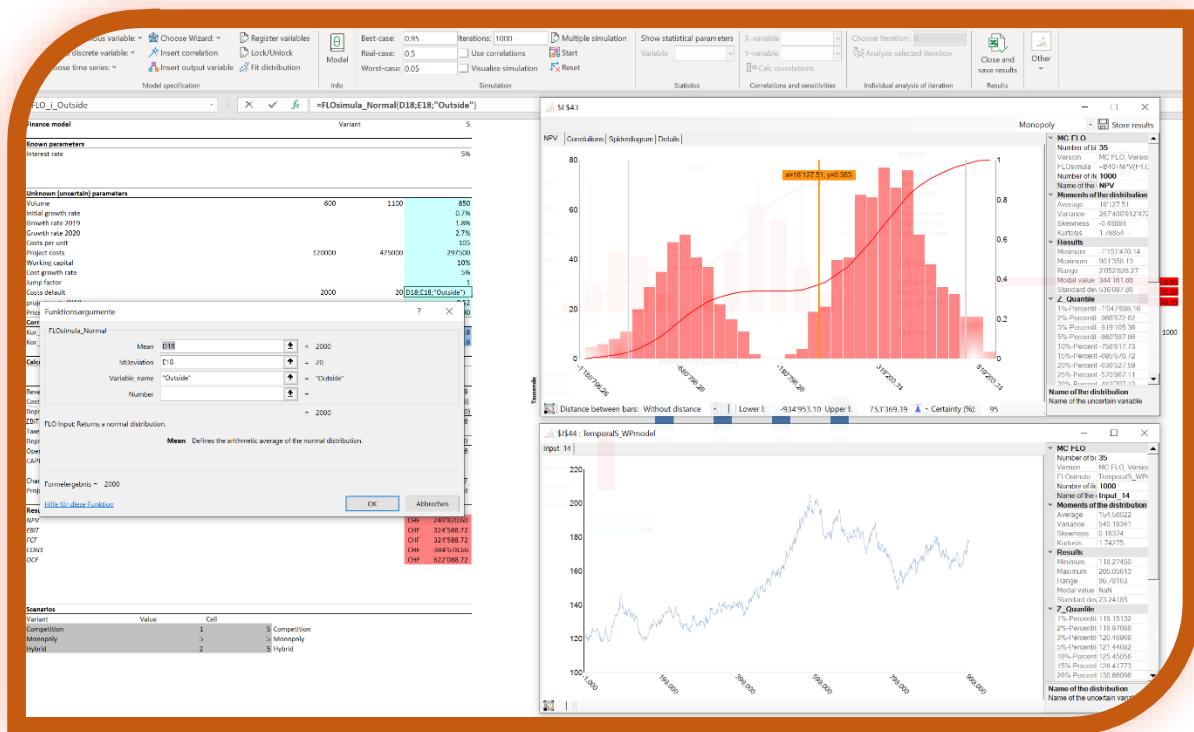


MC FLO: Simulations made easy

Version Fátima III



MC FLOsim
March 2024
support@mcflosim.ch

Summary

MC FLO allows you to simulate uncertain events with the popular Monte-Carlo simulation approach using Microsoft Excel®. Arrange easily variables to a distribution or time series function and create graphically correlations between them. During the Monte Carlo simulation and at each iteration every variable is assigned an allowable random value and the corresponding results are stored. Analyze the results using standardized charts, upload individual iterations into the model and share the results considering different scenarios – all this in the native Excel environment. No separate viewers or other add-ins are required.

Monte Carlo simulations make decisions easier. Whether in the launch of a product, in the context of corporate planning, project controlling, the management of the portfolio, risk identification and more... reveal how likely the results are and where you should pay attention for success. With MC FLO you can judge which impacts can be accepted and which should be avoided to ensure the best possible decision under uncertainty.

Use additionally the integrated time-series based forecasting tools to automatically predict potential future developments (costs, revenues, etc.). MC FLO is complemented by Bayesian statistical methods: Combine assumptions and data to make better decisions. You can use Bayesian statistics to solve linear regressions and make better decisions in classification tasks.

MC FLO is fully integrated into Microsoft Excel®. Call the integrated MC FLO functions easily as any other functions are called in Microsoft Excel®. Familiarize yourself right away with MC FLO – low incorporation costs, maximum efficiency! Use MC FLO to make data-oriented decisions.



SUMMARY	1
INTRODUCTION	3
System	3
Installation	3
More	3
Disclaimer	3
PROCEDURES TO RUN A SIMULATION	4
Model definition	4
Identification of uncertainty	5
Fit distributions and time series	12
Define correlations between inputs	16
Insert the desired output variable	17
Tips and Tricks	18
Input validation	18
Definition of scenarios and simulation mode	19
Starting the simulation and examining the results	19
Tools	32
Forecasting using time-series	46
Frequently Asked Questions	47
SPECIFICATION OF MC FLO	49
LICENSE AGREEMENT	50



Introduction

Welcome to MC FLO. MC FLO is a Microsoft Excel add-in simulation tool for Windows that puts you in the position to calculate and visualize the uncertain future and thus enables you to make more informed decisions.

System

The simulation with MC FLO requires the installation of Microsoft Excel from at least version 2010. MC FLO runs both on both 32-bit and 64-bit Excel systems. In addition to Excel, you need the .Net runtime version 4.8, which can be obtained free of charge from Microsoft. For time restricted licenses an internet connection is necessary to run the simulations. That's all.

Installation

MC FLO doesn't need a specific installation. Unzip the desired version (32-bit or 64-bit) on any local folder and copy the license file (MC FLO.txt) to your "My documents" folder. Then simply click on the MC_FLO_FIRST.xll file and confirm the load procedure. Excel will start automatically with a new ribbon menu called "MC FLO". You can immediately start with the specification of the model with a new or already existing workbook. Info: If you want to start MC FLO automatically with each Excel start, unpack the corresponding .zip file directly in the add-in directory of the corresponding folder, like "C:\Users\your username\AppData\Roaming\Microsoft\AddIns" (the path may differ, see Excel-> File-> Options-> Add-ins-> Go -> Browse the path of your add-in folder)

Important: You must have previously stored the license key in your personal Windows document folder. At <https://www.mcflosim.ch/en/the-product/faq/> you can use a small program to determine the document folder of your computer.

More

If after reading this manual some questions remain unanswered, do not hesitate to contact us (<mailto:support@mcflosim.ch>). More information and educational videos are available at our homepage. See also our blog contributions there. We urge to run the examples shipped with MC FLO to gain better insights of the capabilities and embedded functions.

Disclaimer

The program is designed and developed with the utmost care. However, since dependencies to Microsoft Excel exist, a continuous functional test hasn't be carried out. Nevertheless, by exporting the entire data you have the model results always under control and therefore

maximum transparency. Please note that the manual may be subject to change at any time. The images may differ from the actual program.

The respective rights of use of the names used here is governed by their holders.

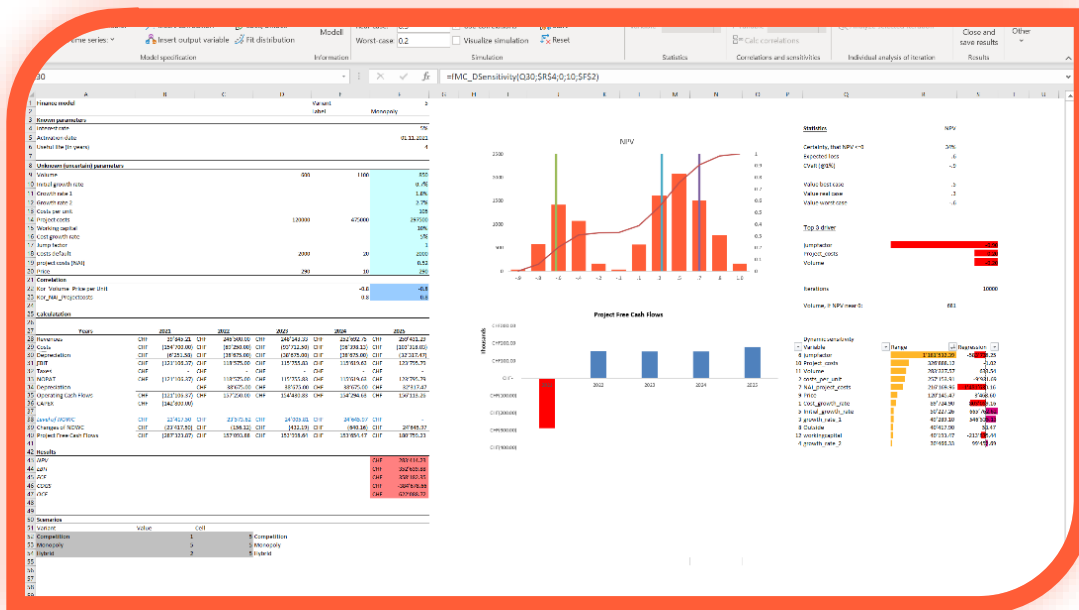
Procedures to run a simulation

The simulation is performed in three steps

- Design a model and determine which variables are uncertain. Uncertain means that you do not know exactly which value may occur in the future. But you have some idea of the possible values.
- Assign each uncertain variable to a proper distribution or time series function and define dependencies (correlations) between these variables.
- Determine which calculated values you want to examine additionally and start the simulation.

