

Project	Document	Filename	Version	Date
FOCUS-2.3	User Guide	FOCUS231 User Guide.doc	1	26/10/2007

FOCUS-2.3.1 - USER GUIDE

Autor: Dr. Lutz Tischendorf
ELUTIS Modelling and Consulting Inc.

Scientific Concept: Jeff Holland & Dan Bert
Implementation: Lutz Tischendorf

1	Introduction.....	2
2	Sampling schemas.....	2
2.1	Plot-Buffer related sampling.....	2
2.2	Plot-Plot related sampling.....	3
3	Configuration File.....	5
3.1	Example configuration for Plot-Buffer sampling.....	6
3.2	Example configuration for Plot-Plot sampling.....	7
4	Input file formats (Plot-Buffer).....	8
4.1	Plot_Distances.....	8
4.2	Predictor_Variables.....	8
4.3	Response_Variables.....	8
5	Input file formats (Plot-Plot).....	9
5.1	Predictor_Variables.....	9
5.2	Response_Variables.....	9
6	Results.....	9
6.1	Raw Data.....	9
6.2	Aggregated Data.....	10
6.3	Plot Data (Plot-Buffer).....	11
6.4	Plot Data (Plot-Plot).....	12
7	Logging.....	12
8	System Requirements.....	12
9	License.....	12
10	Running FOCUS-2.3.....	12
11	Trouble Shooting.....	13
12	Considerations.....	13
12.1	Duplicate Samples.....	13
12.2	Unique Samples.....	13
12.3	Small sample sizes.....	13
13	Change Log.....	14
13.1	Focus 2.3 -> Focus 2.3.1.....	14
13.2	Focus 2.2 -> Focus 2.3.....	14
14	References.....	14

Author	Company	Page
Lutz Tischendorf	ELUTIS – Modelling and Consulting Inc. 681 Melbourne Avenue, Ottawa, Ontario, K2A 1X4, CANADA	1 of 14

Project	Document	Filename	Version	Date
FOCUS-2.3	User Guide	FOCUS231 User Guide.doc	1	26/10/2007

Introduction

FOCUS-2.3 conducts statistical analyses on random samples of spatial data sets. The spatial data sets are comprised of plots, which relate predictor and response variables to space. Sampling is done based on distance constraints between plots. FOCUS-2.3 can sample either plots, which are closer to each other than a specified distance constraint, or plots, which are farther apart from each other than a specified distance constraint. After sampling, FOCUS-2.3 calculates linear regression coefficients and conducts an analysis of variance. The results are reported in comma separated files, which can be further processed in spreadsheet programs like MS Excel. In addition to the linear regression, FOCUS-2.3 can test for the significance of polynomial regression models up to the 5th order, i.e. test for the functional stability of the initial linear regression model. FOCUS-2.3 may also be used for sampling only and can report the sampled variables in a separate file, which can be processed further in spreadsheet or statistical software packages.

1 Sampling schemas

FOCUS-2.3 can sample plot data from two conceptually different schemas, which are referred to as “Plot-Plot” and “Plot-Buffer” hereafter.

1.1 Plot-Buffer related sampling

In this situation, the response variable is measured at the location of a plot and the predictor variable is measured within a spatial context, e.g. buffer or ring, surrounding a plot. FOCUS-2.3 samples response and predictor variables based on the relationship between a chosen distance constraint and both, the inter-plot distance and the corresponding buffer radius.

For an ‘*outside*’ distance constraint only those response variables (plots) will be selected, whose inter-plot distances are equal or larger than the distance of the constraint. For example, if the distance is larger than the outer buffer radius in Figure 1 and Plot 1 is selected first, then also Plot 4 and Plot 6 will be selected for the statistical analysis. For an ‘*inside*’ distance constraint only those plots will be selected, whose inter-plot distances are equal or smaller than the distance of the constraint.

The predictor variables of the selected plots are chosen according to the relationship between the buffer radius and the distance constraint. For an ‘*outside*’ distance constraint buffers with the nearest smaller radius to the distance of the constraint will be selected. For instance, if the distance constraint is larger than the outer buffer radius in Figure 1, Predictor Variable C would be selected for the statistical analysis. For an ‘*inside*’ distance constraint, the buffer with the nearest larger radius to the distance of the constraint would be selected.

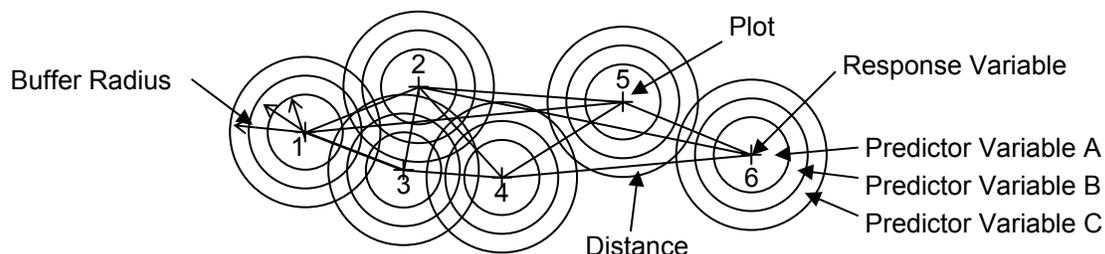


Figure 1: Plot-Buffer related sampling schema

Author	Company	Page
Lutz Tischendorf	ELUTIS – Modelling and Consulting Inc. 681 Melbourne Avenue, Ottawa, Ontario, K2A 1X4, CANADA	2 of 14

