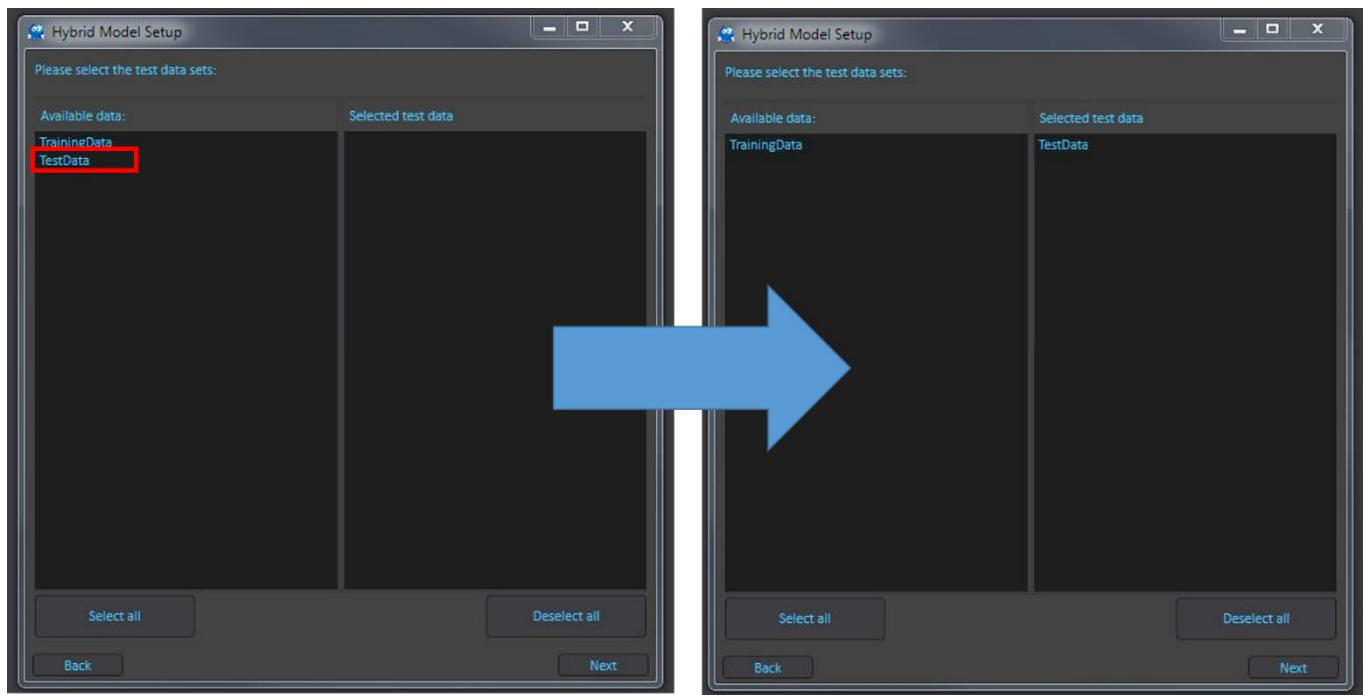
- first, a window will open, where you have to name your model



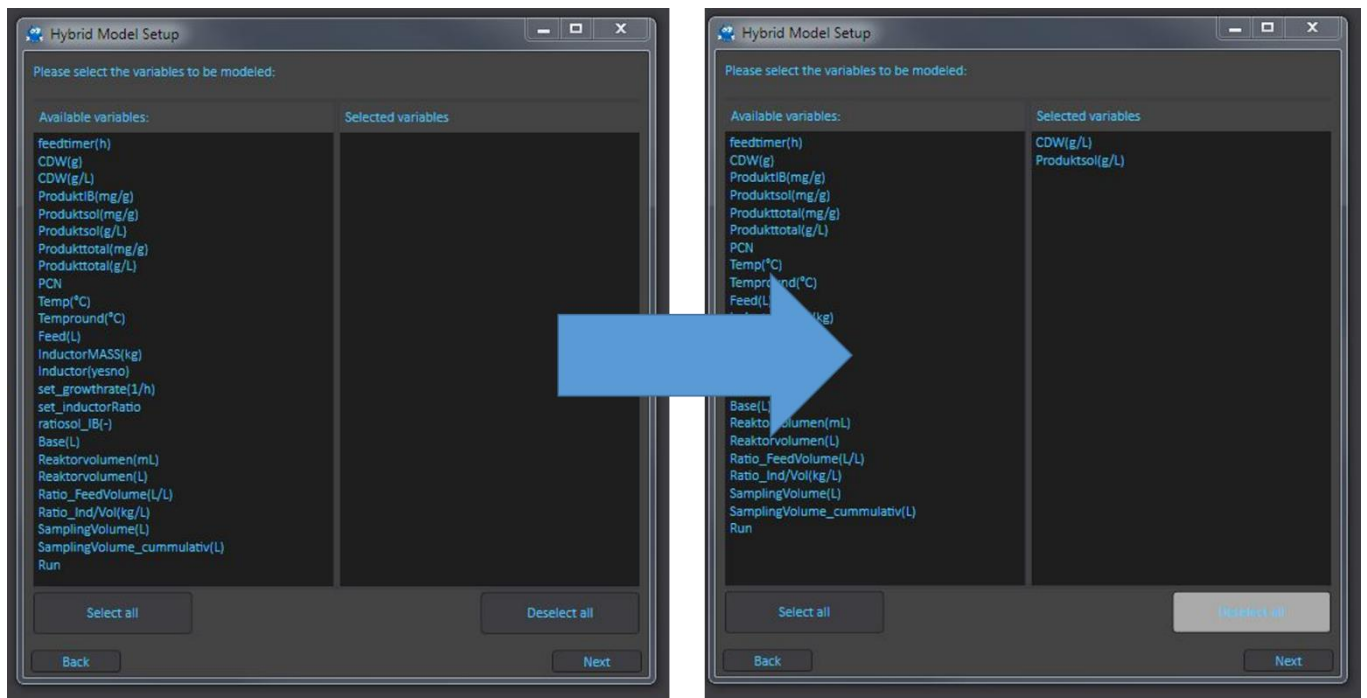
- in the next window, your added data sets are listed. First, select your data set you want to use for training and validation from the previously imported data sets by moving it to the right panel. You are able to add more than one data set if you want to



- likewise, in the next window, you select your data set for testing the established hybrid model from the previously imported datasets by moving it to the right panel. Again, you are able to add more than one data set, if you want to



- in the next window, all available variables out of your imported data sets will be listed. Herein, you select the target variables (state variables) to be predicted from your data set by moving them to the right panel. In our provided data set and project, these state variables are the biomass concentration and the soluble product concentration.

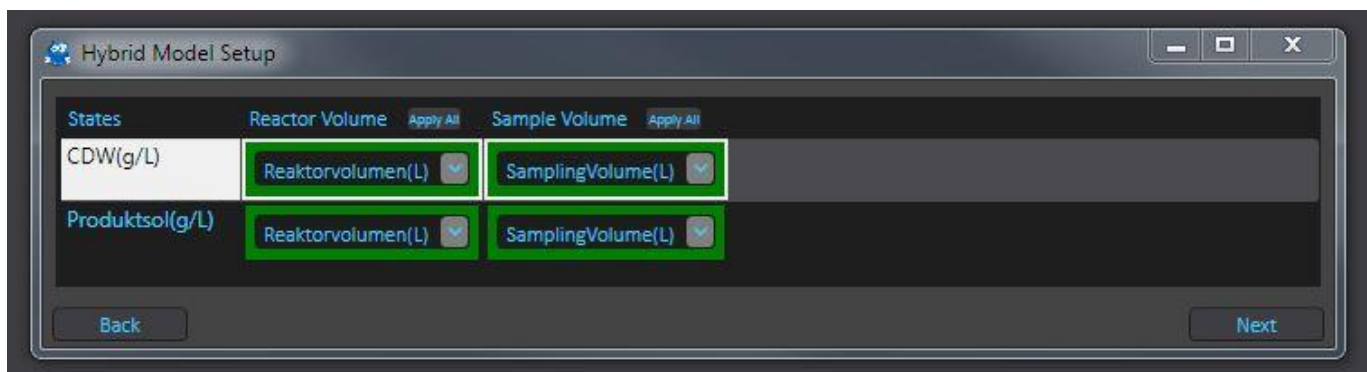


- the next window provides you the opportunity for correcting the volume of your process, e.g., in the case of fed-batch or sampling you have to define the column for volume compensation in your data sets.