Model definition

The key-aspect of a simulation is the appropriate specification of a model that best describes the situation to be examined. Evaluating a business case needs assumptions about the future revenues and service costs incurred. Some assumptions may be given or a certain (prices), other

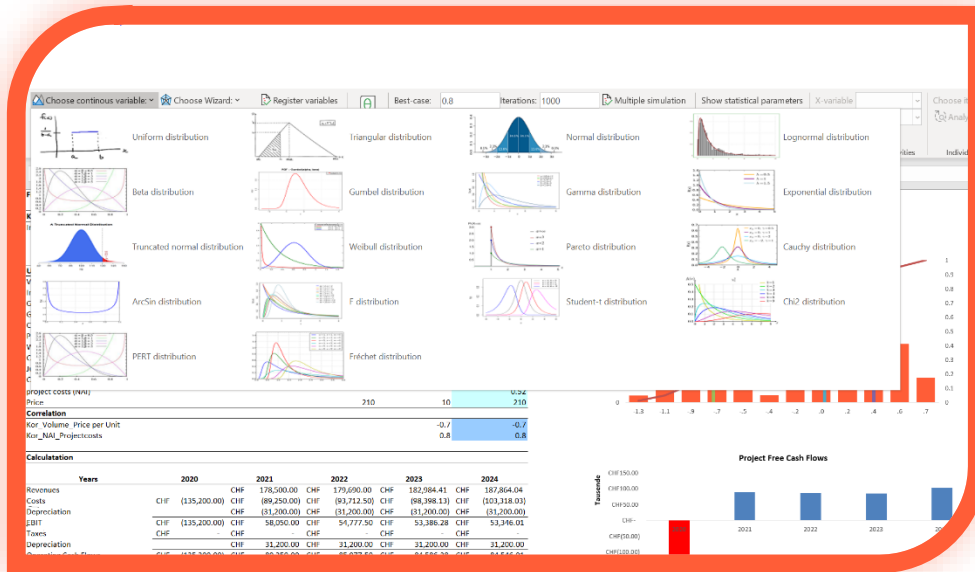


may be considered uncertain (penetration rate, costs of production). Additionally, relationships between the different variables must be considered. So, it may be that the unit cost of production and the sales volume is negatively correlated. An increase in the sales volumes will therefore be linked with a reduction in unit costs.

Document the assumptions. Follow the rule of thumb that a poorly specified model will not be salvaged by an extensive simulation run. Loosely speaking: A model to predict the weather will not be suitable to evaluate a business opportunity.

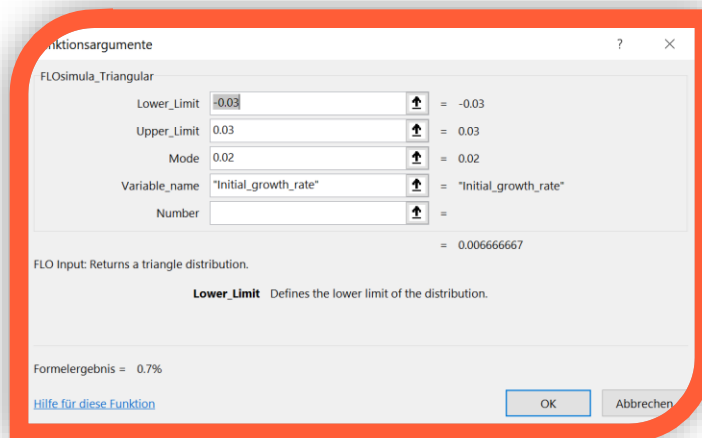
Identification of uncertainty

After the model set-up, the uncertain variables must be assigned to a distribution, a function which best reflect the future state of those variables. Here MC FLO show its glance. Specify which

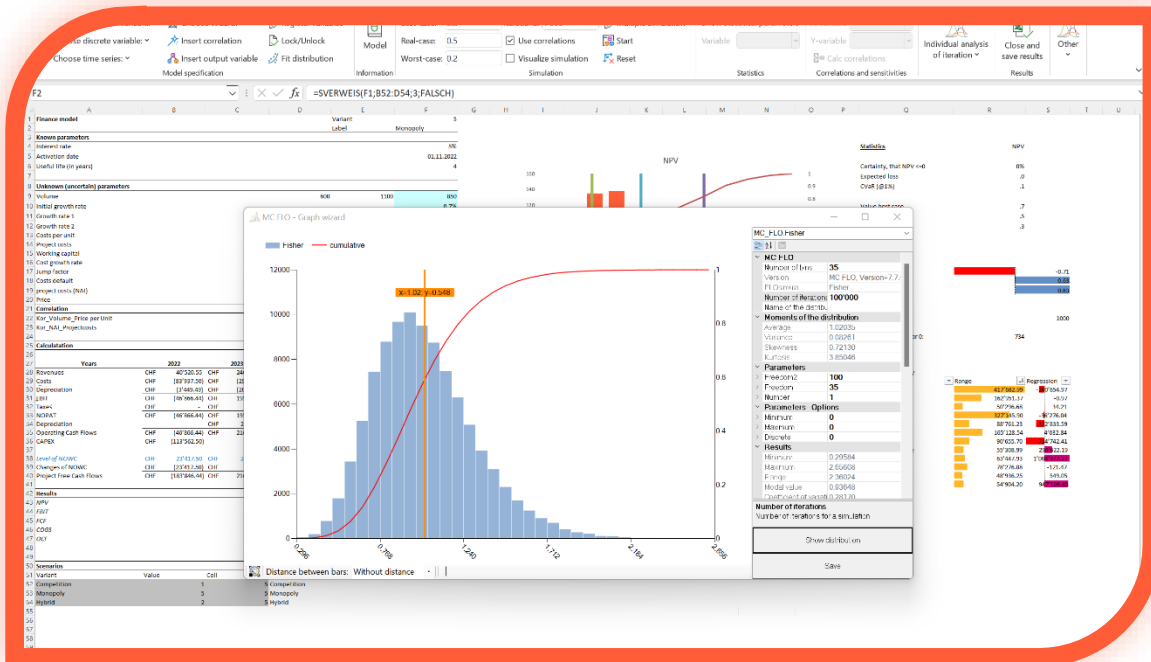


model variable (or driver) should be defined as uncertain and hence to be “modelled”.

After selecting a distribution function using the ribbon fields an Excel dialog, where – as with any other Excel-functions – different parameters must be inserted, whereby absolute values or reference to other cells for number related variables are supported. Note that characters that are not compatible with the name manager are replaced by a «_» for the variable name. After the insert process, the Excel cell will display the expected value for the selected distribution function.



If you are not sure which function you should use, use the graphical wizard to display the functions defined in MC FLO and then import them into Excel. Proceed as follows: go to "Choose Wizard" and select from the various functions the one closest to your expectations. Enter the values



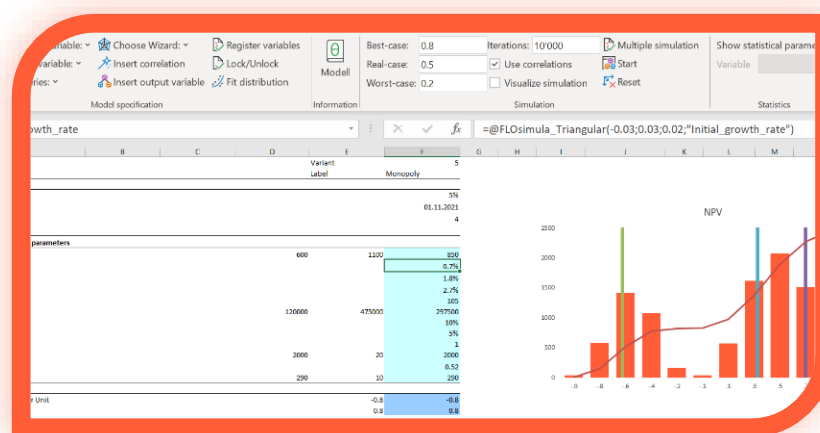
specific to the function under "Parameters". Under "Number of Iterations", enter the number of values to be considered in the simulation and define the number of columns to appear in the diagram under "Number of bins". After pressing "Show", the result of the function (in the above example the triangular distribution is shown) appears after a short time. Note, that for several functions you can also specify a minimum, a maximum and in addition, specify if the numbers should be rounded (under "Parameters - Options"). Several quantiles are displayed alongside the mean value, the minimum, maximum, variance and standard deviation. In the chart a vertical line ("magic stripe") will also be shown which displays the x-values of the distribution and the corresponding quantile values. Just press the left key of the mouse over the stripe and move to the right or to the left and the x and quantile values will automatically be updated. By clicking the right mouse on the canvas, MC FLO loads the empirical data – if any - that has been selected previously in Excel. The historical data will be loaded as "Real-Data" and shown as a line graph. Please note that you will have to reload the historical data every time you choose another distribution function. Enter a name under "Name of distribution" for the function, select an appropriate cell in your model, and after pressing "Save" the function is inserted automatically to the currently selected Excel cell.

