NED White Paper

August 11 2003

Table of Contents

1. Introduction

The goal of this white paper is to introduce you to the Akitisoft NED application. After going through this document you should gain the following:

- A general understanding of NED and what problem it solves.
- A good understanding of the data input structure.
- How data is entered in the web-based GUI.
- How to obtain, view and interpret results.

1.1 Background

In science and engineering, to accurately predict the behavior of a system over different ranges of various independent variables, a "good" mechanistic model is often needed. Theoretical models are preferable for the task since these are models based on theory governed by some underlining assumptions. Often these models are highly complex and nonlinear in parameters (i.e. the derivative of at least one parameter is still dependent on a parameter). Furthermore, for a single system under study several nonlinear models may be postulated. In such situations, experimental techniques such as factorial designs cannot be used to direct an experimental design program. Rather a statistical sequential design procedure based on information theory needs to be employed.

Several techniques are available for the task of designing experiments when nonlinear models are involved. One of the more popular classes of methods breaks the problem into two different parts; design for model discrimination and design for precise parameter estimation. In the first part, given an Operability Region (OR), a criterion is used to design runs to select the best of the various rival models. Once this is done, a separate criterion can be used to design runs to improve the precision of the estimated parameters. Employing such a sequential design strategy ensures that only optimal experimental conditions are chosen to be executed. As a result, the total number of runs required to arrive at the best answer is minimized. This in turn translates to savings in time and money.

Nonlinear Experimental Design (NED) is a simulation application for designing experiments in nonlinear situations. From an Operability Region (OR) supplied by the user, NED will select the best run conditions and the order in which they should be executed, to achieve model discrimination and/or precise parameter estimation in as few runs as possible.

2. Approach

2.1 The design criteria

NED uses the Joint Design Criterion (JDC) and a modified Initial-run Joint Design Criterion (I-JDC) to select runs. In either case, the best set of independent variables for a subsequent run is that which maximizes an objective function over the entire space of independent variables provided by the user. The overall criterion is:

$$C = w_1 D + w_2 E$$

The function D is a discrimination criterion, E is a parameter estimation criterion and w_1 and w_2 are weighting factors that determine the emphasis put on each objective. Incorporated in w_1 is a parameter λ which determines the rate of change of emphasis from one objective to the other. The main difference between the JDC method and the modified I-IJDC method is how initial experimental runs are chosen. With the JDC method, a set of preliminary runs, Np, are required before the method can be used to design new ones. Although this information can be useful it can also lead to erroneous results since the preliminary runs are often arbitrarily chosen. The I-JDC method solves this problem by using the criterion to start the design from the very first run.

NED is equipped with three design modes:

- Mode 1 is used to design runs for model discrimination only (i.e. $w_2 = 0$). At least two models must be involved in this design mode. NED has several model discrimination criteria to select from; they are listed in appendix A. The model probability, which is calculated after each designed run is performed and analyzed, is used as a measure of discrimination. NED has several probability density functions to select from for this purpose (see appendix B).
- Mode 2 is used to design runs for precise parameter estimation only (i.e. $w_1 = 0$). Only one model must be involved in this design mode. NED has only one criterion available for this purpose.
- Mode 3 is used to design runs for both model discrimination and parameter estimation simultaneously. All model discrimination criteria in appendix A and probability computation functions in appendix B are available to select from when using this design mode.

All modes can be used with supplied preliminary runs or with the preliminary runs designed by the criterion. Table 1 summaries this information.

Table 1 NED design modes

Method	Design mode	Type of design	# preliminary experiments	Weight factor W_1	Comments
1	1	Model discrimination	$N_p > 0$	$w_1 = 1$	Supply a set of preliminary runs prior to using the criterion to design further runs to achieve the objective of selecting a preferred model from several rivals.
2	1	Model discrimination	$N_p = 0$	$w_1 = 1$	Use the criterion form the first run onwards to design runs to achieve the objective of selecting a preferred model from several rivals.
3	2	Precise parameter estimation	$N_p > 0$	$w_1 = 0$	Supply a set of preliminary runs prior to using the criterion to design further runs to achieve the objective of estimating with precision the parameters in a selected model.
4	2	Precise parameter estimation	$N_p = 0$	$w_1 = 0$	Use the criterion form the first run onwards to design runs to achieve the objective of estimating with precision the parameters in a selected model
5	3	Joint Discrimination and precise	$N_p > 0$	$0 \leq w_1 \leq 1$	Supply a set of preliminary runs prior to using the criterion to