Project	Document	Filename	Version	Date
FOCUS-2.3	User Guide	FOCUS231 User Guide.doc	1	26/10/2007

1.2 Plot-Plot related sampling

In Plot-Plot related sampling, predictor and response variables are related to the location of plots. This is conceptually different from Plot-Buffer related sampling, because there is no implicit spatial relationship between the response and predictor variables via the spatial context of a plot. Predictor and Response variables are both spatially independent.

In Plot-Plot related sampling, both predictor and response variable sets are sampled according to distance constraints. For an *'outside'* distance constraint plots are selected, whose inter-plot distances are equal or larger than the distance of the constraint. For an *'inside'* distance constraint plots are selected, whose inter-plot distances are equal or smaller than the distance of the constraint. Note that FOCUS-2.3 will apply the same distance constraints to sample both predictor and response variables. In other words, each distance constraint as defined in the configuration file will be used to sample plots from the predictor variable plot set and from the response variable plot set. If the sampled subsets from both plot sets are different in size, the larger subset will be reduced to the size of the smaller subset so that the number of the predictor and response variables is always identical.

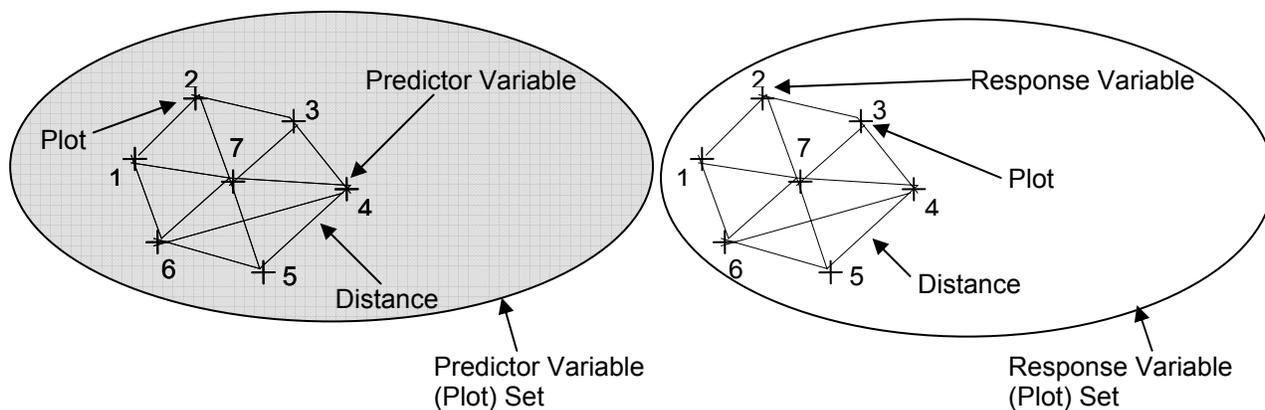


Figure 2: Plot-Plot related sampling schema. Both variable (plot) sets are spatially independent. Predictor and response variables are measured at the plot's locations. Plot locations can also be cells of a grid.

Author	Company	Page
Lutz Tischendorf	ELUTIS – Modelling and Consulting Inc. 681 Melbourne Avenue, Ottawa, Ontario, K2A 1X4, CANADA	3 of 14

Project	Document	Filename	Version	Date
FOCUS-2.3	User Guide	FOCUS231 User Guide.doc	1	26/10/2007