Tip: Zoom-in or out the wizard window if the values of the magic stripe are not visible.

In Excel, the variable is presented both as a formula, starting with "FLO_i_", as well as a number, which is defined as the average value of the defined limits. The cell is also marked with a background colour and registered in the name manager of Excel.

Important: Each variable is stored in Excel in the name manager. This requires that the name of the variable cannot contain empty spaces. Make sure that other variables that are not defined by MC FLO, do not start with "FLO_i_".

You can - like any other function in Excel - adjust the formula in the Excel editor to suit your needs. The combination of several MC FLO functions is fully supported - use the variety and flexibility of Excel! Note, however, that only MC FLO formulas are allowed in an individual cell as the principal function. Otherwise, incorrect values will be shown in the result workbook. As highlighted in the figure, only a MC FLO formula is inserted in a cell. You can use other Excel functions within a MC FLO function, like "=@FLOsimula_DiscreteUniform(if(E4<100,200,300),500)". However, it is not allowed to use other functions outside a MC FLO function. Cell contents like "=@FLOsimula_DiscreteUniform(300, 500) / E4" should be avoided for a simulation. We

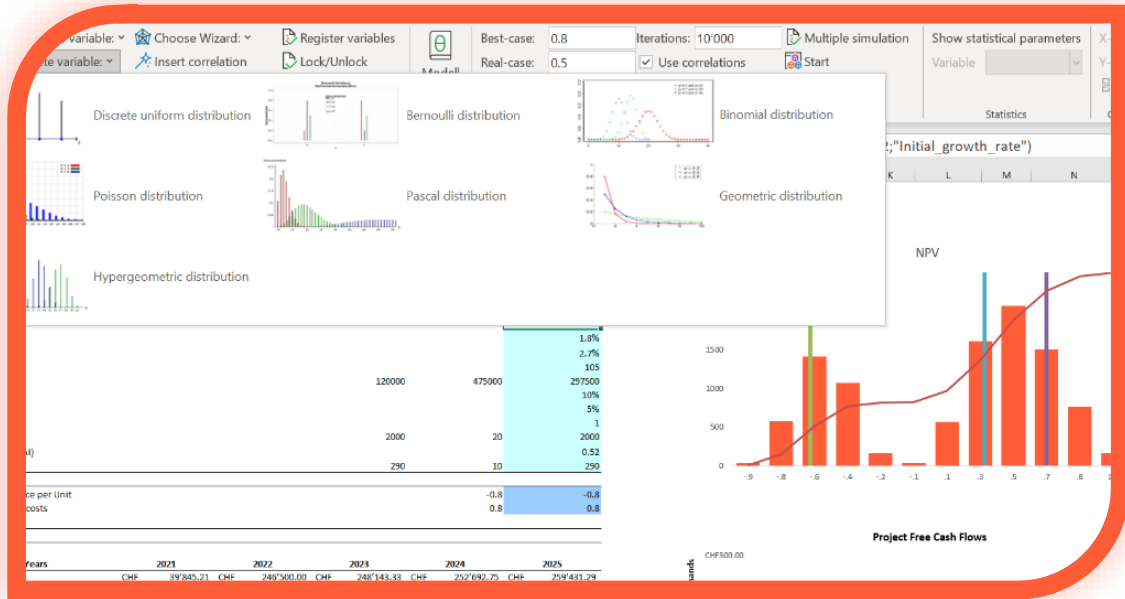


recommend using complex functions outside of a MC FLO function. Use references to the complex function instead.

Important: Make sure that the values fall within the allowable range, otherwise this may lead to false results.

Example: You define in cell A1 a uniform distribution with the lower limit of 5 and a maximum of 10. In cell A1 a value of 7.5 is shown. In cell B1, you want to specify a uniform distribution with lower limit as a reference to cell A1. Since within the simulation cell A1 can also assume a value of 10, the upper limit of the cell B1 must be a number *higher* than 10.

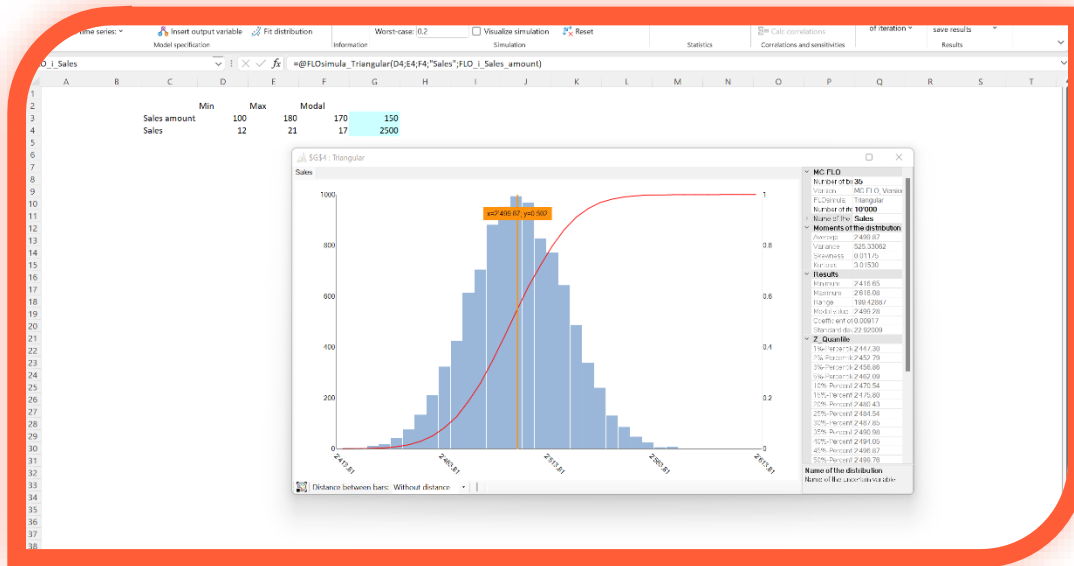
In addition to the continuous distribution functions, discrete distributions are available. These are then particularly useful if yes / no conditions are necessary or if only integer numbers should



appear in the model.

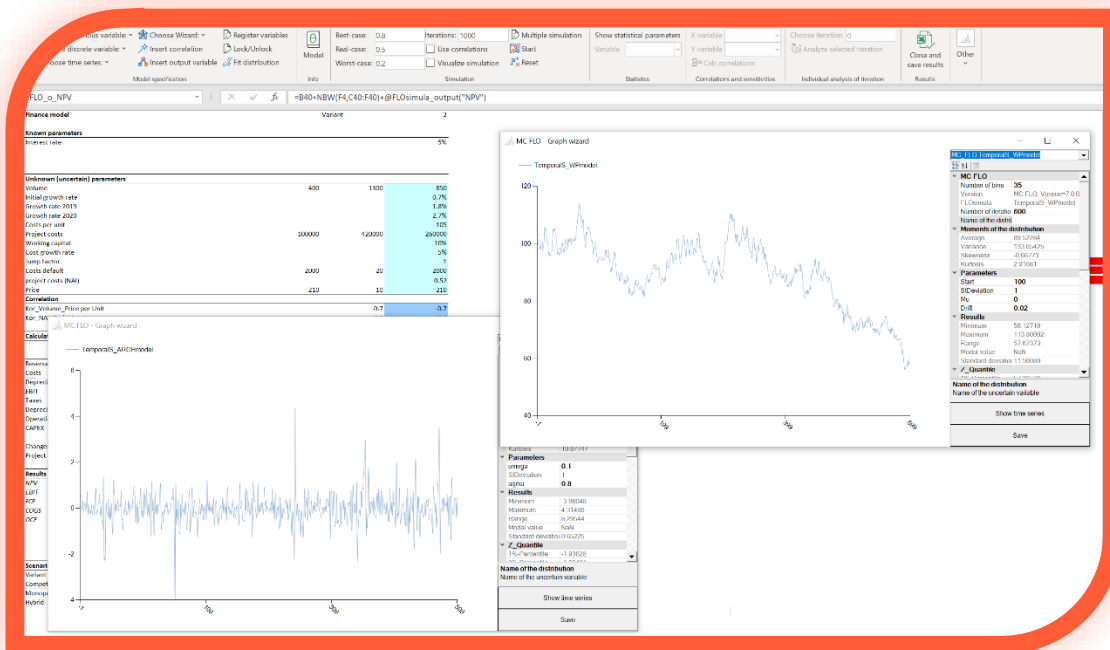
Info: In the discrete uniform distribution with the value range [0,1], the value 0 is shown as standard in Excel.