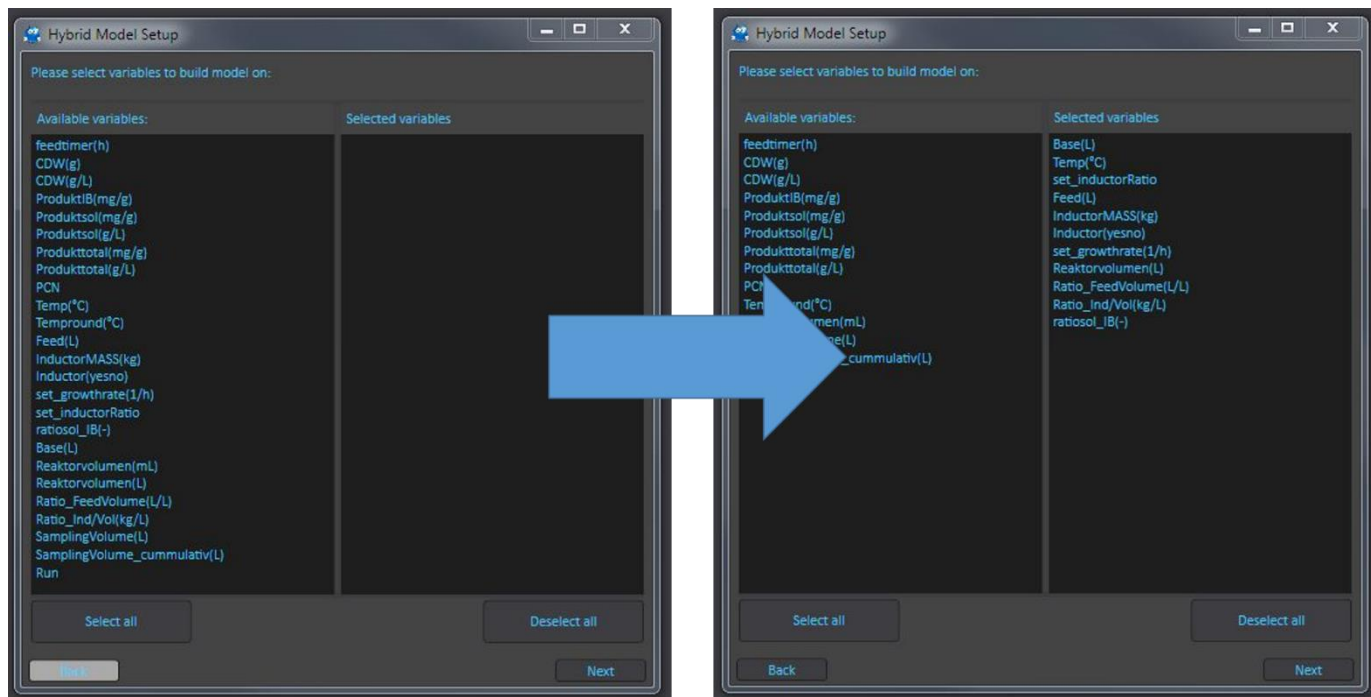
- it is important to provide these variables in kg or L to ensure the correct calculation

- for the reactor volume provide the cumulative volume over time, as it was done in our provided data set

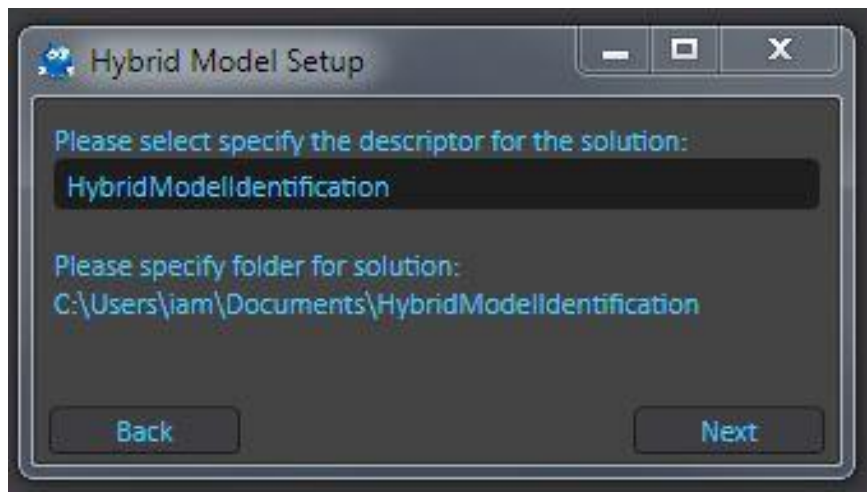
- for the sample volume provide the volume, which is extracted from the bioreactor at each time point; do not provide cumulative volumes, as it was done in our provided data set



- in the next window, add the variables you want to consider for building the data-driven black box part of your hybrid model. Select the wanted input variables by moving them to the right panel



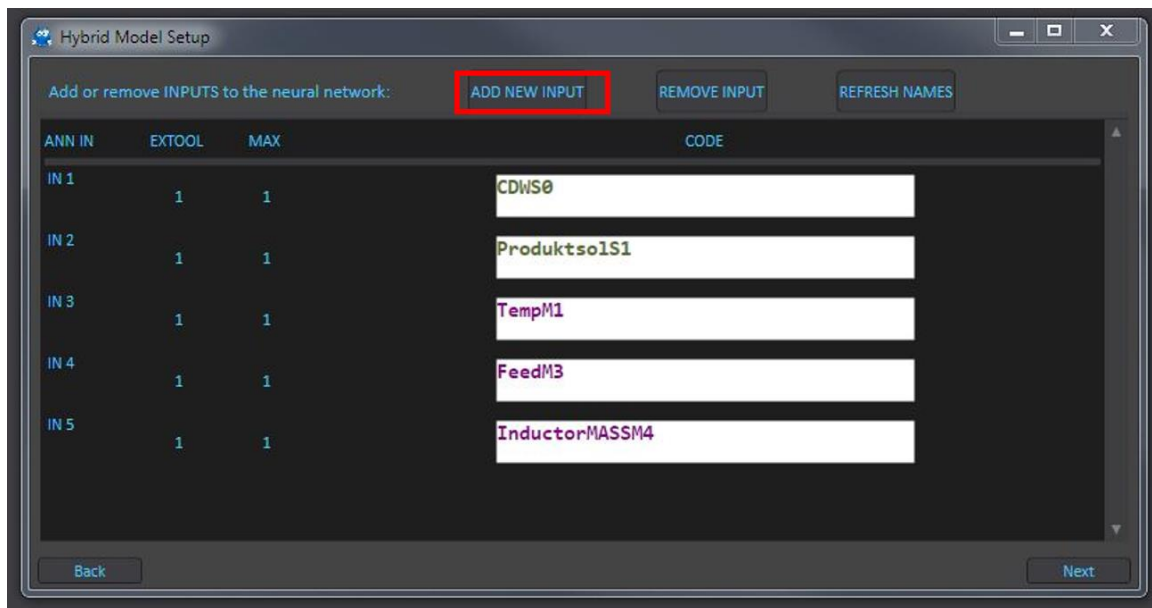
- select where and with which name the data should be stored on your PC



- to create the black box part of your hybrid model, set all wanted inputs for the neural network. All variables you previously added to consider in the model-building are available. You can add more input panels if needed.