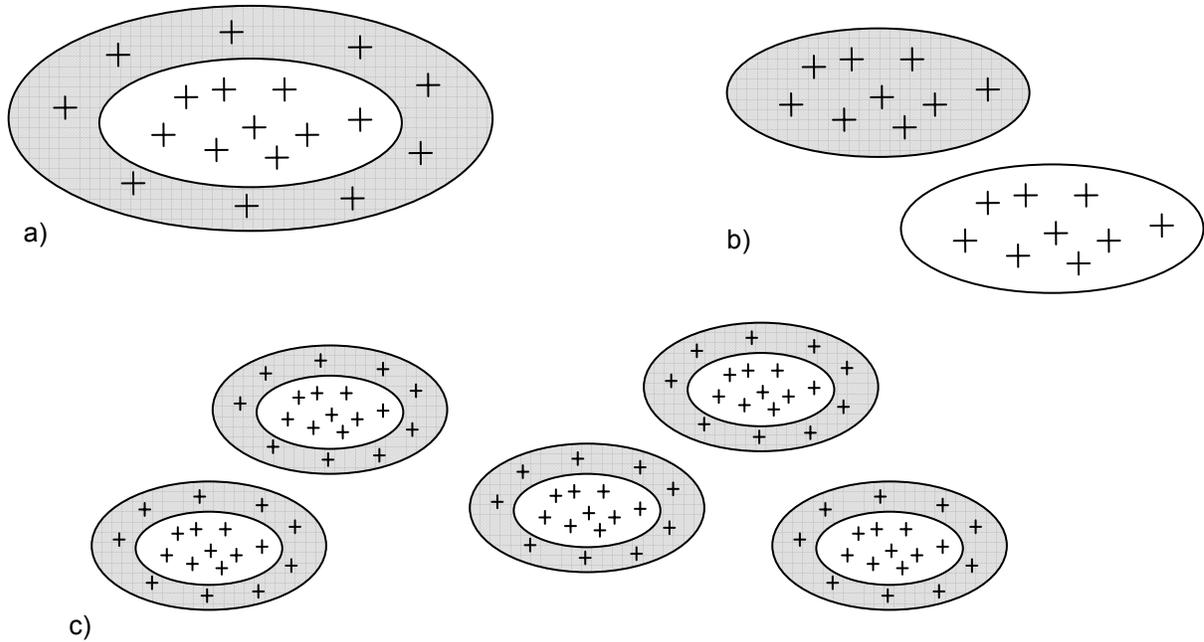


Figure 3: Potential spatial configurations of the independent variable (plot) sets. a) response variables are surrounded by predictor variables. For example, the response variables might be the species diversity within a park, whereas the predictor variables are the road densities surrounding a park. b) response and predictor variables are located side by side. c) multiple sets of response and predictor variables, e.g. for cross-park or global (with respect to a set of selected parks) analysis of the effects of a predictor on a response variable.

Project	Document	Filename	Version	Date
FOCUS-2.3	User Guide	FOCUS231 User Guide.doc	1	26/10/2007

2 Configuration File

FOCUS-2.3 reads a configuration file. This configuration file must be in a specific xml format as shown in the examples below. Xml stands for 'extensible markup language' and allows to structure information by means of tags. The configuration file provides all necessary information for FOCUS-2.3 and allows for controlling the execution of the program. One can either provide FOCUS-2.3 with the name of the configuration file as command line parameter (e.g. FOCUS23.exe config.xml). If this file name is not provided, FOCUS-2.3 will try to read the default FocusConfig.xml file.

tag name	valid range	comments
Number_Of_Iterations	1.. 32767	iterations/runs for each constraint (distance)
Maximum_Plots_Per_Iteration	5.. 32767	maximum number of plots to be sampled during each iteration
Report_Zero_Regressions	yes/no	yes – will report all regressions no – will omit regressions with zero only samples
DuplicateSamples		Considered only if UniqueSamples are not enforced
Allow	yes/no	yes – allows duplicate sets of sampled plots within all iterations for one constraint (distance) no – forbids duplicate sets of sampled plots within all iterations for one constraint (distance)
Retrials	1.. 2147483647	maximum number of retrials to find a new unique set of plots, if FOCUS-2.3 cannot find a new unique set of plots after the maximum number of retrials, the program will be terminated.
UniqueSamples		Has higher priority than DuplicateSamples
Enforce	yes/no	yes – enforce unique/distinct sets of sampled plots within all iterations for one constraint (distance) no – allow sets of sampled plots with common plots within all iterations for one constraint (distance)
FunctionalStability		
Enable	yes/no	yes – enables higher order polynomial regression (functional stability test) no – disables functional stability test
Maximum_Order	2..5	maximum order for the functional stability test
Percentage_Of_Runs	1..100	percentage of runs/iterations for which the functional stability test will be conducted
Statistics	yes/no	yes – enables statistical analysis no – disables statistical analysis, i.e. sampling only
WritePlots	yes/no	yes – write sampled plots (id, predictor, response) to a file no – no plots will be written to a file
PlotSet		
Type	Plot-Plot / Plot-Buffer	Plot-Plot – sampling according to Plot-Plot sampling schema Plot-Buffer – sampling according to Plot-Buffer schema
InputFiles		
Plot_Distances	valid file name	csv file for the plot distance matrix
Predictor_Variables	valid file name	csv file for the predictor variables
Response_Variables	valid file name	csv file for the response variables
OutputFiles		
BaseName	valid file name	base name for the output files, up to 3 files are written
TimeStamp	yes/no	yes – adds time stamp to the output file names no – adds no time stamp to the output file names
Distance		
Value	1..100000000	distance constraint value, many values can be provided here, see example configuration file
Multiplier	1..1000	global multiplier for all distance values, i.e. the distance values will be multiplied with this number
SampleType	inside/outside	inside – plots, which are closer than the distance values are sampled outside – plots, which are farther apart than the distance values are sampled

Note that xml files are ASCII files and can be edited with any text editor. On some computers xml files may be opened automatically by a particular program (e.g. Internet Explorer) due to a predefined association of the xml extension. This may result in difficulties to open the xml files with an ASCII editor, such as NotePad or WordPad. One may either rename the configuration file with another extension or

Author	Company	Page
Lutz Tischendorf	ELUTIS – Modelling and Consulting Inc. 681 Melbourne Avenue, Ottawa, Ontario, K2A 1X4, CANADA	5 of 14

Project	Document	Filename	Version	Date
FOCUS-2.3	User Guide	FOCUS231 User Guide.doc	1	26/10/2007

open the program from within an ASCII editor. The xml extension is not mandatory for FOCUS-2.3. The following table explains each parameter/tag of the configuration file.