Additionally, you can set the number of occurrences per iteration run for each function. This is particularly useful when you want to multiply an uncertain variable with another variable



(constant or uncertain). For example, you can model the uncertain variable "sales amount" and "sale prices" independently of each other and thus aggregate the total to the desired "sales" figures. Commonly this is known as a compound function and can be expressed in MC FLO with the optional "Number" parameter. In the illustrative example above the compound functionality is accomplished by a triangle distribution for the sales amount and also a triangular distribution for the sale prices. You can also correlate the uncertain "Number" parameter with other uncertain functions of MC FLO, as expressed in a correlation between the number of customers and the weather conditions in the mentioned example.

In addition to distributions of events, you can also select certain time series to model price or



volume developments.

MC FLO supports stationary auto-regressive models in form of AR(1), MA(1), ARMA(1,1), additionally the Wiener process (also called geometric Brownian motion) with drift as a random walk process with normal log returns and constant volatility and the ARCH(1) process for non-constant volatility. Please note that for time series the magical stripe is not available. Please note, that a simulation with thousands of iterations for a particular time series variable is equivalent to the determination of one path of a time series process over thousands of periods. However, if you like to simulate a time series process over the next 10 periods, you should enter 10 in the optional "Number" parameter.

For continuous and discrete random variables, you can use the wizard to create a custom distribution. There are two options available for this purpose. The first option consists to mark the corresponding values of the model workbook and to transpose these into a histogram. After pressing "Wizard" and "Custom defined", the data will be loaded into the wizard and you can adjust the histogram with changing the "Number of bins". The second way is to directly import a histogram, using following the six columns as a template. "Min" and "Max" indicate the boundaries of the individual columns, "Number" the number of hits and "Percent" the relative hits. The terms ending with "cum" represent accumulated values. Here is an example where eight values are mapped into a histogram using 12 rows.

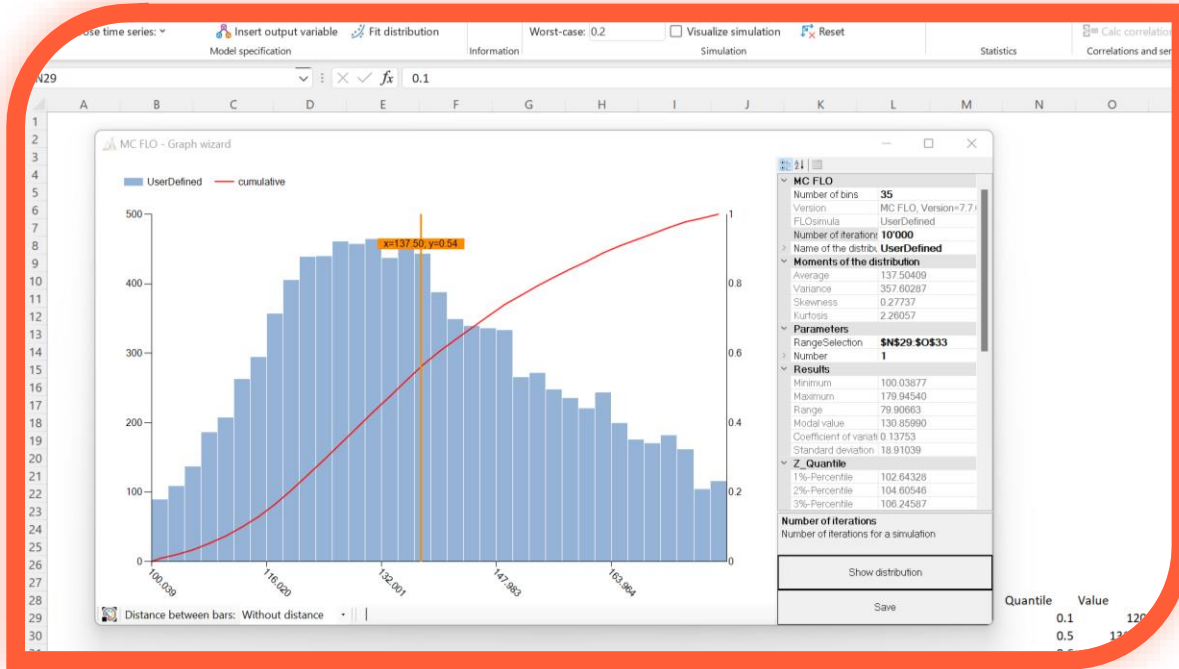
<i>Min</i>	<i>Max</i>	<i>Number</i>	<i>Number cum.</i>	<i>Percent</i>	<i>Percent cum.</i>
266347	296025	1	1	0.125	0.125
296025	325703	0	1	0	0.125
325703	355381	0	1	0	0.125
355381	385059	2	3	0.25	0.375
385059	414736	1	4	0.125	0.5
414736	444414	0	4	0	0.5
444414	474092	1	5	0.125	0.625
474092	503770	1	6	0.125	0.75
503770	533448	0	6	0	0.75
533448	563126	1	7	0.125	0.875
563126	592804	0	7	0	0.875
592804	622481	1	8	0.125	1

(the last three columns are optional)

Note that a custom histogram should be loaded without headings. «Save» saves the user-defined distribution in the model workbook, where the formula = FLOsimula_UserDefined («Data area», «Number of columns», «Name») is stored in the target cell.

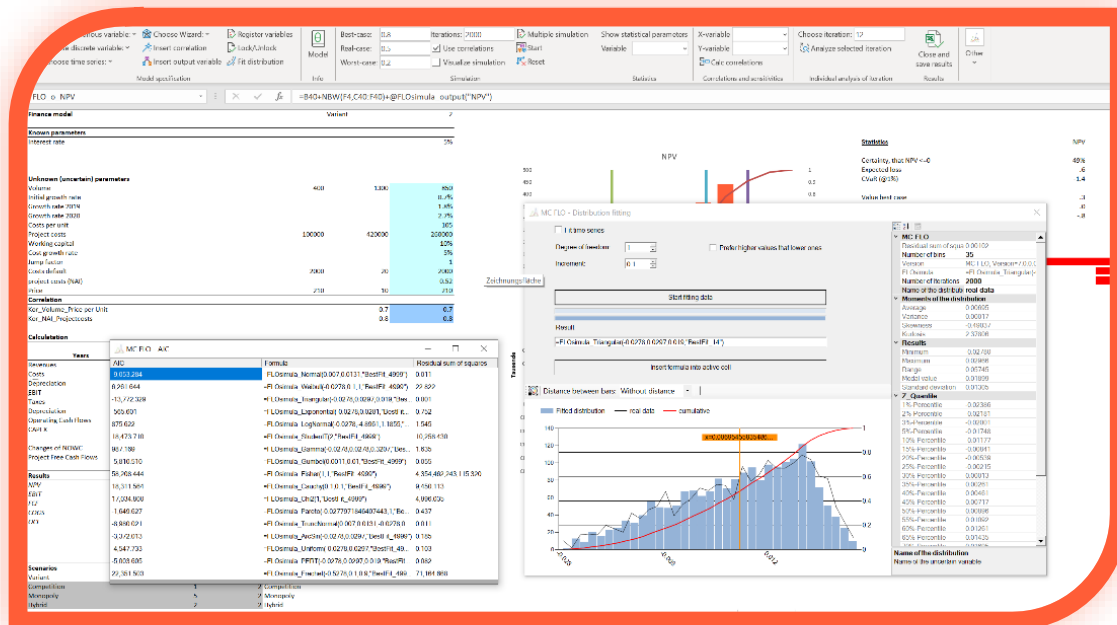
Tip: If you only want to read values distributed over a row or a column, you can prefix a name in the first cell, which is then used automatically as the name of the series when you open the graphic wizard.

Alternatively, you can read in the quantiles and the corresponding quantile values. In the present example, cells O30 to P34 were read; the first four lines contain the quantiles (about 0.1) and the quartile value (120). The last line contains the minimum (100) and maximum (180) values. Based on this information, a distribution is generated via «Select Wizard» -> «User-defined distribution», which almost corresponds to the previously entered values.



Fit distributions and time series

Occasionally important historical data is presented in large quantities (such as data on the useful



life of plants) and not allocated to any distribution function due to a lack of expertise or other

factors. If the historical data is needed for the model, a suitable distribution function is to be found as a first step. We have already introduced the graphical wizard. This presupposes, however, a certain fundamental understanding. Another possibility is to be found under "Fit distribution". Again, the data must be in an Excel. Mark the data and then go to "Fit distribution". Enter the maximum value for "Degree of freedom".

Often it is enough to set the value to 1 and to keep "increment" at 0.1. After pressing "Start fitting data", all the continuous functions implemented in MC FLO are compared with the historical data and the preferred outcome is the distribution which – in combination with the Kolmogórov-Smirnov test - has the least squared deviation from the historical data.