- **green variables**: these are state variables to be modeled. You can also use them in the ANN as inputs. They will only be used in the first iteration for a good approximation of the starting point

- **purple variables**: these are modeling parameters (at best CPPs, which you can control) and will be used for each iteration



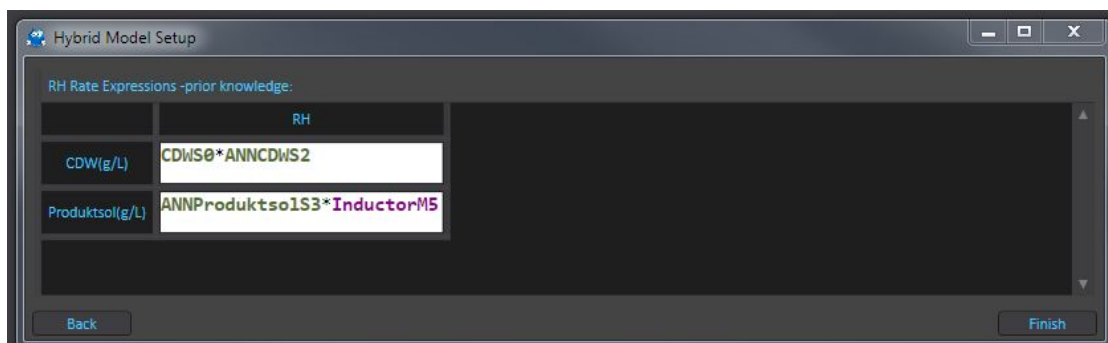
- after setting up the black box part, you define the kinetics for the white box part of your hybrid model

- you define your parametric part according to:

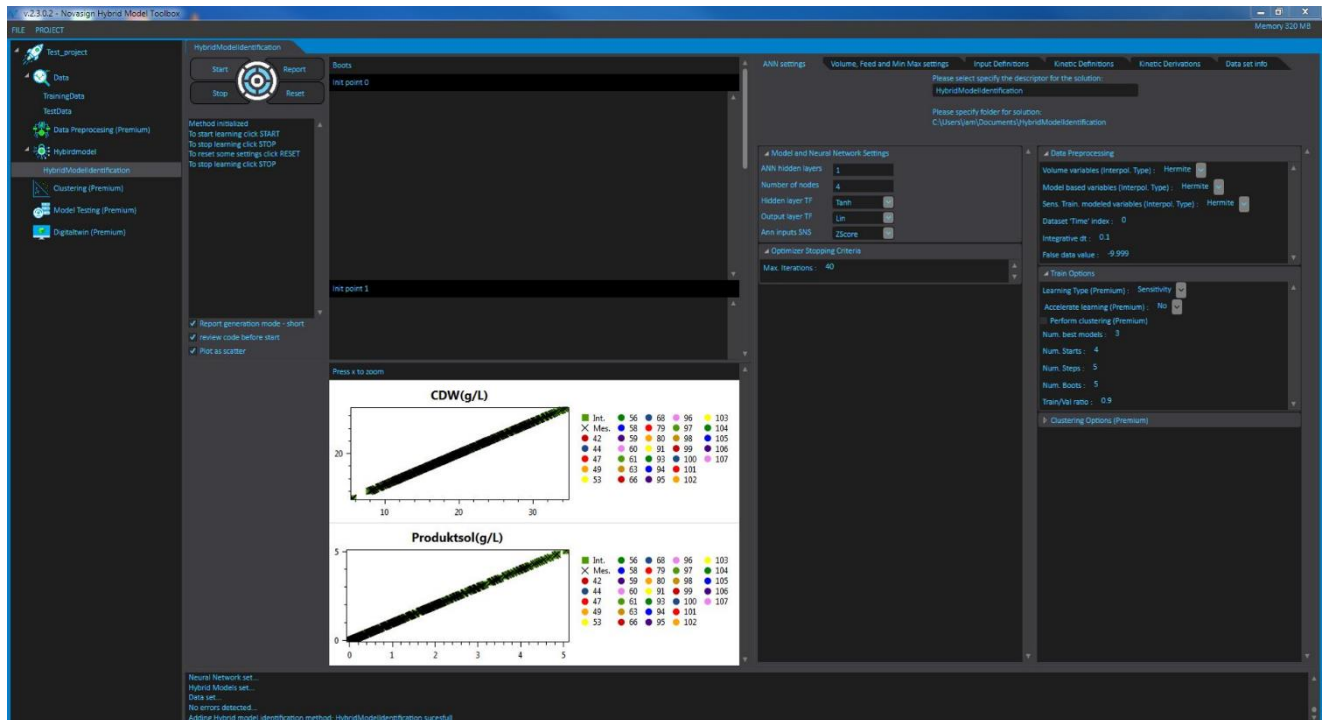
- your process knowledge
- literature

for help ask us via <https://portal.novesign.eu/>

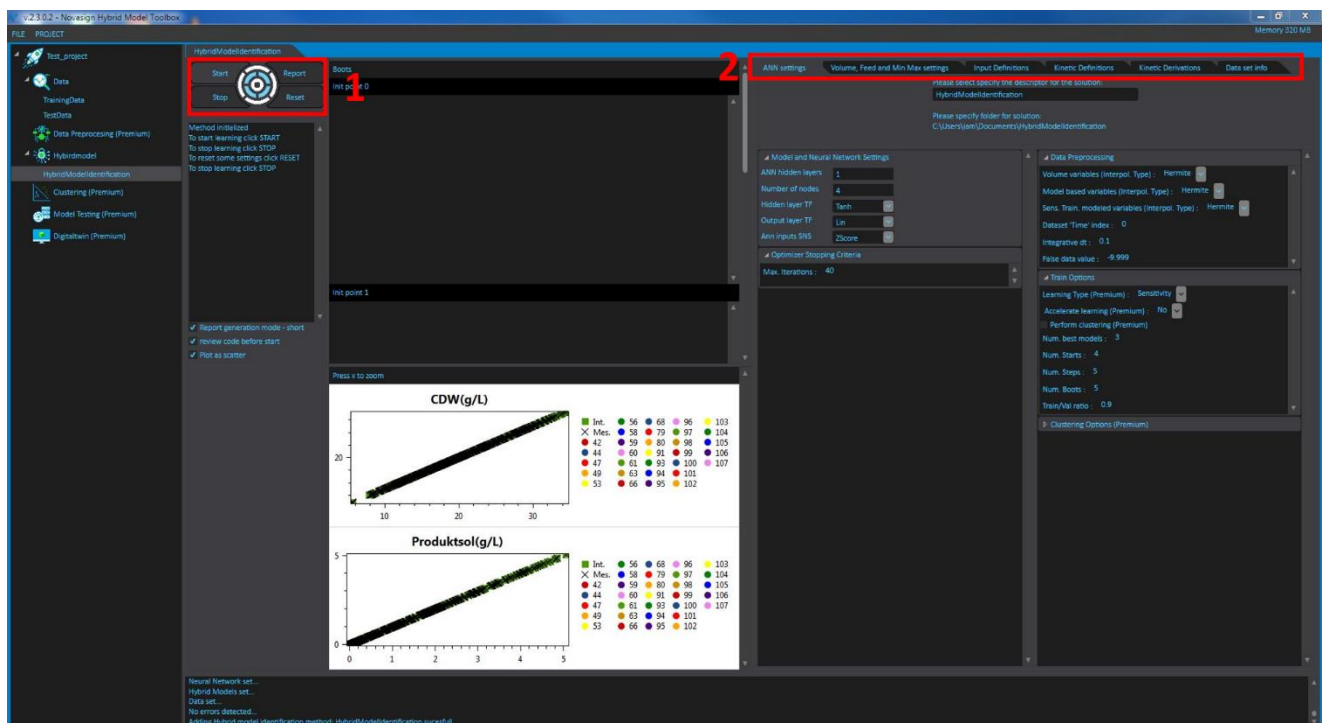
- for our provided project, the white box part takes the state variables (CDWS0), the outcome from the black box (ANNCDWS2 & ANNProduktso1S3) and our process knowledge in respect to the product formation (InductorM5, contains either '0' for no product formation or '1' for product formation) into account



- once everything is set, the respective subpanel in the projects hybrid model panel, with the name you gave your hybrid model, is generated



- you can use the left panel (1) to start ('Start') and stop ('Stop') the model. Once a model is finished with learning, the report can be generated ('Report') or the input parameters can be reset ('Reset'). All previously chosen black box inputs, kinetic expressions and data sets can be investigated in the according tab (2) and also changed after clicking the 'Reset' button. More details about the individual tabs are given below.



- in the tab 'ANN settings', you are able to configure all the inputs to the black box part

- Model and Neural Network Settings

ANN hidden layers: this defines the number of hidden layers you want to create for identifying your model. Most of the times one layer is sufficient

Number of nodes: this defines how many nodes your hidden layer/network should consist of. A good number to start with is the amount of ANN inputs -1

Hidden Layer TF: this will allow you to set the transfer function of your hidden layer. Most of the time 'Tanh' (tangens hyperbolicus) works best

Output TF: likewise, this sets the transfer function of your output layer. In almost any situation we encountered "lin" (linear) worked best

ANN input SNS: this allows you to choose the scaling/normalization/standardization method for your ANN inputs, ensuring fair conditions. This is necessary in order to compare different inputs to the ANN, typically all having different dimensions. Zscore is a good standard used for machine learning

- Optimizer Stopping Criteria

Max. Iterations: this defines the maximum number of performed iterations before a new Jacobian is set

- Data Preprocessing

Volume variables: here, you choose the type of interpolation used for the volume variable, to correct for the increasing volume in a fed-batch and the volume reduction due to sampling

Model-based variables: this sets the interpolation type of your input parameters you build your model and the ANN on

Sens. Train. modeled variables: this sets the interpolation type of your target/state variables.