2.1 Example configuration for Plot-Buffer sampling

According to the settings in this example configuration file, FOCUS-2.3 will conduct a Plot-Buffer related sampling. It will read 3 input files (Distance_Matrix.csv, Predictor_Context_Variables.csv, Response_Variables.csv). For each constraint (i.e. distances specified within the Distance tag), FOCUS-2.3 will sample plots, which are farther apart than the specified distances (SampleType = outside). Distances will not be multiplied, because the Multiplier tag is set to 1. For each constraint, FOCUS-2.3 will conduct 100 cycles (iterations/runs) of sampling with a maximum of 10 sampled plots. Duplicates are allowed during the 100 iterations. (The tag 'Retrials' is ignored when 'Allow' is set to 'yes').

```

<Focus>
  <Configuration>
    <Number_Of_Iterations>100</Number_Of_Iterations>
    <Maximum_Plots_Per_Iteration>10</Maximum_Plots_Per_Iteration>
    <Report_Zero_Regressions>no</Report_Zero_Regressions>
    <DuplicateSamples>
      <Allow>yes</Allow>
      <Retrials>100</Retrials>
    </DuplicateSamples>
    <UniqueSamples>
      <Enforce>yes</Enforce>
    </UniqueSamples>
    <FunctionalStability>
      <Enable>yes</Enable>
      <Maximum_Order>5</Maximum_Order>
      <Percentage_Of_Runs>50</Percentage_Of_Runs>
    </FunctionalStability>
    <Statistics>yes</Statistics>
    <WritePlots>yes</WritePlots>
  </Configuration>
  <PlotSet>
    <Type>Plot-Buffer</Type>
    <InputFiles>
      <Plot_Distances>Distance_Matrix.csv</Plot_Distances>
      <Predictor_Variables>Predictor_Context_Variables.csv</Predictor_Variables>
      <Response_Variables>Response_Variables.csv</Response_Variables>
    </InputFiles>
    <OutputFiles>
      <BaseName>TestRun</BaseName>
      <TimeStamp>no</TimeStamp>
    </OutputFiles>
    <Constraints>
      <Distance>
        <Value>600</Value>
        <Value>900</Value>
        <Value>1200</Value>
        <Value>1500</Value>
      </Distance>
      <Multiplier>1</Multiplier>
      <SampleType>outside</SampleType>
    </Constraints>
  </PlotSet>
</Focus>

```

Zero regressions (i.e. regressions where all response variables are zero) will not be reported. FOCUS-2.3 will conduct a functional stability test for all sampled plots up to the order (of the polynomial regression) of 5. The sampled plots are recorded and written to a file. All results are written to 3 output files with the BaseName of 'TestRun' (TestRun_raw.csv, TestRun_aggr.csv, TestRun_plots.csv) without a time stamp

Author	Company	Page
Lutz Tischendorf	ELUTIS – Modelling and Consulting Inc. 681 Melbourne Avenue, Ottawa, Ontario, K2A 1X4, CANADA	6 of 14

Project	Document	Filename	Version	Date
FOCUS-2.3	User Guide	FOCUS231 User Guide.doc	1	26/10/2007

in the file names. (Note, existing output files are detected by FOCUS-2.3 and require the users explicit permission to be overwritten).

2.2 Example configuration for Plot-Plot sampling

According to the settings in this example configuration file, FOCUS-2.3 will conduct a Plot-Plot related sampling. It will read the 2 input files PlotPlotPredictor.csv and PlotPlotResponse.csv. (The 'Plot_Distances' file will not be read when 'Type' is set to 'Plot-Plot'. For each constraint (i.e. distances specified within the Distance tag), FOCUS-2.3 will sample predictor and response variable plots, which are farther apart than the specified distances (SampleType = outside). Distances will be multiplied by 2, because the Multiplier tag is set to 2. For each constraint, FOCUS-2.3 will conduct 50 cycles (iterations/runs) of sampling with a maximum of 100 sampled plots.

```

<Focus>
  <Configuration>
    <Number_Of_Iterations>50</Number_Of_Iterations>
    <Maximum_Plots_Per_Iteration>100</Maximum_Plots_Per_Iteration>
    <Report_Zero_Regressions>yes</Report_Zero_Regressions>
    <DuplicateSamples>
      <Allow>no</Allow>
      <Retrials>100</Retrials>
    </DuplicateSamples>
    <UniqueSamples>
      <Enforce>no</Enforce>
    </UniqueSamples>
    <FunctionalStability>
      <Enable>yes</Enable>
      <Maximum_Order>5</Maximum_Order>
      <Percentage_Of_Runs>50</Percentage_Of_Runs>
    </FunctionalStability>
    <Statistics>yes</Statistics>
    <WritePlots>yes</WritePlots>
  </Configuration>
  <PlotSet>
    <Type>Plot-Plot</Type>
    <InputFiles>
      <Plot_Distances>Distance_Matrix.csv</Plot_Distances>
      <Predictor_Variables>PlotPlotPredictor.csv</Predictor_Variables>
      <Response_Variables>PlotPlotResponse.csv</Response_Variables>
    </InputFiles>
    <OutputFiles>
      <BaseName>PP_Test</BaseName>
      <TimeStamp>no</TimeStamp>
    </OutputFiles>
    <Constraints>
      <Distance>
        <Value>300</Value>
        <Value>600</Value>
        <Value>900</Value>
      </Distance>
      <Multiplier>2</Multiplier>
      <SampleType>outside</SampleType>
    </Constraints>
  </PlotSet>
</Focus>