For the historical data, the mean (average), minimum (min), maximum (max), variance (variance), range, modal value and standard deviation (SDeviation) are displayed on the right above the quantiles (or percentiles). The selected number of historical data is indicated as "Number of iterations". Additionally, a new window, called "AIC" will be opened displaying the Akaike Information criterion for all the continuous distributions ($AIC = 2 * \text{Number of parameter} + \text{Number of iterations} * \ln(RSS)$), where RSS is the residual sum of squares (on distributions, which need the minimum and/or the maximum, the corresponding parameters are considered with a value of 1.1 for each one). Now change to the model worksheet and the formula is automatically transferred using the "Insert formula into current cell". Please note, that only continuous distributions are fitted.

Tip: If only after a substantial change of the parameters a new distribution is proposed, it is recommended to use the previous distribution to prevent an "overfitting". Note that in the result window and in the AIC window MC FLO displays the "residual sum of squared errors", by which you can judge if the improvement is acceptable (as MC FLO uses the least squared approach, the RSS should be minimized).

You can simulate every formula listed in the AIC window. Simply mark the formula, whereupon it will be displayed in the main window.

Tip: If you like to import another distribution that MC FLO has provided as the best solution, you can do so following these steps: Copy the desired formula from the AIC window (pressing Ctrl+c) and then insert it (pressing Ctrl-v) under "Result" in the "Fit distribution" window. Alternately you can also copy all the results to the clipboard by marking first all cells of the AIC windows with Ctrl+a and after clicking the menu bar on the AIC window all entries will be copied with Ctrl+c.

In addition to the stochastic time series processes, the result of the most appropriate Holt-Winter model (additive or multiplicative) will be displayed. Note that the Holt-Winters models can't be used as variables in a simulation. On the other hand, with MC FLO you can create predictions based on these models and perform automated predictions in Excel (see also Tools).

Determination of scenarios / option of multiple simulation

In some cases, you may want to investigate several options. Instead of creating your own model for each option / variant, you can perform multiple simulations in MC FLO. To do this, select the option and the corresponding parameters according to the following logic and by pressing on «Multiple simulation» the selected rows will be registered for a multiple simulation:

Variant	Value	Reference
Competition	1	=F1
Monopoly	5	=F1

Under «Variant» enter a unique name for the variant / option. Under "Value" assign a unique value (integer). The selection is completed with the cell reference (cell), which refers to the selected option. The cells are highlighted with a gray background. Details can be found in the example model supplied with MC FLO. Note that the selection of the variants has to be done without the headings. If you want to reset the option of the multiple simulation, mark any empty cell in the model and by pressing «Multiple Simulation» again the selection will be canceled. Info: When the model is closed, the selection of the multiple simulation is automatically reset. Therefore, confirm the saving of the model, so that the background color of the marked cells can be reset. Note that the example file shipped with MC FLO uses this option.

Define correlations between inputs

The definition of the input variables is the first important step. Next you specify dependencies (correlations) and their intensity selecting "Insert correlation" in the menu bar. A new window

The screenshot displays the MC FLO software interface. The main window shows a spreadsheet with a table of parameters and a calculation table. The 'Insert correlation' dialog box is open, allowing the user to select two variables (Variable 1 and Variable 2) and set a coefficient of correlation (ranging from -1 to 1). The dialog box also allows selecting a correlation type (Normal, Student-t, Archimedean copulas, etc.). The background shows a chart with two time series, 'time 1' and 'time 2', and a scatter plot of the two variables.

appears, from which you can select graphically the two input variables. Under "Correlation coefficient" a possible value between (-1,1) must be selected. After confirmation by OK a new formula is inserted in the called cell and highlighted. A variable starting with "FLO_c_" will be registered.

You can choose between elliptical correlation relationships (normal, student-t [with 4 degrees of freedom], archimedean copulas ([r-]Clayton, [r-]Gumbel, where r stands for the different rotations) and, as a special case, non-symmetric correlations (asy. Gumbel, asy. Triangular and asy. Fréchet). The correlation type is stored as an integer between 0 and 7 in the correlation formula. Note that for some correlation relationships, Kendall's Tau is used as an approximation. Under the entered correlation coefficient, the resulting rank correlation coefficient appears in grey colour.

Info: The correlations are computed only during the simulation. Therefore, the variables are not updated after the correlation has been defined. Please note that for time series the correlation is based on the random component. Therefore, a correlation between two time series process can only be modelled if the variables behave time-invariant at each iteration. You can't correlate time series where the parameter "Number" is greater than 1.

Info: Archimedean copulas can only be defined in MC FLO as a bivariate relationship. For example, you define an Archimedean correlation between "quantity" and variable "x". Consequently, it is not allowed to define a further correlation for either "quantity" or "x". Note that the definition of an Archimedean correlation relationship (copula) means that all correlations in the model must be of type bivariate.

Attention: During the simulation, MC FLO checks whether the correlation relationships are valid. If this is not the case, MC FLO automatically adjusts them and displays a corresponding message

Correlation relationships of continuous and discrete distributions in the range of the 5% and 95% quantiles are highlighted in green and red colour. Also note that only 1'000 records are displayed in the window.

Entering the correlations concludes the modelling part. By linking the uncertain variables and parameters defined as safe, you can represent the desired result.

Insert the desired output variable

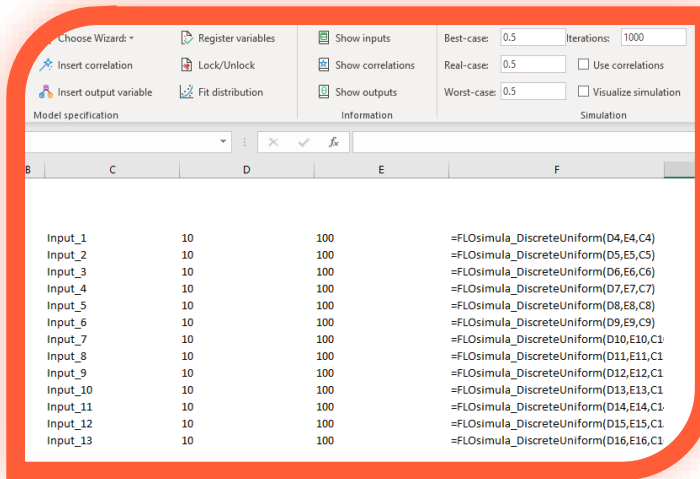
As the last step and for evaluation/analysing purposes output variables, that contain the result of the calculation (ex. net present value), must be defined.

With confirmation of the "Insert output variable" button, you are prompted to enter a name. Pay attention to the conditions applicable to the input variable name conventions. After the name specification is finished a formula called "= + FLOsimula_output ("variable name")" is inserted and the target cell highlighted. A variable starting with "FLO_o_" will also be registered in the name manager of Excel. Before the "+" in the formula the reference to the cell which is to be analysed has to be inserted.

By entering at least one input variable and one output variable the model specification is formally completed and you can start the simulation.

Tips and Tricks

Very often it can happen that you want to assign the input, correlation and output variables continuously in a column or row. Instead to call the input mask or the wizard for each cell



individually, you can proceed as follows: Go into the cell which will contain the variable and enter the formula by hand. Start with "=FLOsimu". The functions defined in MC FLO will appear. Choose the right one which you want to include. Take as the variable name a reference to another cell. Confirm the entry for all parameters. Then tighten

the definition to the range which should be filled with the new variables. Please note that for compatibility reasons the parameter "Number" cannot be passed as a parameter. You must introduce the number of occurrences after the automatic processes has been finished.

Close the manual entry with "Register variables". As a result, the variables will be registered in MC FLO.

For constant and discrete distributions, you can define barriers with the distribution function «FLOsimula_Bounded». This function must be entered manually using the Excel editor and then completed by confirming with "Register variables". The function «FLOsimula_Bounded (Reference, LowerLimit; UpperLimit)» contains as first argument the MC FLO variable which has to be bounded, the parameters «LowerLimit» give the lower barrier and «UpperLimit» the upper barrier. As explained before, this function can be resembled by using directly bounds on the corresponding distribution function.

Input validation

Check the selected parameters by calling the appropriate "Show" button. You will see a summary of all the variables in the selected category (Input, correlations, Output). The affected cell is highlighted in Excel by clicking on a column of the summary. Info: If the preview mode is enabled, the distribution and correlations will be displayed graphically. This will also be possible after a simulation run (see below).

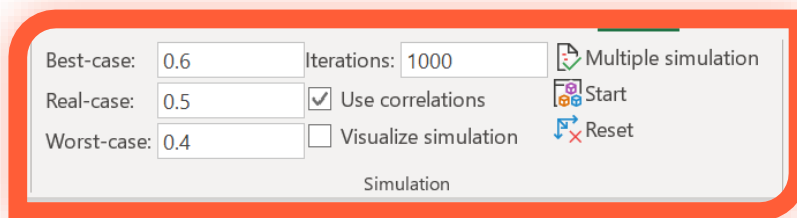
Important: If you do make changes during or after completion of the model, take note of the variables involved and remember the referenced values. If you want to delete an input variable

that is referenced by another input variable, change the specification of the last one before the former is deleted.