		parameter estimation			design further runs to achieve the dual objectives of selecting a preferred model from several rivals and estimating its parameters with precision.
6	3	Joint Discrimination and precise parameter estimation	$Np = 0$	$0 \leq w_1 \leq 1$	Use the criterion form the first run onwards to design runs to achieve the dual objectives of selecting a preferred model from several rivals and estimating its parameters with precision.

2.2 Data input

NED uses a single input file to specify a problem. The Akitisoft web application GUI in form mode will guide you through the specification of input data. However, in text mode, the data is to be entered in a text file in the structure recognized by the backend program. Knowing how to enter a problem in this manner also helps when trying to troubleshoot complex problems. In section, we will illustrate how data is entered in the input file. The input file must have the extension **.ned** for the program to recognize it as a input data file.

Problem title

The first entry in the data input file is the title. It may not exceed 60 characters including spaces and it should be followed by a semi-colon.

Example

> Test problem;

Random error generator seed values

NED uses a random number generator to add experimental error to the response as a means of simulating real experiments. The seed for starting the random number sequence consists of four odd integers between 0 and 4091. These numbers are entered on the next line followed by a semi-colon.

Example

> 13 87 19 3;

Data specification line

The third line in the input file is the specification line. The entries are integers and all are mandatory. A brief description, the range of values, and default values are summarized in Table 2. The first value is the discrimination criterion (1 – 6). The detailed descriptions of the various criteria can be found in Appendix B. The next value is the probability computation method (1 – 4). More detailed descriptions of each method can be found in Appendix C. Then, the number of candidate points is specified (1 – 500). Next, the number of prior experiments; they may take values from 0 to 20. If this number is greater than 0, then it means that preliminary runs data will be supplied later. The number of

independent variables (1 – 10) follows this. It is important to note that all the models under study should have the same number of independent variables. The next entry is the number of models under study (1 - 10). The user has the choice of using a constant or a non-constant error variance. This is indicated in the next entry with a 0 or 1 for a constant or a non-constant error variance respectively. NED has three design modes as discussed above. The desired mode is entered as the next entry (1 – 3). The number of required preliminary runs follows this. This value should be one more than the number of parameters in the model with the most parameters. This value, although mandatory, is only used with the initial-runs method. The next entry is the number of experimental runs to be designed (1 – 100). The user has the choice of generating random error or specifying the values. A 0 is entered if error is to be generated using the seed or a 1 is entered if the values are to be specified later. Note that the seed values should be entered even if the error values are to be specified. The model chosen as correct for simulation is entered next, (1 – m) where m is the number of models under study. Note that the models are 1 to m as they are entered. NED has four different nonlinear estimation methods. They are Gauss-Newton, Levenberg-Marquadt, Global and Simulated annealing. They are options 1 through 4 respectively.

Table 2 – Specification line data

Entry #	Description	Range	Default
1	Discrimination method	1 - 6	6
2	Probability computation method	1 - 4	4
3	Number of candidate points	1 - 500	---
4	Number of prior experiments	0 - 20	0
5	Number of independent variables	1 - 10	---
6	Number of models	1 - 10	---
7	Variance flag	0 – constant 1- constant	0
8	Operation modes	1-3	3
9	Number of preliminary runs	1- max param + 1	---
10	Number of runs to design	1 - 20	10
11	Random error flag	0 – generated 1 - specified	0
12	Correct model for simulation	1 – $m(m_{max}=10)$ models	1
13	Nonlinear estimation method	1 - 4	1

Example

| 1 | 4 | 8 | 0 | 1 | 2 | 0 | 3 | 2 | 20 | 0 | 1 | 1 ; |

Variable names

On the 4th line of the data file the names of the independent variables and the response are specified. Each may not exceed 4 characters in length and a semi-colon should follow the last entry.