```

Duplicates are not allowed during the 50 iterations and FOCUS-2.3 will try up to 100 times to find a new unique set of plots. (If FOCUS-2.3 cannot find another unique sample, the program will be terminated) Zero regressions (i.e. regressions where all response variables are zero) will be reported. FOCUS-2.3 will conduct a functional stability test for 50 percent of all sampled plots up to the order (of the polynomial regression) of 5. The sampled plots are recorded and written to a file. All results are written to 3 output

Author	Company	Page
Lutz Tischendorf	ELUTIS – Modelling and Consulting Inc. 681 Melbourne Avenue, Ottawa, Ontario, K2A 1X4, CANADA	7 of 14

Project	Document	Filename	Version	Date
FOCUS-2.3	User Guide	FOCUS231 User Guide.doc	1	26/10/2007

files with the BaseName of 'PP_Test' (PP_Test_raw.csv, PP_Test_aggr.csv, PP_Test_plots.csv) without a time stamp in the file names. (Note, existing output files are detected by FOCUS-2.3 and require the users explicit permission to be overwritten).

3 Input file formats (Plot-Buffer)

3.1 Plot_Distances

For Plot-Buffer sampling, FOCUS-2.3 expects a distance matrix for all plots. The file must be a comma separated file (permitted delimiters are comma, semicolon and white space). The first row and the first column must contain numerical plot ids. The following is an example for a valid distance matrix for 6 plots.

```
Plot_id,101,102,103,104,105,106
101,0.98.95,249.13,319.68,374.3,358.72
102,98.95,0,154.24,266.49,362.92,393.32
103,249.13,154.24,0,191.62,342.56,438.07
104,319.68,266.49,191.62,0,165.76,307.7
105,374.3,362.92,342.56,165.76,0,174.4
106,358.72,393.32,438.07,307.7,174.4,0
```

3.2 Predictor_Variables

In Plot-Buffer sampling the predictor variables are tied to a plot's spatial context, i.e. a circular buffer with a certain buffer radius (see section 2.1). FOCUS-2.3 expects a comma separated (permitted delimiters are comma, semicolon and white space) predictor variable file of the following format. The first column must contain numerical plot ids. The first row must contain the buffer radiuses resulting in a matrix of predictor variable values for each plot id and each buffer radius. The following is an example for a valid predictor variable file for 6 plots and 6 buffer radiuses.

```
Plotid,20,40,60,80,100,120
101,0.9936407,0.829278022,0.640827562,0.498975973,0.41592043,0.368230566
102,0.9936407,0.99621061,0.999023264,0.877766122,0.747729308,0.660390346
103,0.990461049,0.828480255,0.726602735,0.671761826,0.638096456,0.615160868
104,0.9936407,0.94256083,0.819836619,0.745591688,0.663042088,0.611097542
105,0.994435612,0.998603909,0.999023264,0.988111294,0.924811309,0.852765504
106,0.992845787,0.997606701,0.999556029,0.940656376,0.870634515,0.81377534
```

3.3 Response_Variables

Response variables for Plot-Buffer related sampling have to be provided in a comma separated file (permitted delimiters are comma, semicolon and white space). The first column must contain numerical plot ids and the second column the values of the response variables. The following is an example for a valid response variable file for 6 plots.

```
101,4
102,1
103,2
104,0
105,0
106,0
```

Author	Company	Page
Lutz Tischendorf	ELUTIS – Modelling and Consulting Inc. 681 Melbourne Avenue, Ottawa, Ontario, K2A 1X4, CANADA	8 of 14

Project	Document	Filename	Version	Date
FOCUS-2.3	User Guide	FOCUS231 User Guide.doc	1	26/10/2007