You can fix individual variables to the initial value as required. This keeps the variables unchanged for each iteration. To do this, select the input variable to be fixed and confirm the selection with «Lock/Unlock». The selected input variable is highlighted in red. During the simulation, the locked variables remain unchanged with the expected value of the distribution. The fixation is cancelled by selecting the variables again and pressing "Lock/Unlock". Make sure that the input variables are red-coloured only having used the MC FLO "Lock/Unlock" function, otherwise they will be blocked.

Definition of scenarios and simulation mode

MC FLO supports the definition for three scenarios ("worst", "real" and "best"). For each one you must specify the required confidence level or quantile. The confidence defines the region in which the required quality is very likely to happen. Example: You plan the investment budget for the



next year. Values between 10,000 and 80,000 may result from the simulation. Which value should be considered in the budget? You define in advance for the

real case that with a "probability" of 10% a budget overrun can be tolerated. In the worst case, you want to allow enough funds in the budget, so that a budget overrun occurs only with a 5% "probability". For the real case, you define a confidence level of 0.9 (1 minus "residual probability") and for the worst case the 0.95 quantile is defined. After the simulation run the corresponding value for each scenario is displayed in the result workbook.

You can define in advance how Excel will behave in each iteration. By checking "Visualize simulation" the values of the input, the outputs and the embedded Excel charts are actualized at each iteration. Note that this simulation mode is very slowly and should be used only for demonstration purposes.

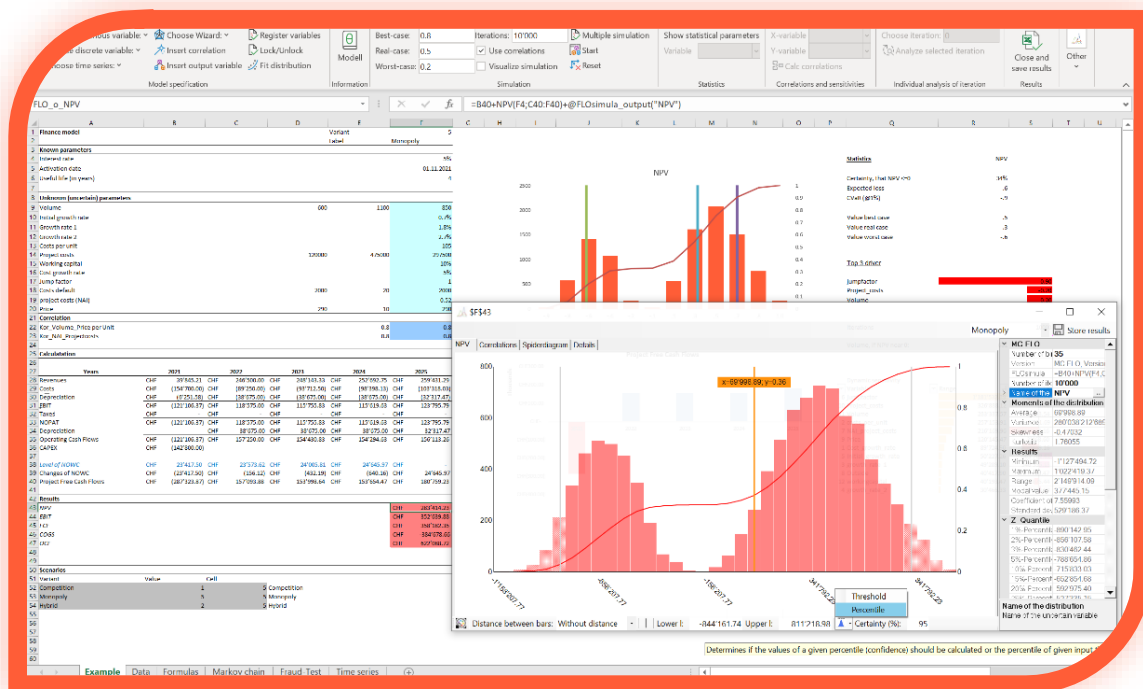
Starting the simulation and examining the results

Now is time to start the simulation. Select an output variable, specify the number of iterations at "# iterations" and start the simulation with "Start". On the status bar of Excel, the progress of the calculation is displayed. The end of the calculations will be completed with the appearance of a separate window containing the results from the selected output variable. As with the wizard interface you can also use the magic stripe to display the x-values of the resulting distribution and the corresponding quantiles. Press the left mouse button on the magic stripe and move it to the right or to the left and the corresponding values will be displayed automatically.

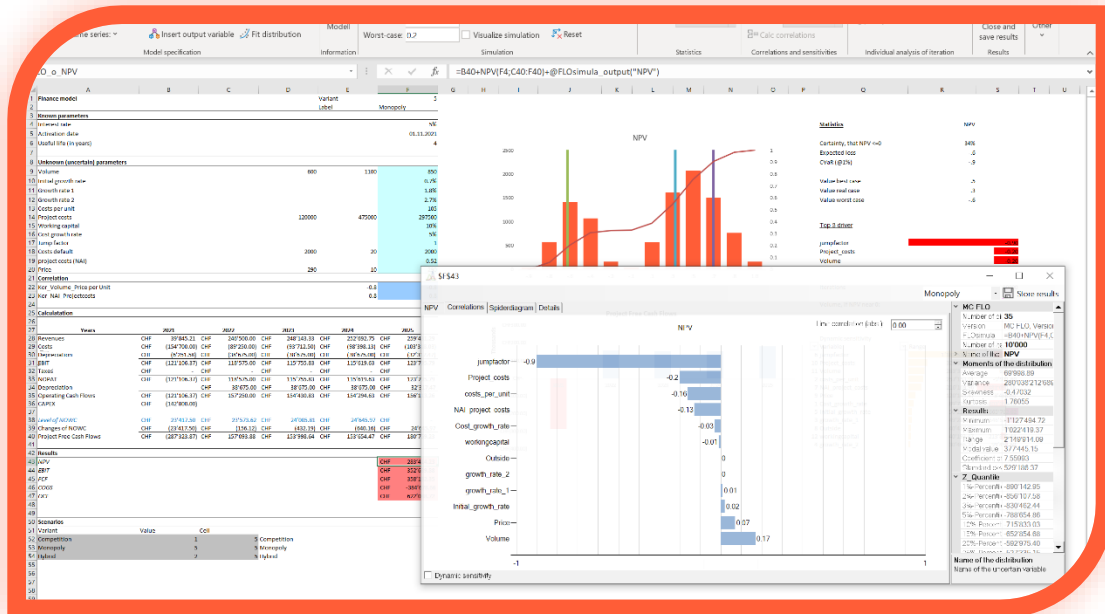
In the lower range, the values for the 95% confidence level (security) appear, also the corresponding bars appears in highlighted color. You can adjust the level or under by selecting "Thresholds" determine new lower (Lower L.) and upper (Upper L.) boundaries and the associated confidence level (see also <https://youtu.be/ZcpmDhRs33I>).

Resize the window if necessary in order to view the values on your screen. From the resulting output window you can store the entire results in a separate workbook by pressing "Store Results".

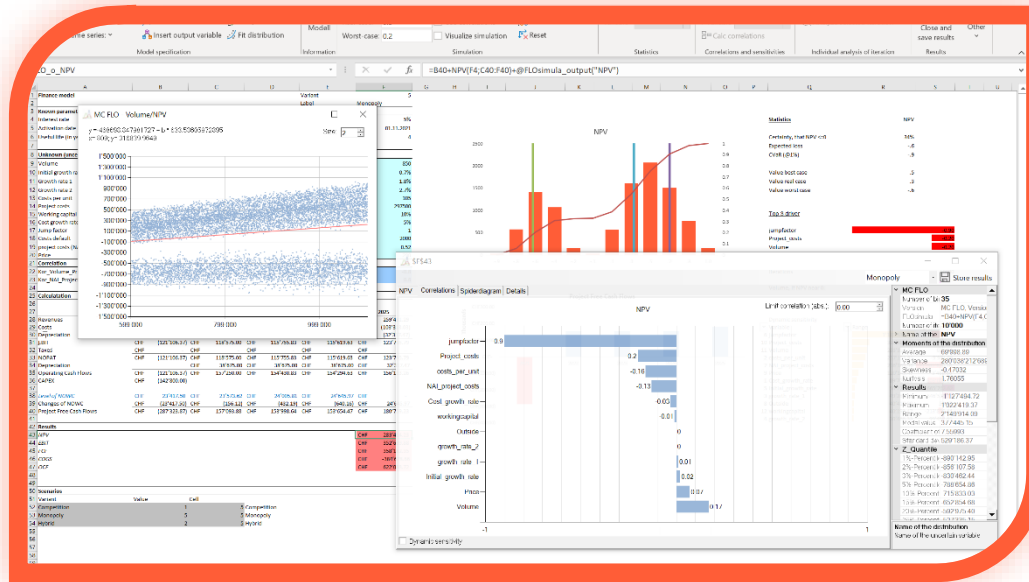
By adjusting the number of columns in the window or by changing the color or the number of bins you can adjust the histogram for your purposes. For output variables you can additionally use a tornado graph to display the correlations (according to Pearson) between the selected output variable and all input variables. This allows you to quickly identify the crucial input variables and thus to make better decisions. However, please note that the Pearson correlation coefficient assumes a linear relationship.