Remember that you need a minimal number of observations for each interpolation type, as explained in section 'III. Prepare the Data'

Dataset 'Time' index: defines the column in which the time is displayed (0 stands for 1). Most of the time it is the first column, if not, define it here, e.g., if the time is in column 8, write 7

Integrative dt: defines your time step, i.e., how time-resolved you want to model. Do not use a value below your minimum online sampling time

False data value: likewise as in the plotting panel earlier, define the false/invalid data value you assigned to your empty fields while preparing your data

- Train Options

Num. best models: here, you can set how many of the best performing models per start point will be available to report, after the model-building process

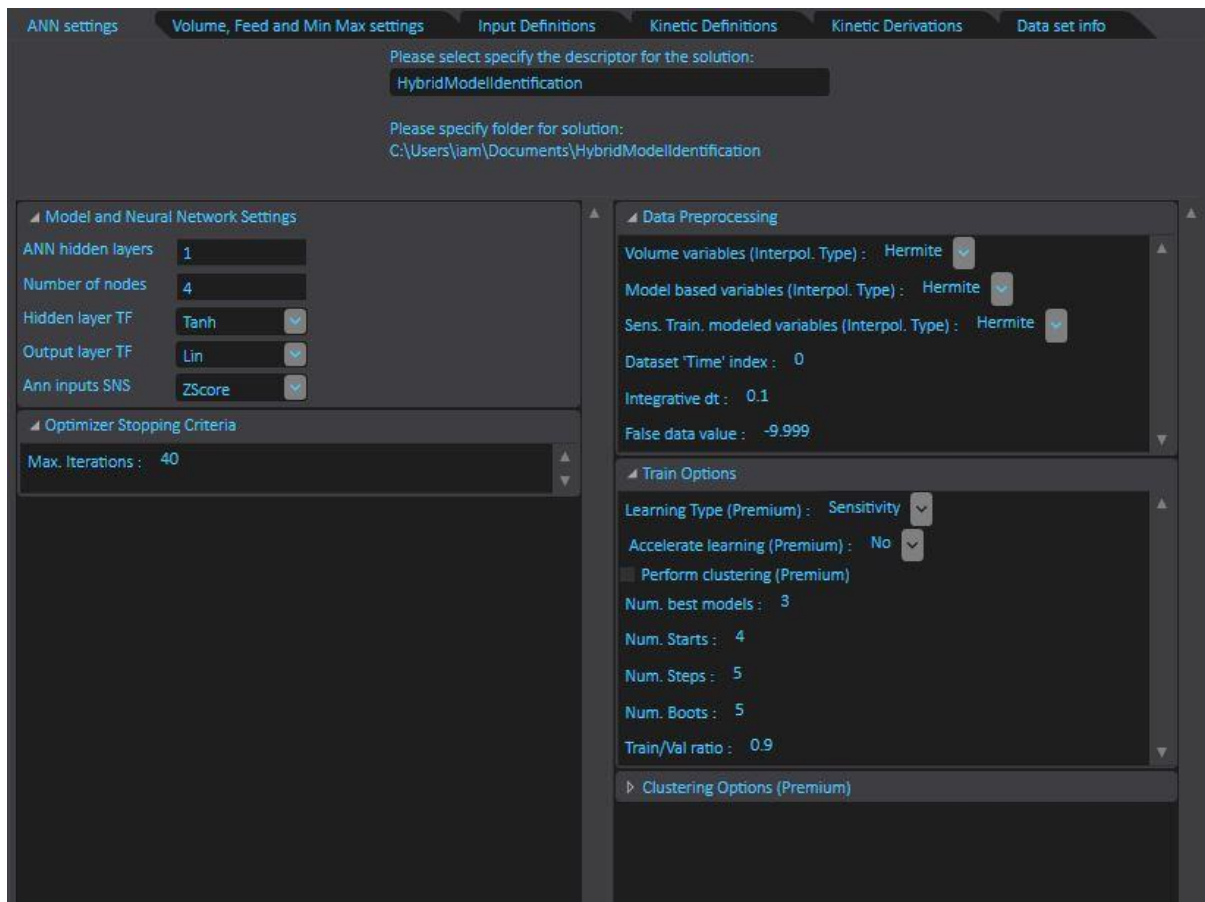
Num. Starts: this number defines how often your ANN is initiated with new random weights, to find the global minimum. The more starts you use, the more likely you will find a good solution to avoid local minima

Num. Steps: defines how often the Jacobian Matrix is reset (5 is a good number to start with)

Num. of Boots: defines how often you want to shuffle your data with the set Train/Val ratio. For example: with 10 runs and a 0.8 ratio, 45 different combinations are possible (using 5-10 boots always gives you a good guess of your setup quality)

Train/Val ratio: splits your train/val dataset into training and validation runs according to your chosen ratio. For example 10 runs in the training/validation set and a 0.8 ratio → 8 training and 2 validation runs will be chosen

Recommended setup: for 4-10 runs: 0.75, for 10-30 runs: 0.8, for > 30 runs: 0.9



Example: with 4 train/val runs and a 0.75 ratio → 3 train and 1 validation runs will be chosen per boot. Setting up a hybrid model with 4 boots, each with 10 steps, 5 starts, and 20 iterations will take approximately 15 minutes without parallelization → 1-3 with parallelization (Premium)

- in the tab 'Volume, Feed and Min-Max settings' you are able to investigate and change the settings you defined for the volume compensation earlier

ANN settings
Volume, Feed and Min Max settings
Input Definitions
Kinetic Definitions
Kinetic Derivations
Data set info

States

Reactor Volume

Apply All

Sample Volume

Apply All

CDW(g/L)

Reaktorvolumen(L)

SamplingVolume(L)

Produktsol(g/L)

Reaktorvolumen(L)

SamplingVolume(L)

STATE VARIABLE	STD VARIABLE	IXMED VARIABLE	MIN	MAX	OUT VARIABLE	MIN
CDW(g/L)	<input checked="" type="checkbox"/>	Base(L)	-1E+20	1E+20	CDW(g/L)	0
Produktsol(g/L)	<input checked="" type="checkbox"/>	Temp(*C)	-1E+20	1E+20	Produktsol(g/L)	0
		set_inductorRatio	-1E+20	1E+20		
		Feed(L)	-1E+20	1E+20		
		InductorMASS(kg)	-1E+20	1E+20		
		Inductor(yesno)	-1E+20	1E+20		
		set_growthrate(1/h)	-1E+20	1E+20		
		Reaktorvolumen(L)	-1E+20	1E+20		
		Ratio_FeedVolume(L/L)	-1E+20	1E+20		
		Ratio_Ind/Vol(kg/L)	-1E+20	1E+20		
		ratiosol_IB(-)	-1E+20	1E+20		

- in the tab 'Input Definitions', you are able to modify the set inputs to the black box part. You can add or delete variables to see if and how the performance of your model changes

ANN settings

Volume, Feed and Min Max settings

Input Definitions

Kinetic Definitions

Kinetic Derivations

Data set info

	EXTOOL	MAX	CODE
IN 1	1	1	CDWS0
IN 2	1	1	ProduktsolS1
IN 3	1	1	TempM1
IN 4	1	1	FeedM3
IN 5	1	1	InductorMASSM4

	CDW(g/L)	Produktsol(g/L)
dOut1	1	0
dOut2	0	1
dOut3	0	0
dOut4	0	0
dOut5	0	0

Change Input Definitions

- the tab 'Kinetic Definitions' allows you to investigate and modify the white box part you previously defined. Likewise, as for the black box inputs, you can change the white box part to see if and how the performance of your model changes

ANN settings Volume, Feed and Min Max settings Input Definitions **Kinetic Definitions** Kinetic Derivations Data set info

K-Matrix of stoidchometirc or yield coefficients:

	CDW(g/L)	Produktsol(g/L)
CDW(g/L)	1	0
Produktsol(g/L)	0	1

RH Rate Expressions -prior knowledge:

	CDW(g/L)	Produktsol(g/L)
RH	CDWS0*ANNCDWS2	ANNProduktsolS3*InductorM5

Change Kinetic Definitions

- the tab 'Kinetic Derivations' allows you to investigate the derivatives of the defined white box part from the tab 'Kinetic Definitions'

ANN settings Volume, Feed and Min Max settings Input Definitions Kinetic Definitions Kinetic Derivations Data set info