4 Input file formats (Plot-Plot)

4.1 Predictor_Variables

For Plot-Plot sampling, the predictor variable file for FOCUS-2.3 must comply with the following format. The first column must contain numerical plot ids, the second column the X coordinates and the third column the Y coordinates of the plots locations. The fourth column must contain the values of the predictor variables. All values must be separated by comma, semicolon or white space. The following is an example for a valid predictor variables file for 6 plots.

```
PLOT_ID,X-COORD,Y-COORD,Predictor
300,8581.804865,7904.110843,50.17619926
301,9699.148534,9797.112949,57.0088919
302,7776.20777,8359.782708,68.8810492
303,8329.935606,9730.643635,52.3400918
304,8595.629749,7590.258492,50.43668991
305,8672.994171,9147.709586,53.61285005
```

4.2 Response_Variables

For Plot-Plot sampling, the response variable file for FOCUS-2.3 must comply with the following format. The first column must contain numerical plot ids, the second column the X coordinates and the third column the Y coordinates of the plots locations. The fourth column must contain the values of the response variables. All values must be separated by comma, semicolon or white space. The following is an example for a valid response variables file for 6 plots.

```
PLOT_ID,X-COORD,Y-COORD,Response
1,3202.459792,2897.274697,8
2,415.5095065,856.0441908,6
3,3110.904263,4408.246101,3
4,3317.667165,486.6176336,4
5,1429.639576,2483.901486,3
6,2564.622944,1983.092746,3
```

5 Results

The results are written in up to 3 output files. The output file names are composed of the base name as provided in the configuration file and appended with 'raw', 'aggr', 'plots' and a time stamp, if set in the configuration file. If the base name is 'TestRun' the resulting file names without a time stamp are:

TestRun_raw.csv, TestRun_aggr.csv, TestRun_plots.csv

If the tag 'TimeStamp' is set to 'yes' in the configuration file, the resulting file names are:

```
TestRun_raw_mm_dd_yyyy_hh_mm_ss.csv
TestRun_aggr_mm_dd_yyyy_hh_mm_ss.csv
TestRun_plots_mm_dd_yyyy_hh_mm_ss.csv
```

Note, that dd, mm, yyyy, hh, mm, ss will be replaced by the actual numbers of the day and time.

5.1 Raw Data

The raw data are written into the 'raw' output file if statistics is enabled (Tag 'Statistics' = yes) in the configuration file. This file contains detailed information on the constraint, plots, regression and anova for each iteration of each constraint. Data are organized in columns. The meaning of the column headers and the corresponding data are described in the following table.

Author	Company	Page
Lutz Tischendorf	ELUTIS – Modelling and Consulting Inc. 681 Melbourne Avenue, Ottawa, Ontario, K2A 1X4, CANADA	9 of 14

Project	Document	Filename	Version	Date
FOCUS-2.3	User Guide	FOCUS231 User Guide.doc	1	26/10/2007

column header	comment
RUN	run/iteration counter for each constraint
DUPLIC	YES/NO – shows whether duplicates were allowed or not
UNIQUE	YES/NO – shows whether unique samples were enforced or not
CONSTR	constraint name - Distance
S_TYPE	sample type – INSIDE/OUTSIDE as provided in the conf. file
C_VALUE	constraint (Distance) value
C_MULTIP	constraint multiplier
C_MVALUE	actual applied constraint (Distance) value, distance * multiplier
M	slope of the linear regression line
B	intercept of the linear regression
#PLOTS	actual number of plots sampled for this run/iteration
SST	total sum of squares
SSR	explained sum of squares
SSE	residual sum of squares
MSE	error mean square
F	F value
R-SQUARE	R ² value
R	r value
P_1... P_XXX	ids of the sampled plots

The output is identical for Plot-Plot and Plot-Buffer sampling with one exception. For Plot-Buffer sampling the Ids of the sampled Plots are reported. For Plot-Plot sampling the Ids of the sampled predictor and response variable plots are reported as follows: P_id/R_id where P identifies the plot as predictor variable, R identifies the plot as response variable and the id is substituted by the actual numerical id of the sampled plot.

5.2 Aggregated Data

The aggregated data are written into the 'aggr' output file if statistics is enabled (Tag 'Statistics' = yes) in the configuration file. This file contains information aggregated over the runs/iterations for each constraint on the plots, regression, anova and significance of higher order models (functional stability). Data are organized in columns. The meaning of the column headers and the corresponding data are described in the following table.

column header	comment
CONSTR	constraint name - Distance
S_TYPE	sample type – INSIDE/OUTSIDE as provided in the conf. file
C_VALUE	constraint (Distance) value
C_MULTIP	constraint multiplier
C_MVALUE	actual applied constraint (Distance) value, distance * multiplier
RUNS	number of runs/iterations
PLT-MEAN,MIN,MAX,VAR,KON,SDE	mean, min, max, variance, 95%-confidence interval and standard error for number of plots sampled within all runs/iterations for this constraint
SST-MEAN,MIN,MAX,VAR,KON,SDE	mean, min, max, variance, 95%-confidence interval and standard error for total sums of squares calculated for all runs/iterations for this constraint
SSR-MEAN,MIN,MAX,VAR,KON,SDE	mean, min, max, variance, 95%-confidence interval and standard error for explained sums of squares calculated for all runs/iterations for this constraint
SSE-MEAN,MIN,MAX,VAR,KON,SDE	mean, min, max, variance, 95%-confidence interval and standard error for residual sums of squares calculated for all runs/iterations for this constraint
MSE-MEAN,MIN,MAX,VAR,KON,SDE	mean, min, max, variance, 95%-confidence interval and standard error for error mean squares calculated for all runs/iterations for this constraint
F-MEAN,MIN,MAX,VAR,KON,SDE	mean, min, max, variance, 95%-confidence interval and standard error for F values calculated for all runs/iterations for this constraint