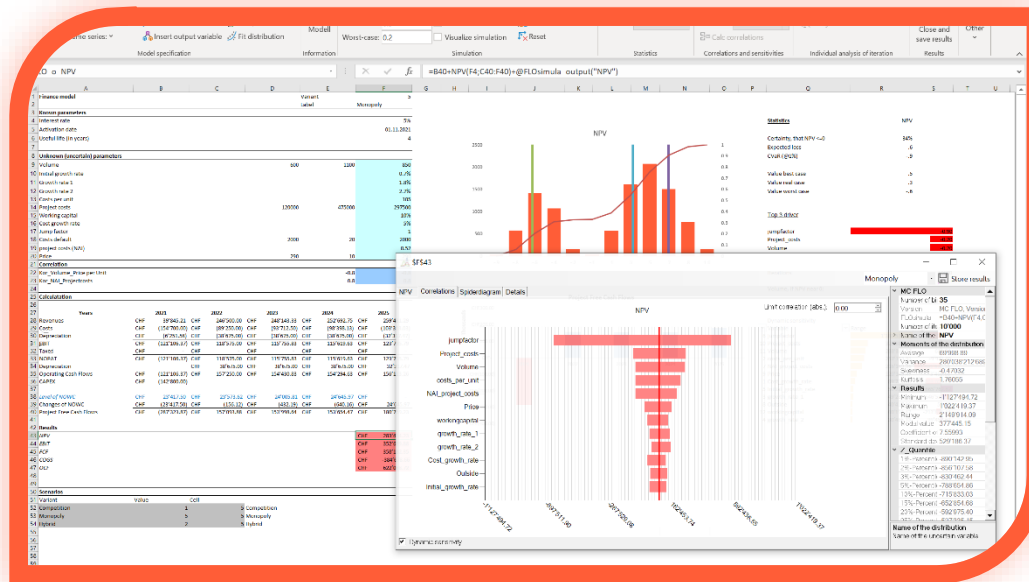
Navigate through each input, correlation or output variable after a simulation run and at each time a new window appears, summarizing the outcome with standard statistics. Input variables are shown in blue, output variables in orange color.



By pressing any of the correlation bars a new window will open, which will show the details in a chart. This will give you a visual check if Pearson's linear correlation is feasible, showing the linear

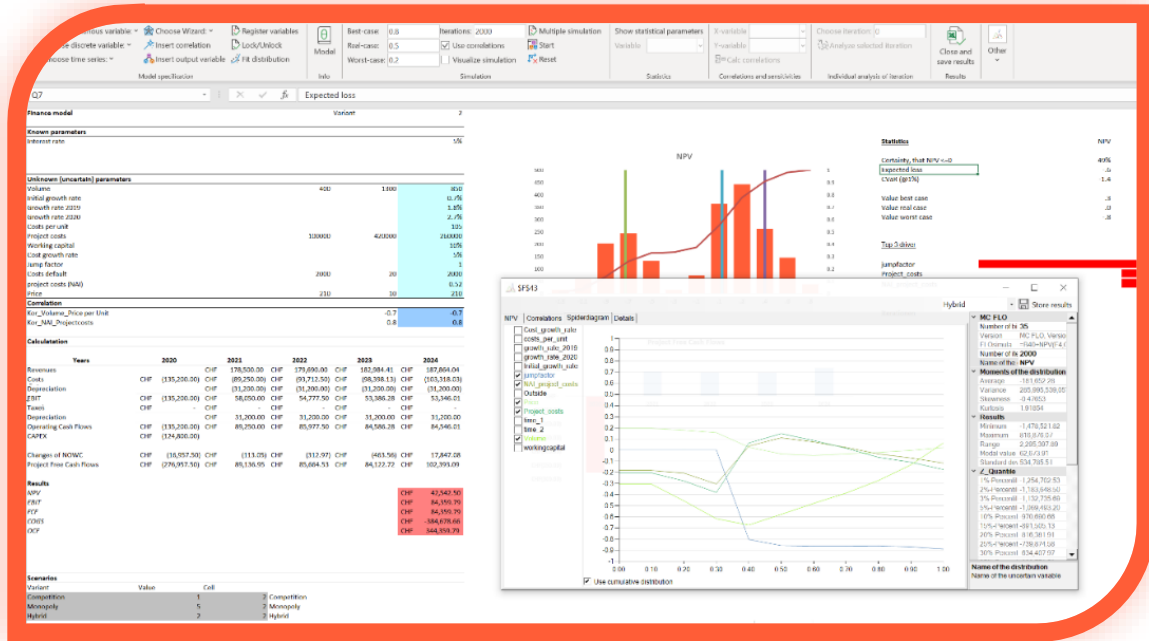


regression equation. In addition, you can also see the effects using a dynamic sensitivity.

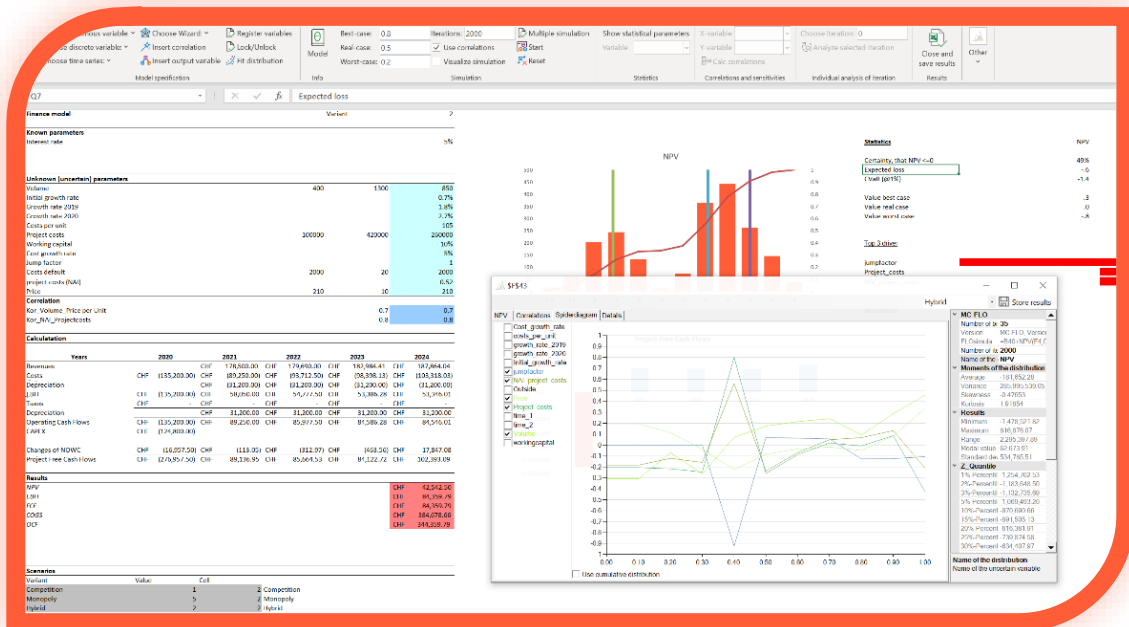


In this case the minimum and maximum of an interval-based relationship between the input (where the input variable is initially sorted in ascending order) and the output variable is shown.

As a complement MC FLO facilitates through a spider diagram the correlations by quantile, either in isolation or cumulatively. At the outset, the input variables that have a correlation equal to or



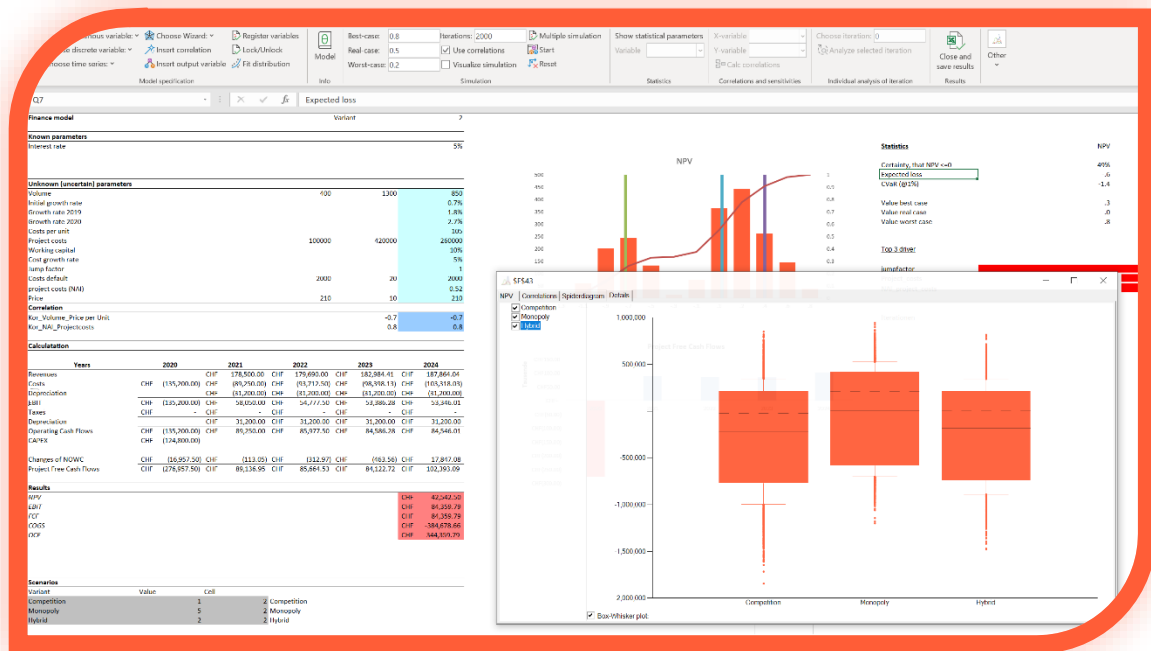
greater than +/- 0.3 in any of the quants will be highlighted in any of the quants. In the example