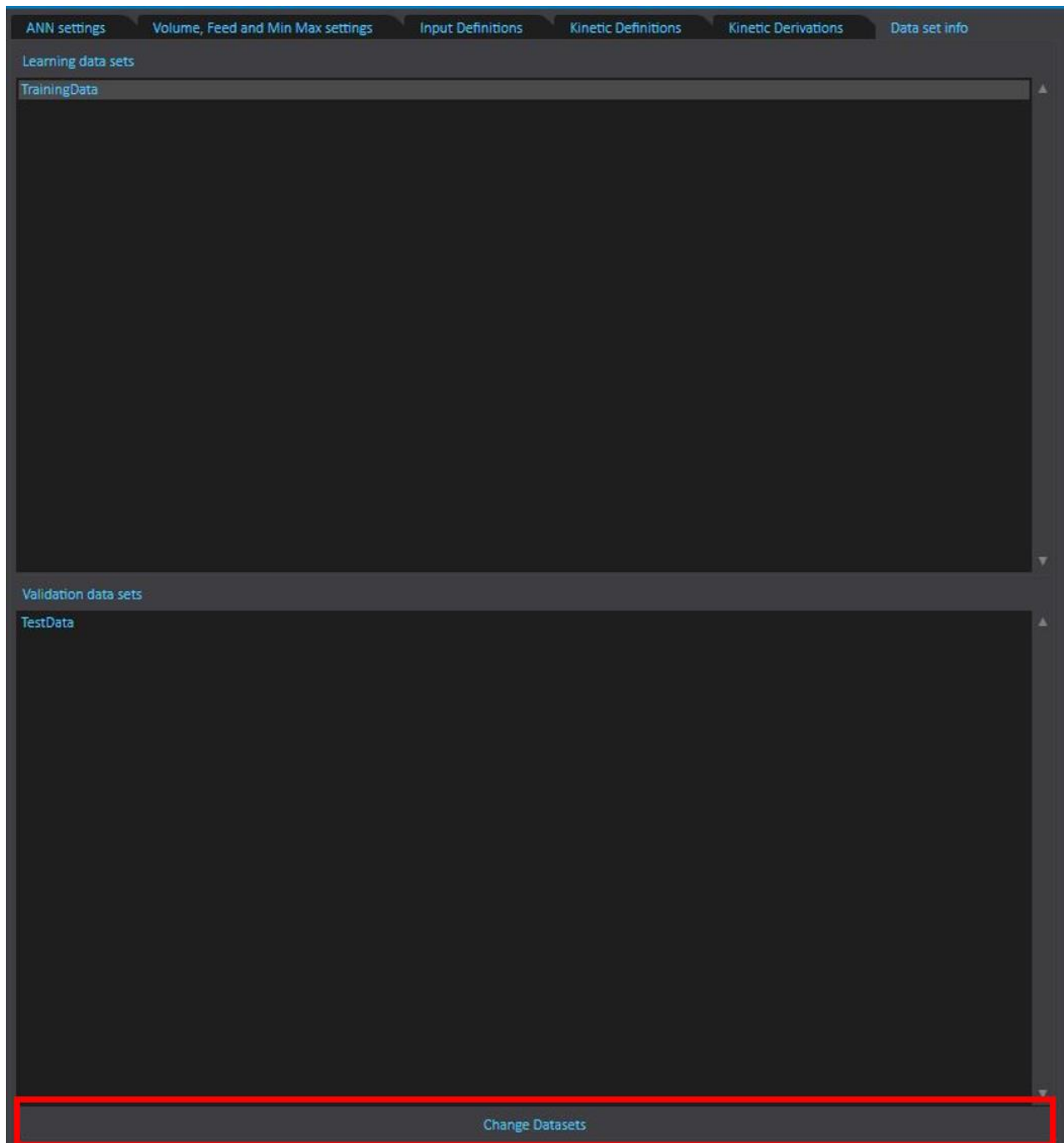
dRH/dX:

	CDW(g/L)	Produktsol(g/L)
CDW(g/L)	ANNCDWS2	0
Produktsol(g/L)	0	0

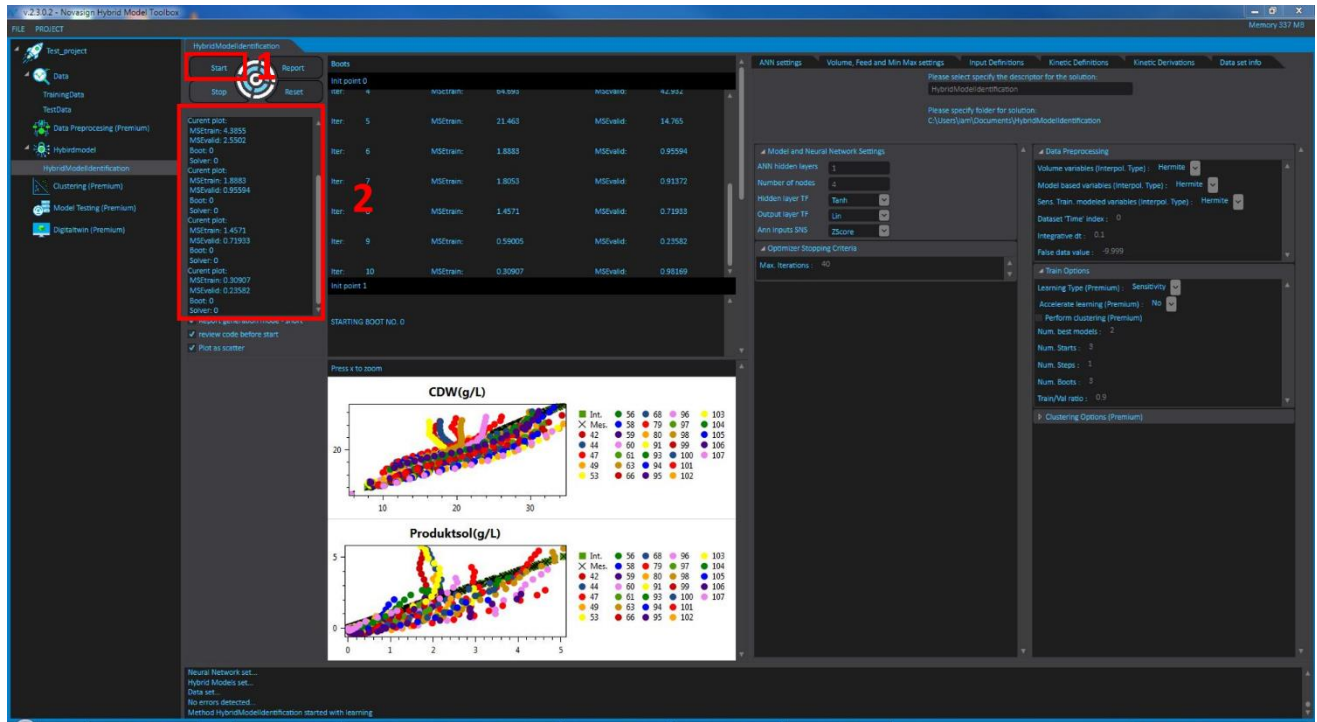
dRH/dRO:

	CDW(g/L)	Produktsol(g/L)
CDW(g/L)	CDWS0	0
Produktsol(g/L)	0	InductorM5

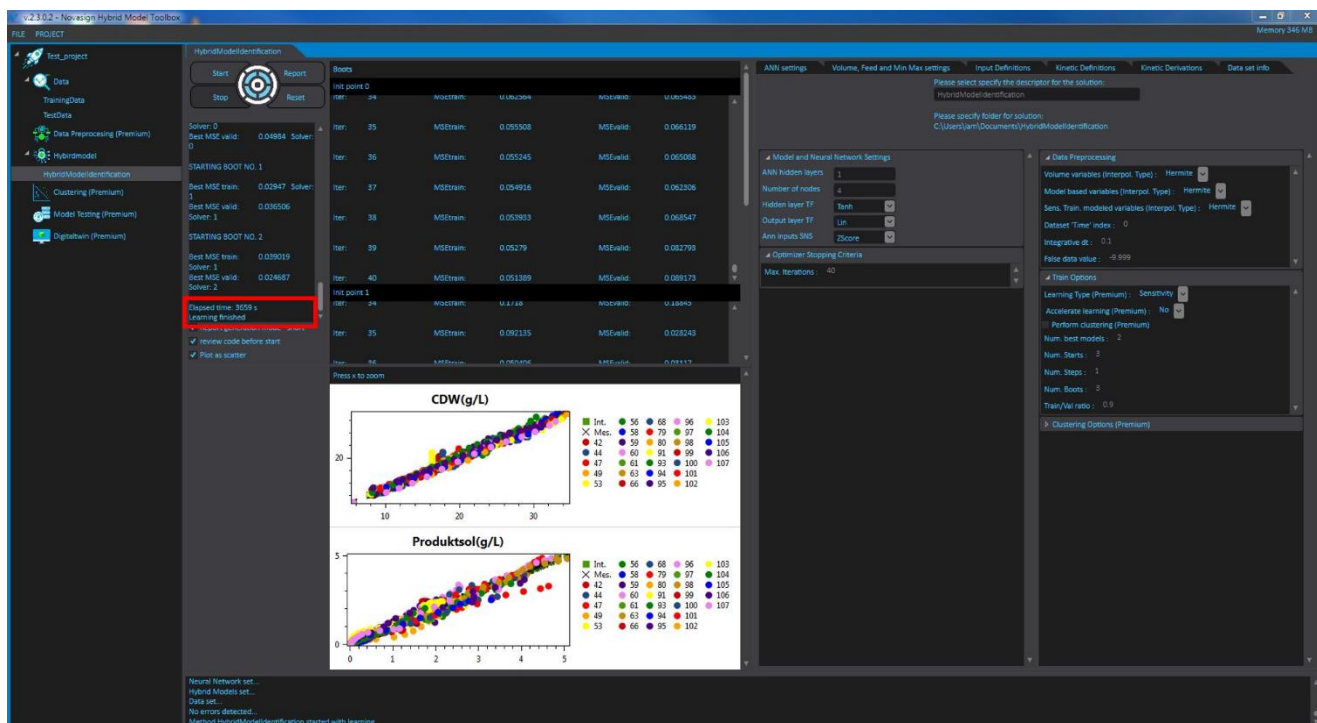
- the tab 'Data set info' allows you to investigate the used data sets you have chosen for training/validation and testing of the model. You can also change these data sets to any of the previously imported data sets



- once you have defined all settings, you can start the model-building with the chosen properties. To start the model, click on 'Start' in the upper left panel (1). The current progress and the performance of the best model are displayed in the panel below (2). Also, the scatter plots, for the training/validation data set, of the state variables for the currently best performing model are displayed

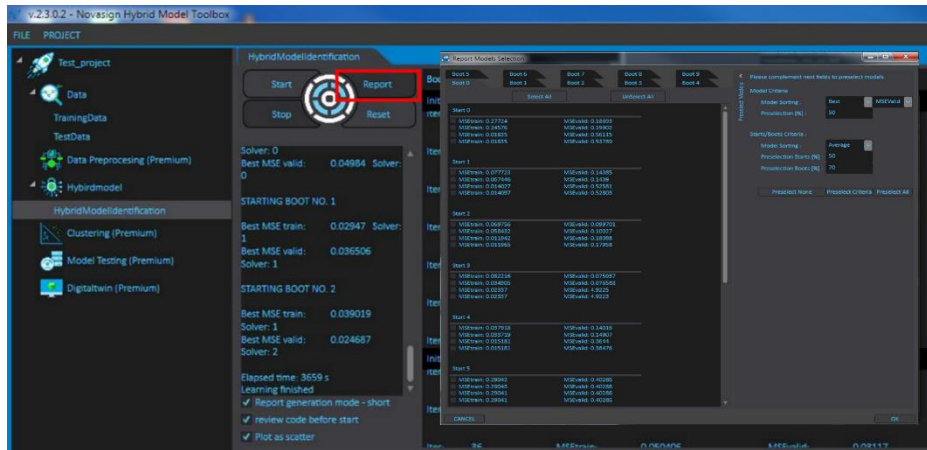


- once the model is finished, the elapsed time and the finishing message are displayed in the panel. Now you are able to investigate and evaluate your generated hybrid models by reporting them



VII. Evaluate the Results

- to create the report for the created models, click on the 'Report' button in the upper left panel. They are sorted into the respective boot and herein subdivided in the number of starts, as you previously defined them

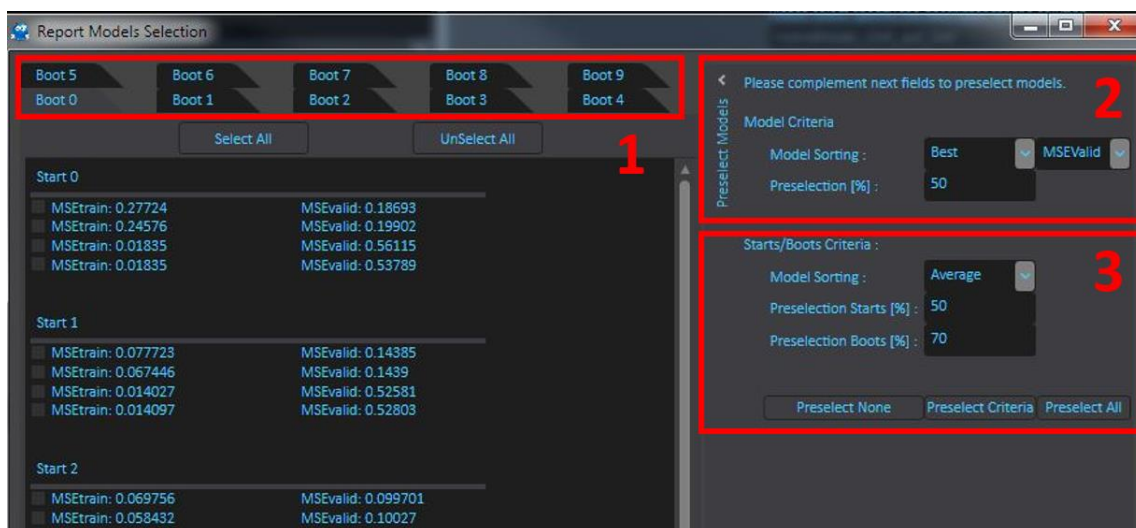


- the box will show you a tab, where you can switch between all performed boots (1). Herein, the chosen number of best performing models is displayed for each starting point. each checkbox enables you to display the respective model in the report section. The models respective mean squared error (MSE) for training and validation, indicating the models' performance on the internal validation, is given. Out of all displayed best models, select the models you want to create a report for

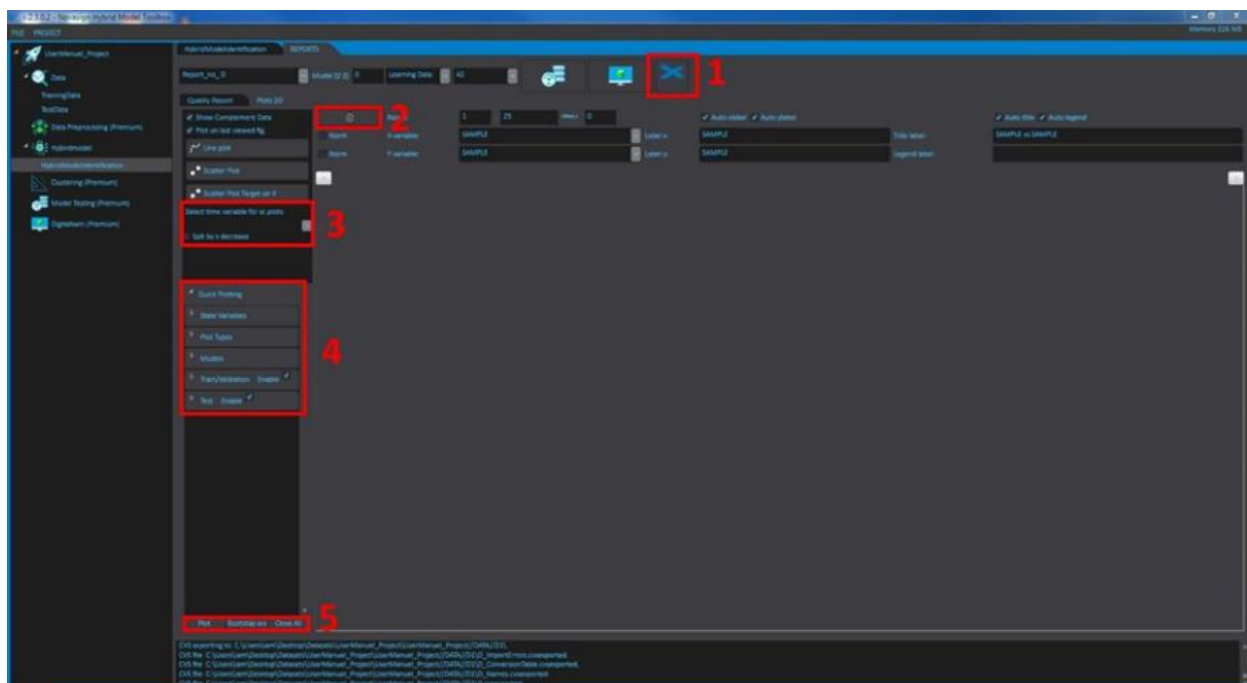
- in case you need a little help, you can use our options for preselecting models, according to the criteria you set