Author	Company	Page
Lutz Tischendorf	ELUTIS – Modelling and Consulting Inc. 681 Melbourne Avenue, Ottawa, Ontario, K2A 1X4, CANADA	10 of 14

Project	Document	Filename	Version	Date
FOCUS-2.3	User Guide	FOCUS231 User Guide.doc	1	26/10/2007

RR-MEAN,MIN,MAX,VAR,KON,SDE	mean, min, max, variance, 95%-confidence interval and standard error for R ² values calculated for all runs/iterations for this constraint
R-MEAN,MIN,MAX,VAR,KON,SDE	mean, min, max, variance, 95%-confidence and standard error for r values calculated for all runs/iterations for this constraint
FS_PERC	percentage of runs/iterations used for calculating higher order polynomial regression (functional stability test), only, if functional stability was enabled in configuration file
FS_OBS	actual number of runs/iterations used for calculating higher order polynomial regression (functional stability test), only, if functional stability was enabled in configuration file
ORDER_1... ORDER_5	counter for highest order of a significant polynomial regression for this constraint, only, if functional stability was enabled in configuration file

5.3 Plot Data (Plot-Buffer)

The data of the sampled plots are written into the 'plots' output file if 'WritePlots' is enabled (Tag 'WritePlots' = yes) in the configuration file. There are 3 columns for each run/iteration of each constraint. The column headers are named as follows: I_CVAL_ID, I_CVAL_P, I_CVAL_R, whereas 'I' is replaced by the number of the actual run and 'CVAL' is replaced by the actual constraint (Distance) value. The column I_CVAL_ID contains the ids of the sampled plots. The column I_CVAL_P contains the predictor values of the sampled plots and the column I_CVAL_R contains the response values of the sampled plots. See the following example for 3 columns of the 'plots' output file. The first column contains the ids of the plots sampled in the first iteration/run of the constraint (Distance) = 20. The second column contains the predictor variable for this constraint and the third column contains the response variable for this constraint.

1_20_ID	1_20_P	1_20_R
1002	0.98966616	3
102	0.99364072	1
502	0.7734499	0
1803	0.99205089	0
1703	0	0
305	0.66613674	0
1704	0.99602544	3
1705	0.99364072	0
1001	0.99364072	0
2906	0.98569155	1

Note, that the number of columns in this file corresponds to 3 times iterations times constraints. For example, if the configuration file is set for 5 distance values and 100 iterations, 1500 columns will be written into the 'plots' output file. MS Excel can only handle 265 columns at this time, but other spread sheet programs might be able to handle more. The user must be aware, that further processing of the 'plots' output file depends on the capability of the corresponding computer program. One can still sample one constraint with about 80 iterations, which will result in 240 columns.

Author	Company	Page
Lutz Tischendorf	ELUTIS – Modelling and Consulting Inc. 681 Melbourne Avenue, Ottawa, Ontario, K2A 1X4, CANADA	11 of 14

Project	Document	Filename	Version	Date
FOCUS-2.3	User Guide	FOCUS231 User Guide.doc	1	26/10/2007

5.4 Plot Data (Plot-Plot)

The data output for the sampled plots is essentially the same as for Plot-Buffer sampling. However, because two sets of plots are sampled during Plot-Plot sampling and each plot refers either to a response variable value or to a predictor variable value, the output is slightly different. The example below shows the sampled plots for the first iteration of the distance constraint 300. The first row shows the plot ids for the sampled predictor plots. The second row provides the corresponding predictor variable values. The third row shows the plot ids for the sampled response plots. The fourth row provides the corresponding response variable values.

1_300_ID	1_300_P	1_300_ID	1_300_R
338	61.3990669	135	3
354	44.9511299	118	5
465	55.9895859	76	4
375	57.3732796	154	5
356	68.0641174	25	3
357	63.5735779	65	5
346	60.9063492	164	2
387	60.5413704	173	1
310	57.3787842	102	2
476	43.8312225	101	5

6 Logging

FOCUS-2.3 writes detailed information into a focus20.log file. This file may become very large during the course of one experiment, but will be recycled at the next start of FOCUS-2.3, if the file size is larger than 10 MB. Logging information is useful for the author and will help to track potential errors or problems with FOCUS-2.3.

7 System Requirements

FOCUS-2.3 should run on any MS Windows OS, such as NT, ME, 98, 2000, XP. FOCUS-2.3 has been tested on Windows XP. The program FOCUS-2.3 requires about 400 KB free space on the hard drive, but may need much more during operation and for storing results. One should have 200 MB free space available on the hard drive, when FOCUS-2.3 is executed. FOCUS-2.3 will use about 3-5 MB of RAM, but may use much more, in particular when 'WritePlots' is enabled in the configuration file. The sampled plots are stored in the program during execution and are written after all iterations over all constraints have been completed.