Example

X y;

Operability Region (OR) candidate treatment points

The candidate points that make up the OR are specified on the lines following the variable names. They are real numbers and one treatment should appear on the same line followed by an index number. The index number is used to identify candidates in the outputs. Any treatments that have already been run should appear on the top of the list in the order in which they were run. If a candidate has been duplicated it should appear twice. The overall number of candidate point entries should be equal to the value entered in the specification line for the total number of candidate points, (third entry in the specification line). A semi-colon should follow the last index number.

Example

```
1.0     1  ⎫
0.5     2  ⎬        ←————————————————————    Preliminary runs
0.5     3  ⎭
0.0     4
0.25    5
0.50    6
0.75    7
1.00    8 ;
```

Previous responses

If there were any preliminary runs executed, the values of their responses, which are real numbers, are to be entered on the next few lines. One response is entered per line, and each response should correspond to the candidate point that was used to generate it. A semi-colon should follow the last entry. If no prior runs were conducted or are to be used (i.e. using the initial-runs method), this line should be omitted.

Example

```
0.97
0.52
0.48 ;
```

Variance values

On the next line the variance of the independent variables and the response are specified as real numbers. The variances of the independent variables are entered first followed by that of the response followed by a semi-colon. Remember that the independent variables are 1 to n as they appear in the models. If the variances on the independent variables are nil, 0.0 should still be entered.

Example

```
0.0              1.0e-1;
```

Number of parameters in rival models

The next line consists of integer values that indicate the number of parameters in each model. On the same line, the first entry is for Model 1 the second is for Model 2 and so on. A semi-colon should follow the last entry.

Example

```
1      1;
```

Model parameter guesses

The initial parameter guesses are specified as real numbers on the next few lines. For a given model, these guesses are on the same line from parameter 1 to p as the parameters appear in that model. This is done for each model 1 to m. A semi-colon should follow the last entry.

Example

```
2.0
```

5.0;

Model a priori probabilities

The next line is for the initial model probabilities. Although this line is mandatory, it is only used with the initial-runs method in modes 1 and 3. The values are entered as real numbers and should be on the same line from models 1 to *m*. A semi-colon should follow the last entry. The sum of the probabilities should equal unity.

Example

0.5   0.5;

Criteria weighting factors

The weights, w_1 and w_2, that determines how much emphasis is put on model discrimination and precise parameter estimation are entered as real numbers on the next line followed by a semi-colon. If a 0.0 is entered for the first entry then the second entry is λ. Although this line is mandatory it is only used in design mode 3.

Example

0.0   0.7;

Nonlinear estimation data

The next line is the estimation routine data specification line. There are seven entries on this line. There are four methods to select from; the Gauss-Newton method, the Levenberg-Marquardt method and two global optimization methods.

The Gauss-Newton nonlinear regression method

The first entry is an integer that determines the choice of estimation routine. To specify the Gauss-Newton method enter a 1. The next entry is the machine precision. The next value is the absolute tolerance for the step vector (relative to the vector of current estimates). The 4th value is the absolute tolerance for the Euclidean norm of the residual vector. The next entry is the maximum number of function evaluations. The sum of squares scale factor is next. The last entry is the maximum number of halvings of a calculated Newton step vector. Note that if an invalid value is entered for any of these variables the default value will be used. A semi-colon should follow the last entry. A brief description of each parameter along with their default values can be found in Table 3.

Table 3 - Gauss-Newton nonlinear regression routine descriptions and defaults

Entry #	Description	Default value
1	The method	1
2	Machine precision	1.0×10^{-5}
3	Relative tolerance for step vector	1.0×10^{-5}
4	Absolute tolerance for step vector	1.0×10^{-5}
5	Maximum number of function evaluations	1000
6	Sum of squares scaled factor	1.0
7	Maximum number of halvings of the Newton step	15.0

The Levenberg-Marquardt nonlinear regression method

The first entry is an integer that determines the choice of estimation routine. To specify the Levenberg-Marquardt method enter a 2. The next entry is the machine precision. The next value is the relative tolerance for the difference between the ultimate and penultimate residual vector. The 4th value is the absolute tolerance for the difference between the ultimate and penultimate residual vector. The next entry is the maximum number of function evaluations. The sum of squares scale factor is next. The last entry is the initial Marquardt parameter. Note that if an invalid value is entered for any of these variables the default value will be used. A semi-colon should follow the last entry. A brief description of each parameter along with their default values can be found in Table 4.