- in the first field you can preselect the general model criteria you want to choose automatically (2). This includes the option for 'Model Sorting' (Best or Worst) and the respective criteria (MSETrain, MSEValid or MSETrainValid) and how many of all these models you want to preselect (in percentage)

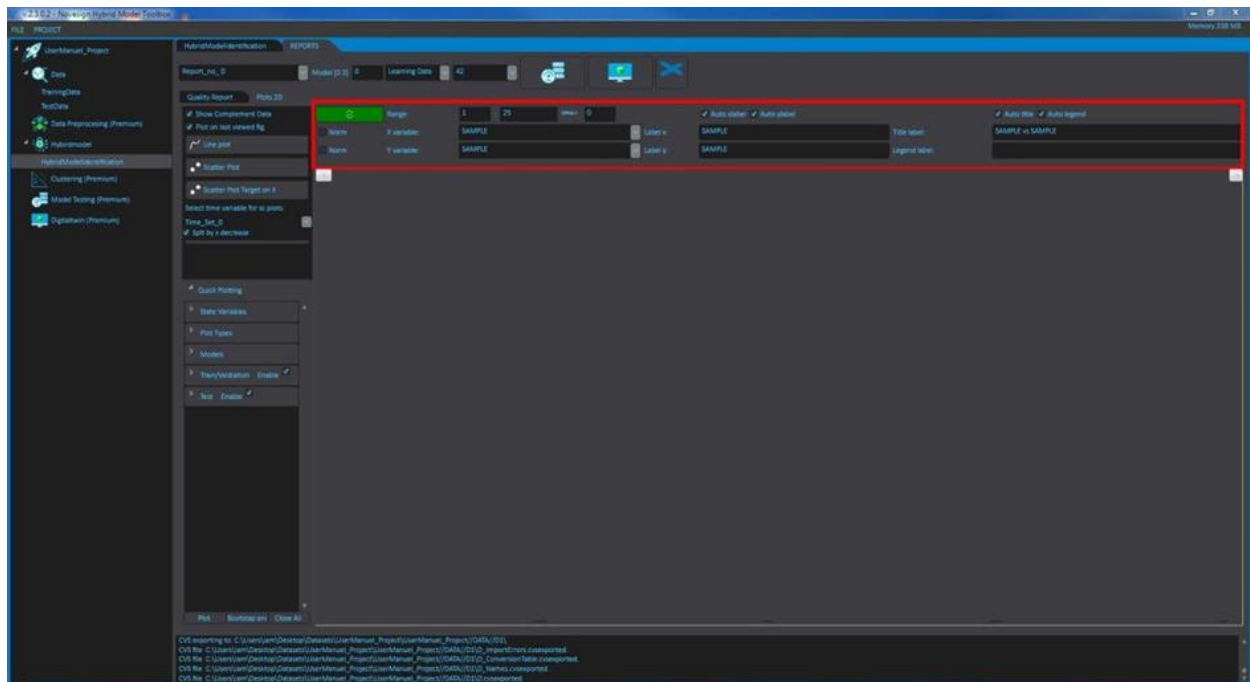
- the Starts/Boots Criteria (3) let you choose how many starts and boots, in percentage, you want to consider for preselecting the best performing models



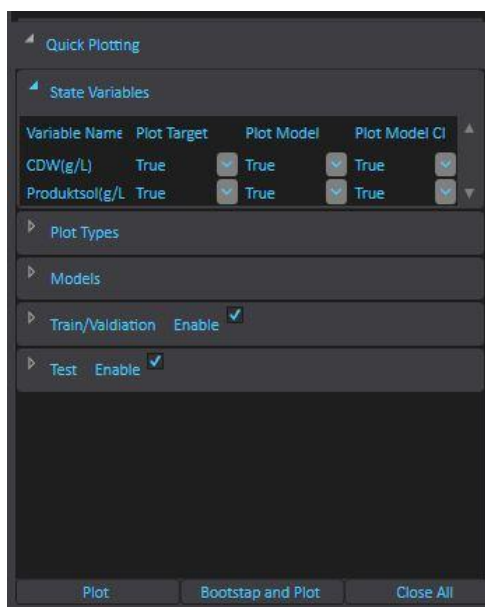
- once the report is created, a new tab, called 'REPORTS', appears above the upper left panel, next to the panel named as your hybrid model. This tab contains all reports that were and will be created
- In the tab 'Quality Report' the metrics for all of your reported models can be investigated and different ways to describe the models are listed
- the tab 'Plots 2D' contains all the plotting options for the created reports. You can switch between all created reports by choosing the wanted report in the upper left panel of this tab
- if wanted, the report can be deleted (1), e.g., to save memory. For simple plotting as before in the 'Data' section, set the specific value of non-defined data (2) and separate the runs (3)
- the advanced settings for plotting are displayed in an extra panel (4). Here, you are able to display the models as scatter and line plots for the state variables. Different plots for the model performance on the training/validation set and the test set will be displayed. These can be plotted, bootstrapped and again closed via the bottom left panels (5)
- when clicking 'Plot', all of your models are displayed; they are sorted by increasing errors



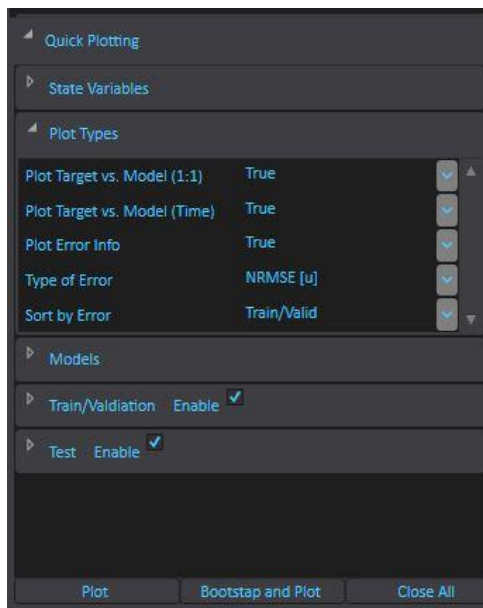
- for simple plotting, the same options as before in the 'Data' panel for plotting can be chosen. In the marked panel the options for labeling, choosing the x and y variables and the wanted plot type are seen. You can use these options likewise as before



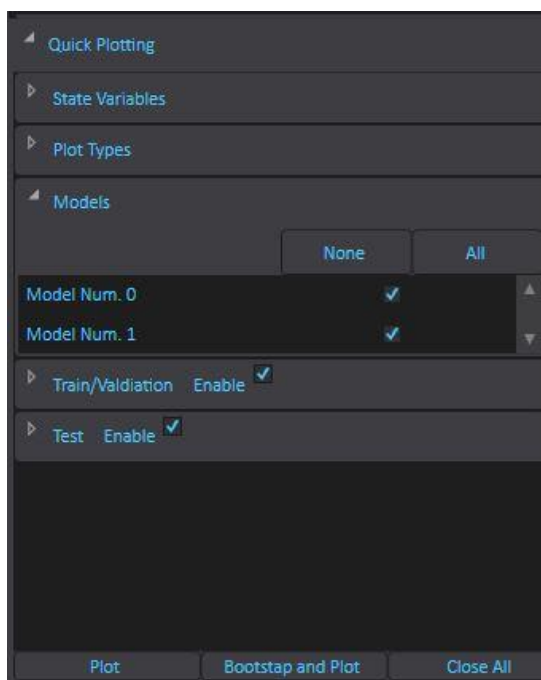
- for the advanced model plotting, you can choose between several options in the first panel. You can choose if the state/target variables and the models predicted value are displayed (true) or hidden (false) in the plots. Likewise, for the confidence interval, which is only possible for bootstrapped models.