8 License

FOCUS-2.3 is released under the GNU General Public License. This means, FOCUS-2.3 is free software, i.e. it will be released with the source code to allow for free and unrestricted redistribution and/or modification under the terms of the GNU General Public License. This particular license is inherited from third party libraries FOCUS-2.3 is build on.

9 Running FOCUS-2.3

Copy the following files into one folder on your computer:

- focus23.exe
- configuration file (e.g. FocusConfig.xml)
- predictor variables file

Author	Company	Page
Lutz Tischendorf	ELUTIS – Modelling and Consulting Inc. 681 Melbourne Avenue, Ottawa, Ontario, K2A 1X4, CANADA	12 of 14

Project	Document	Filename	Version	Date
FOCUS-2.3	User Guide	FOCUS231 User Guide.doc	1	26/10/2007

- response variables file
- distance matrix file (for Plot-Buffer sampling only)

Open a DOS shell on an MS Windows computer and navigate to the folder where the above mentioned files are stored. Make sure that the configuration file has correct settings and that the file names in the configuration file are identical to the file names stored on your computer. Verify the data format of the input files.

Execute FOCUS-2.3 by typing "focus23.exe FocusConfig.xml" after the DOS prompt in your DOS shell. Adjust the name of the configuration file if it differs from this example.

FOCUS-2.3 will write output to the screen indicating progress of execution. After successful completion, the output files as described in section 6 can be viewed and processed further.

10 Trouble Shooting

FOCUS-2.3 has been tested and initial bugs have been fixed. Jeff Holland compared the results with those obtained from FOCUS-1.0. The results are identical and therefore trustworthy. This can be stated here since FOCUS-1.0 and FOCUS-2.3 have been developed completely independent from each other. If, for some reason, FOCUS-2.3 does not 'behave' as described in this user guide, first verify the configuration file, the data format in the input files and the consistency between configuration file and input data. If you believe to have detected a problem, please send a short description together with the focus23.log file and all input data to the author.

11 Considerations

11.1 Duplicate Samples

When duplicate samples are NOT allowed in the configuration file, FOCUS-2.3 will maintain a history of all sampled plots during the iterations for one constraint. FOCUS-2.3 will verify that each new sample has not been sampled before, hence no duplicate samples can occur. Since FOCUS-2.3 samples the plots randomly, it will try to find a new unique sample (related to the history of already sampled plots) up to the number of retrials as specified in the configuration file. Consequently, FOCUS-2.3 will sample the entire set of plots until either a new unique sample is found or until the maximum number of retrials is reached. If FOCUS-2.3 cannot find a new unique sample within the maximum number of retrials, the program will terminate with an error message. For certain constraints, the actual number of unique samples might be smaller than the required number (number of iterations/runs). In this situation, FOCUS-2.3 will always terminate, because the requirement cannot be fulfilled. Without the maximum retrial setting, FOCUS-2.3 would than execute in an indefinite loop.

11.2 Unique Samples

Unique Samples is a more stringent restriction than Duplicate Samples. It means that each plot can be sampled only once throughout all iterations for one constraint. This reduces the number of available plot samples drastically and therefore the number of possible iterations to be conducted for one constraint. FOCUS-2.3 will use the same approach as described under 12.1 to ensure that each plot sample is a unique set of plots. If FOCUS-2.3 keeps terminating with an error message, try to reduce Maximum_Plots_Per_Iteration or Number_Of_Iterations according to the capacity of the data set.

11.3 Small sample sizes

Regression and anova will be skipped when the sample size is smaller than 2. If functional stability is enabled, the sample size must be larger than the maximum order. Otherwise, the polynomial regression and anova will be skipped.

Author	Company	Page
Lutz Tischendorf	ELUTIS – Modelling and Consulting Inc. 681 Melbourne Avenue, Ottawa, Ontario, K2A 1X4, CANADA	13 of 14

Project	Document	Filename	Version	Date
FOCUS-2.3	User Guide	FOCUS231 User Guide.doc	1	26/10/2007

12 Change Log

12.1 Focus 2.3 -> Focus 2.3.1

- Bug fix in unique sample constraint
- Added data validation for plot buffer sampling schema
- extended value for duplicate sample retrieval parameter
- minor editorial corrections in user guide

12.2 Focus 2.2 -> Focus 2.3

- Unique sample constraint added

13 References

Holland, J., Bert, D. 2003. Focusing on Relevant Ecological Scales. in review.

Holland, J.D, D. Bert, and L. Fahrig (2004), Determining the spatial scale of species response to habitat. Bioscience 54(3):227-233.

Author	Company	Page
Lutz Tischendorf	ELUTIS – Modelling and Consulting Inc. 681 Melbourne Avenue, Ottawa, Ontario, K2A 1X4, CANADA	14 of 14