Table 4 - Levenberg-Marquardt nonlinear regression routine descriptions and defaults

Entry #	Description	Default value
1	The method	2
2	Machine precision	1.0×10^{-5}
3	Relative tolerance for step vector	1.0×10^{-5}
4	Absolute tolerance for step vector	1.0×10^{-5}
5	Maximum number of function evaluations	1000
6	Sum of squares scaled factor	1.0
7	Initial Marquadt parameter	0.001

The Global minimization via local search

The first entry is an integer that determines the choice of estimation routine. To specify the first global minimization routine enter 3. The next entry is the conversion criterion, number of significant digits. The next two entries are the lower and upper bounds of the parameter search. The next entry is the maximum number of function evaluations. The sum of squares scale factor. The last entry is the number of starting points selected for evaluation. Note that if an invalid value is entered for any of these variables the default value will be used. A semi-colon should follow the last entry. A brief description of each parameter along with their default values can be found in Table 5.

Table 5 – Global minimization routine descriptions and defaults

Entry #	Description	Default value
1	The method	3
2	Number of significant digits (convergence criterion)	5
3	Lower bound for parameter search	1.0×10^{-5}
4	Upper bound for parameter search	100000
5	Maximum number of function evaluations	10000
6	Sum of squares scaled factor	1.0
7	Number of starting points	15

The Global minimization via simulated annealing

The first entry is an integer that determines the choice of estimation routine. To specify the SA global minimization routine enter 4. The next entry is the conversion criterion, tolerance for termination. The next two entries are the temperature reduction factor and the initial temperature. The next entry is the maximum number of function evaluations. The sum of squares scale factor is next. The last entry is the number of iterations before temperature reduction. Note that if an invalid value is entered for any of these variables the default value will be used. A semi-colon should follow the last entry. A brief description of each parameter along with their default values can be found in Table 6.

Table 6 – SA Global minimization routine descriptions and defaults

Entry #	Description	Default value
1	The method	4
2	Tolerance for termination (convergence criterion)	1.0×10^{-5}
3	Temperature reduction factor	1.0×10^{-5}
4	Initial temperature	100
5	Maximum number of function evaluations	10000
6	Sum of squares scaled factor	1.0
7	Number of iterations before temperature reduction	50

An example of the parameter estimation specification line for Gauss is shown below.

Example

> 1 1.0×10^{-5} 1.0×10^{-5} 1.0×10^{-5} 1000   1.0   15

Random error values

If a 1 was entered for the random error flag, entry 11 in the data specification line, then the error values are to be specified on the next few lines. The values, real numbers, are entered in the order in which they were run and they should be on separate lines. The number of entries should not be less than the number of experiments to be designed. A semi-colon should follow the last entry.

Example

> 0.01
> 0.067
> 0.056
> 0.10;

Model Specification

The models under study are specified next. The number of models should equal the 6th value entered in the specification line. The "true" parameter values are specified in the models and are multiplied by the guesses during the design. Therefore a guess is a multiplier of a "true" parameter and should converge to unity as the estimates improve. For example a model,

$$y = k_1 \, x1^2$$

will be specified as follows;

> y=6.0*b1*x1^2;

In this case 6.0 is the "true" parameter value. It is used to simulate the response data. The independent variables may be x1, x2, …,xn or the values entered on line 4 of the input file. However the guesses must be b1, b2,..bn or else the program will generate an error (note the change from k_1 in the first equation to b1 in the second). A semi colon should follow each model. A table of available mathematical functions that may be used when specifying a model can be found in appendix D.

2.3 Data output

2.3.1 Text output
NED displays the results in three textural output files, Summary1, Summary2 and Output. All three files present a summary of the input data.

Summary1

The Summary1 output file presents, for each model and run; the run number, the weight on discrimination w_1, the best candidate point for the next run, the response with error, the model probability and the estimated parameters.