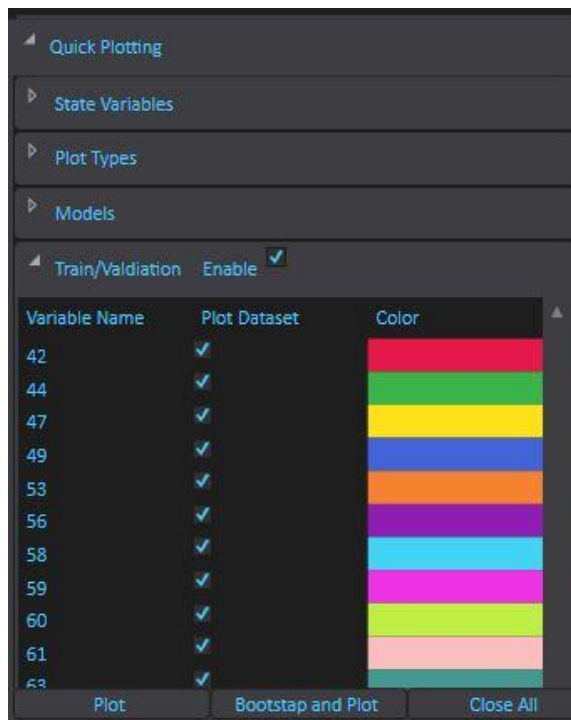
- in the panel 'Plot Types', you are able to choose if you want to display the scatter plot, target over time plot or both types
- you are also able to choose if the error information will be displayed. You can set different types of errors being used and the sorting criteria, which will be used to display the chosen models
- To sort the models by error, the balanced minimum error of training and validation is pre-set. You can also sort only by the best training or validation. You are also able to sort the models according to the overall error including the test data (global)



- in the 'Models' panel, you can select all reported models for which the plots should be created by enabling the respective box. If you want to deselect or select all models, simply click on the 'None' or 'All' button



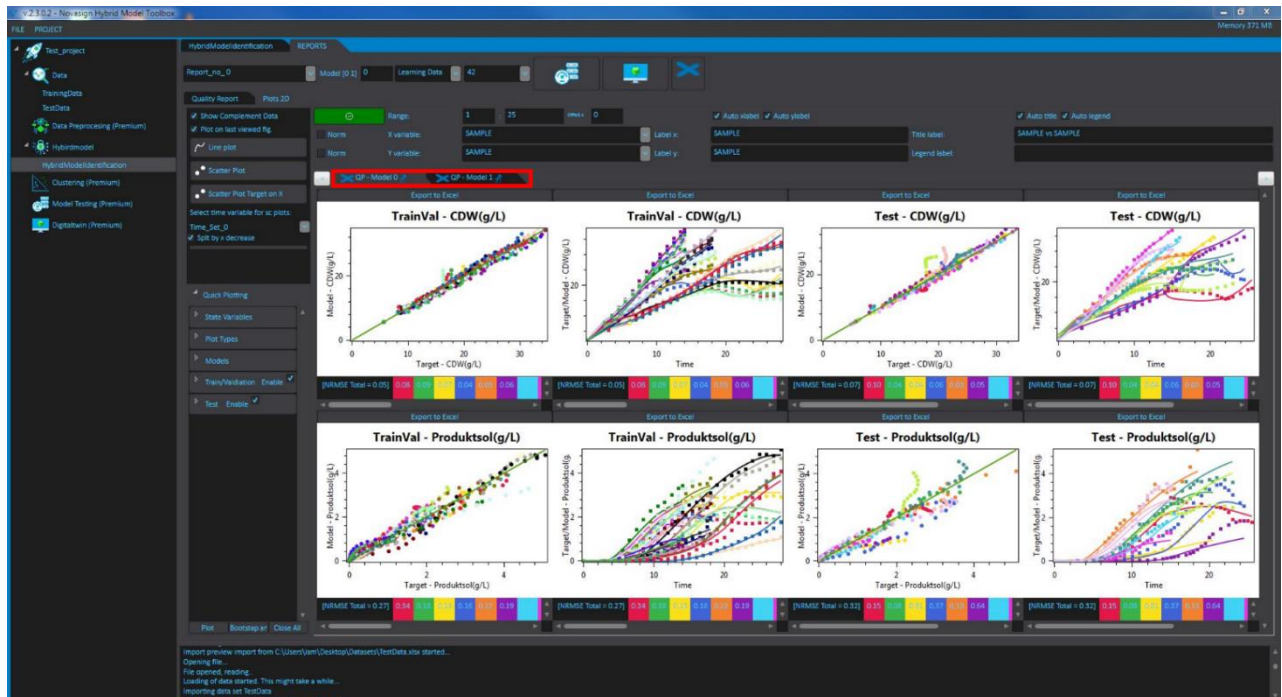
- the 'Train/Validation' panel enables you to choose if the train/validation plots should be displayed, either in general or choose only particular runs. The pre-set option is, that all models are plotted



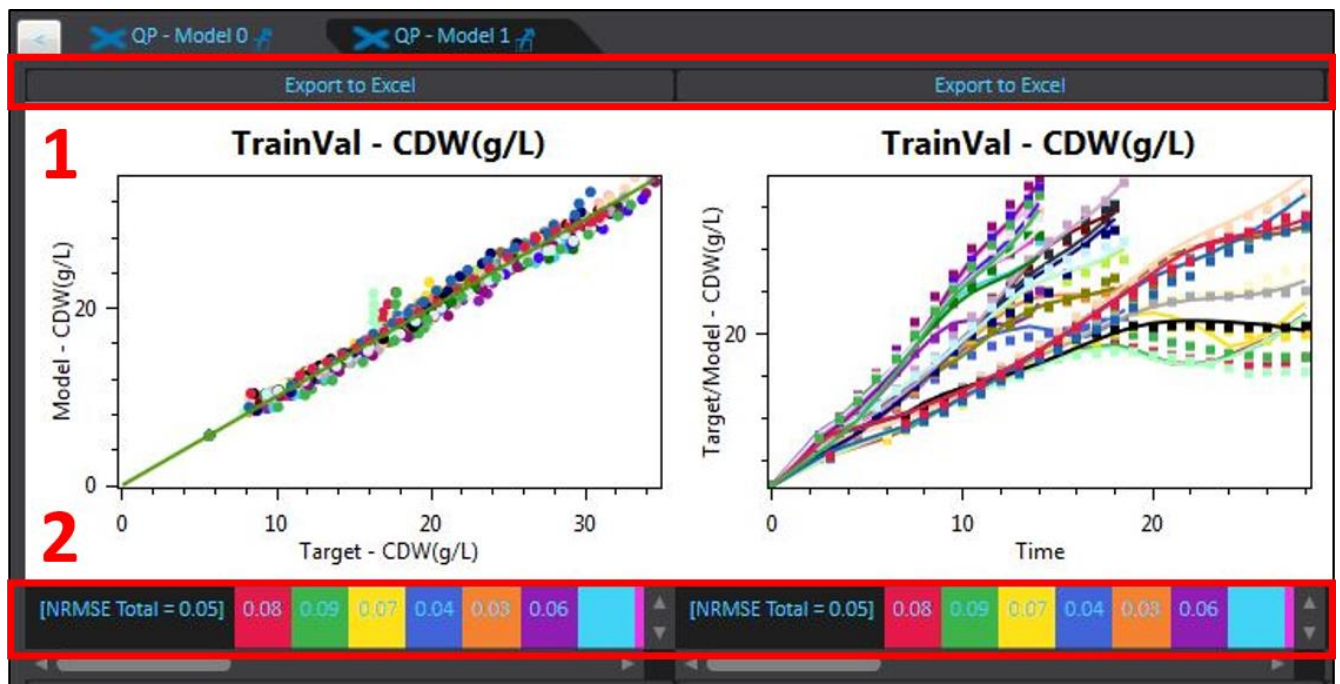
- likewise, the 'Test' panel enables you to choose if the test plots should be displayed, either in general or choose only particular runs. The pre-set option is, that all models are plotted



- if you choose the advanced plotting, define all plotting options and click on 'Plot', the tabs for the chosen models are displayed above the plots to switch between them
- on the left side of the plotting panel, you see the plots and error information on the training and validation data set. To the right, you can see how your model performs on the test data set. In this case, we have two targets to be modeled. Displayed are the 1:1 and output over time plots for each variable



- if you have a closer look at these created advanced plots, you see that each of these has the option to automatically export the herein included data to an excel sheet (1). Also, the respective chosen error is displayed on the bottom (2); the overall error and colored for each run separately



- if you want to bootstrap your models, choose the wanted models from the 'Model' panel (1). Therefore, only the wanted models have to be selected and the 'Bootstrap and plot' (2) button has to be used. The same plot structure as before is created, but the bootstrapped models including the confidence intervals are displayed. Again, you can export them to excel, including an additional column for the values of the confidence interval