Summary2

In the Summary2 output file for each run and each candidate point the values of; the weight on discrimination w_1, the discrimination criterion D, the weight on parameter estimation w_2, the precise parameter estimation criterion E and the Initial-runs Joint Design Criterion C are presented. Then also after each run for each model; the observed response, the predicted response and the residual for all the runs up to and including the current run are displayed. And finally, also after each run, the confidence limits for each parameter in each model are presented.

Output

The Output output file presents the most information. First, for all the candidate points, the values of the G^TG determinant divided by the error variance are printed. Then the values of the discrimination criterion, the precise parameter estimation criterion and the Initial-runs Joint Design Criterion are presented for each candidate point. From these values the maximum value of C is printed, followed by the chosen candidate point to be used for the next run. The simulated response without error, the simulated experimental error and the simulated response with error are printed next in that order. Then for each model, the nonlinear regression convergence criteria met, the estimated parameters and the sum of squares of the residuals are displayed. Also printed are the observed response, the predicted response and the residuals for all the runs up to and including the current run. The G^TG determinant after the parameters have been modified is printed. The covariance matrix for that model follows this. After all the above information for each model is given, the variance of the predicted response and the G^TG determinants of each model are presented again together. This is followed by the probabilities of all the models for the current run.

2.3.2 Graphical output

Six plots can be generated.
- The discrimination plot shows the probability of all the models versus the run number.
- The weight on discrimination plot shows w_1 versus the run number.
- The selected candidates plot shows the best candidate chosen for each run.
- The random error plot shows the standard deviation of the generated error versus the run number.
- The response plot shows the observed response for a given model versus an independent variable, the predicted response or the run number.
- The residual plot displays the raw residuals for a given model versus an independent variable, the response or the run number.

3. Sample problem

In this section we will illustrate how NED is used to solve a nonlinear experimental design problem, by stepping through a simple reactor design problem.

3.1 Problem statement

Suppose you are asked to design an isothermal continuous stirred tank reactor (CSTR) in which the reaction

$$A \rightarrow B$$
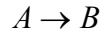
is to take place. The volume of the reactor is given by,

$$V = \frac{\upsilon_o C_{Ao}}{-r_A}$$

where υ_0 is the volumetric flow rate, C_{A0} is the initial concentration of species A and r_A is the rate of reaction. In other to predict the reaction rate a rate mechanism is required. Suppose that based on different theoretical assumptions the following two mechanisms were postulated,

Model 1
$$r_A = kC_A$$
and

Model 2
$$r_A = kC_A^2$$

Now the task is to determine which of these two mechanisms best describes the above reaction. Assume that there are no prior experimental runs available for this problem.

3.2 Solution

To solve our problem, we will use NED to design experimental runs form the first onwards to discriminate between these two models and estimate with precision the parameter in the best model. This is method 6 in Table 1.

3.2.1 Problem specification in Akitisoft

Since NED is a web-based application, your first step is to connect to the Internet. Once online, access the Akitisoft page. Next, login to your account. If you are not already a member, sign up by following the instructions after clicking the Open an Account link to the left of the web page. Then click on Applications and select NED. The NED start page will appear. To specify a NED problem you will step through 5 input pages.

Page 1

After selecting **New** under the NED logo, the first input page will appear; it has 14 entries.

- The first entry is the title. We will call our problem **Rate model determination**.
- The next set of entries are the seed values for starting the random error sequence. The default values presented should work just fine.
- The next entry is the discrimination method. We will use the **Box-Hill** criterion. It is one of the more popular criteria for model discrimination (see Appendix B for more details).
- The next entry is the probability computation method. We will use the **Stewart-Box-Hensen** method.

- The number of candidate treatment points in the Operability Region (OR) is entered next. In our problem we will specify concentrations values from 0 to 1 at increments of 0.25, hence we have **5** candidate points.
- The next value is the number of prior experiments. Since we are designing runs form the first one onwards, we will enter **0**.
- The next entry is the number of independent variables. In our case we have **1**, the concentration of A.
- Next the number of models under investigation must be specified. We have **2** proposed models.
- Next we must indicate whether the variance is constant or non-constant. In our case we will assume a **constant** variance.
- The next entry is the design mode. We are designing for **discrimination and estimation**.
- The next entry is number of preliminary runs. Since all models have one parameter, the value is **2**.
- The number of experimental runs to design is entered next. We will design **10** runs.
- In the next entry we must specify if the random error is generated using the built in random number generator or if the values will be specified. We will use program **generated** values.
- The last entry is for the model from which experimental response data will be generated. We will use model 1. Hence at the end of our simulation we should arrive back at this model.

When you are done hit submit and the next page will appear.

Page 2

- The first entry section on the second input page is for the names of the independent variables and the response. There are default values, however we will use others. We will specify the concentration as **Ca** and the reaction rate as **ra**.
- The next section is for the candidate points in the OR. This section will resemble Table 7

Table 7 Operability region

Variable 1	INDEX
0.00	1
0.25	2
0.50	3
0.75	4
1.00	5

Hit submit when done.

Page 3

- The first set of entries on the third input page are for the variances. Since we are using a constant variance, the variance on the independent variable will be **0.00**. We will assume **0.001** as the variance on the response.
- The next entry is for he number of parameters in each model. In our case each model has only **1** parameter.

Hit submit when done.

Page 4

- The first entry on page 4 is for the initial parameter guesses in each model. We will accept the default values.

- The second set of entries are the initial (a priori) probabilities for each model. Here too we will accept the defaults.
- The next set of entries is for the initial weights on each objective (discrimination and estimation). Again we will accept the defaults.
- The last entry of the page is the estimation method used to fit each model to the experimental data. Since this is a very simple problem, the Gauss-Newton method should suffice.

Hit submit when done.

Page 5

- The first set of entries on the last entry page, are the estimation parameters for the chosen routine, in our case Gauss-Newton. The defaults presented should be adequate.
- Next the models under consideration are entered. In our case ra=b1*Ca is entered for the first model and ra=b1*Ca^2 for the second. Note that our true rate constant value is taken as 1, hence no multiplier is necessary for our true model (model 1).

Click Run when you are done.

3.2.2 Results

If the run was successfully there will be 5 options to choose from.
- You may save your complete input data file (note you may save files with partial input information from page 2 onwards).
- You may view the summary1 text results.
- You may view the summary2 text results.
- You may view the output text results.
- You may view graphical results.

Going through the various results pages is very intuitive. For our exercise, the results show that, as expected Model 1 is select as the better of the two. The model is selected in the minimum number of runs and the parameter estimates improve over the 10 designed runs.

The example presented here is very simple in that the models are not nonlinear in parameters, but it serves as good illustration as to how NED can be of great value. Usually, for a given reaction, many more mechanisms are postulated and some if not all are nonlinear in parameters. The use of NED can greatly reduce the time and cost of experimentation to arrive at a suitable model. NED is not restricted to kinetic modeling, it can be used in any discipline where the task is to design experimental runs to find the best suitable model, form many alternatives, that describes a phenomenon.

4. Conclusion

NED is one of the applications on the Akitisoft web application site. Use NED when your goal is to design experiments when nonlinear models are involved. As you have seen here, NED is an extremely powerful application that is simple to use . With most analysis packages today, to get started you need to have or download and install a copy of the software. This may bring up the question of disk space. With the growing number of applications on your computer you may need extra space to house the new application. This can be costly. Also, how easily can you and your colleague verify your results? Unless the other party has a similar application with the same settings, verifying results can be difficult. With Akitisoft you and your colleagues, worlds apart, can solve the same problem and compare results in near time. Furthermore, how steep is the learning curve? How much time and money are you willing to spend on training? With the Akitisoft intuitive web-based application GUI almost no training is required.

Appendix A Terminology

Empirical models – These models are limited in that they may predict a response only over the range of independent variables chosen by the experimenter. Extrapolations may predict incorrect responses.

Model discrimination – Favoring a particular model over several alternatives.

Nonlinear model – A model that, in its original form, the derivative with respect to a particular parameter still depends on a parameter in the model. Thus, in nonlinear experimental design, the values of the parameters affect the design at all stages.

Operability region – A space of independent variables from which a particular run condition may be chosen to be performed.

Precise parameter estimation – Evaluating the parameters, with confidence intervals, in a model.

Sequential Design – In sequential design, the next run condition is selected based on a well-defined rule that uses information from previous runs.

Theoretical model – Attempts to describe a phenomenon as closely as possible using theoretical means. A theoretical model is the result of a better understanding of the physical mechanism of the phenomenon. Such models make it possible to design more efficient systems.

Appendix B Available discrimination criteria

List of discrimination criteria used in NED

NED No.	Year	Authors	Criterion
1	1967	Box and Hill	$$D_0 = \frac{1}{2} \sum_{i-1}^{m-1} \sum_{j=i+1}^{m} \Pi_{i,n} \Pi_{i,n} \times \left\{ \left[\frac{(\sigma_i^2 - \sigma_j^2)^2}{(\sigma^2 + \sigma_i^2)(\sigma^2 + \sigma_j^2)} + (\hat{y}_{i,n+1} - \hat{y}_{j,n+1})^2 \right] \times \left(\frac{1}{\sigma^2 + \sigma_i^2} + \frac{1}{\sigma^2 + \sigma_j^2} \right) \right\}$$
2	1971	Hsiang and Reilly	$$C_0 = \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} \left| \hat{y}_{i,n+1} - \hat{y}_{j,n+1} \right| \left[\Pi_{i,n} + \Pi_{j,n} \right]$$
3	1983	Buzzi-Ferraris and Forzatti	$$T_0 = \frac{\sum_{i=1}^{m-1} \sum_{j=i+1}^{m} (\hat{y}_{i,n+1} - \hat{y}_{j,n+1})^2}{(m-1)(m\sigma^2 + \sum_{i=1}^{m} \sigma_i^2)}$$
4	1965	Roth	$$Z_0 = \sum_{i=1}^{m} \left[\Pi_{i,n} \prod_{\substack{j=1 \\ j \neq i}}^{m} \left| \hat{y}_{j,n+1} - \hat{y}_{i,n+1} \right| \right]$$
5	1965	Hunter and Reiner	$$S_0 = \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} (\hat{y}_{i,n+1} - \hat{y}_{j,n+1})^2$$
6	---	Akiti and Eckert	$$A_o = \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} \left| \hat{y}_{i,n+1} - \hat{y}_{j,n+1} \right|$$

Appendix C Available probability computation functions

List of probability density functions used for probability computations in NED

NED No.	Year	Authors	Density function
1	1966	Hill	$p(M_i \mid y)\, \alpha\, \exp\left[-\dfrac{1}{2\sigma^2} S_{i,n}\right]$
2	1967	Box and Hill	$p(M_i \mid y)\, \alpha\, \exp\left[-\dfrac{1}{2(\sigma^2 + \sigma_i^2)} S_{i,n}\right]$
3	1969	Box and Henson	$p(M_i \mid y)\, \alpha\, \exp\left[-\dfrac{1}{2\sigma^2}(S_{i,n} + p_i\sigma^2)\right]\left[\dfrac{G_{i,p}^T G_{i,p}}{G_{i,n}^T G_{i,n}}\right]^{\frac{1}{2}}$
4	1996	Stewart et al.	$p(M_i \mid y)\, \alpha\, 2^{-\frac{p_i}{2}} \exp\left[-\dfrac{1}{2\sigma^2} S_{i,n}\right]$
5	1996	Stewart et al. (soon)	$p(M_i \mid y)\, \alpha\, 2^{-\frac{p_i}{2}} S_{i,n}^{-\nu_e/2}$

Appendix D Available mathematical functions

Mathematical functions

Math function	Usage
$\cos^{-1}(x)$	acos (x)
$\sin^{-1}(x)$	asin (x)
$\tan^{-1}(x)$	atan (x)
$\cos(x)$	cos (x)
$\sin(x)$	sin (x)
$\tan(x)$	tan (x)
$\cosh(x)$	cosh (x)
$\sinh(x)$	sinh (x)
e^x	exp (x)
$\ln(x)$	log (x)
$\log_{10}x$	log10 (x)
y^x	y ^ x
$\sqrt{x}$	sqrt (x)
$|x|$	abs (x)
$\int_{a}^{b} f(x)dx$	I(xn,f(xn),a,b)
$\min(f(x),a,b)$	z(xn,f(xn),a,b)