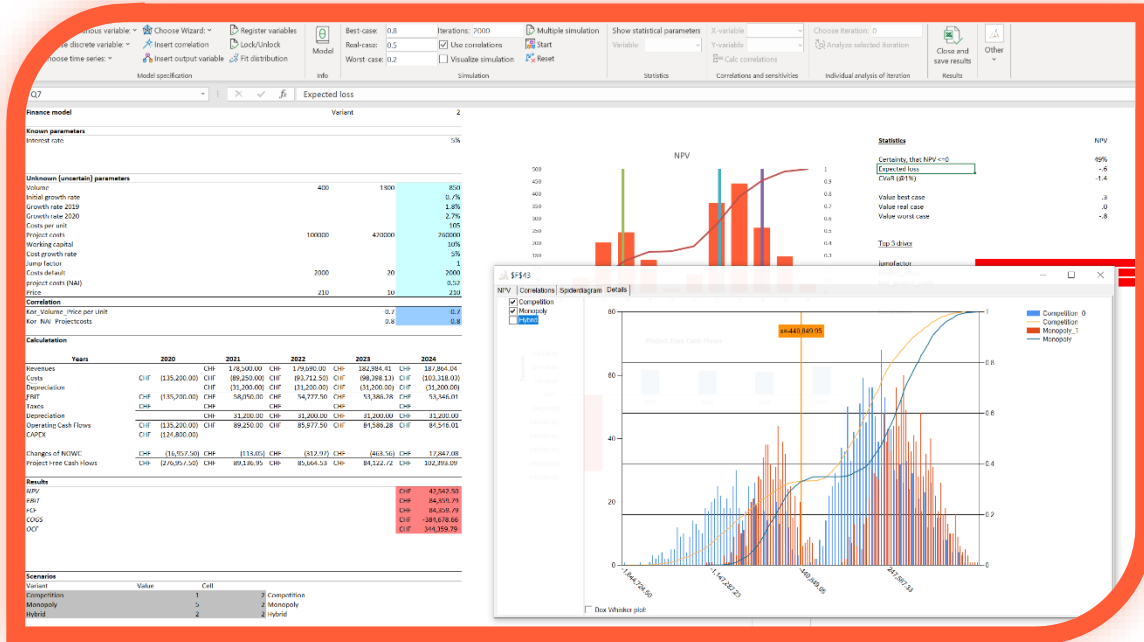


above the correlation between the variable "jumfactor" and "NPV" is zero in the quantile between 0 and 0.1, comprising corresponding Values of the NPV between -1'048M and -748k.

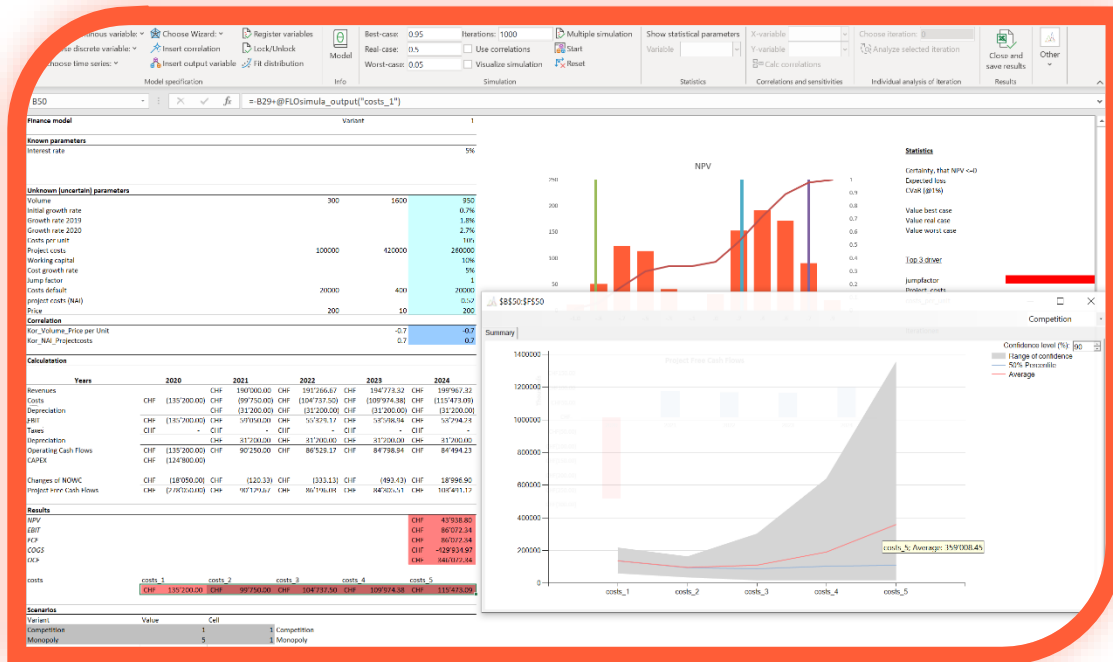
If you use the "Use cumulative distribution" option instead, the correlation will be displayed cumulatively (the "jumpfactor" variable has a correlation of approximately -0.9 between the NPV minimum (-1'048M) and quantile 0.4 (equivalent to 140k). In quantile 1 the cumulative correlation is equivalent to the calculated correlation.

If you have activated the multiple simulation, you can make a graphical comparison between the variants under "Details" for the selected output variable. You can also select the individual variants of the corresponding variables in the main window, which triggers Excel to refresh the cells and graphics.





In addition, you can select several input and/or output variables at once and display the corresponding values in a summary window, optionally adjusting the confidence level. Please note that only 1'000 variables can be displayed at once.



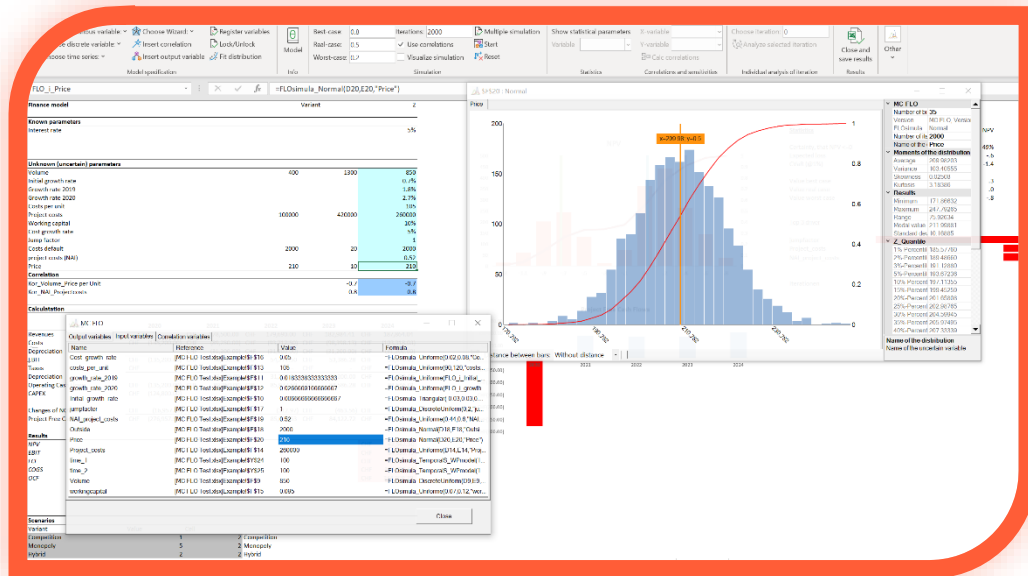
Info: To minimize the calculation time no further action with Excel is possible after starting the calculations.

In order to simplify the exchange with users who do not want to have MC FLO installed, you have the option to store the input variables, correlations and the results as a background image in the corresponding cell as a comment. By double-clicking on the corresponding diagram, the window is saved in the corresponding cell. Note, however, that existing comments will be overwritten.

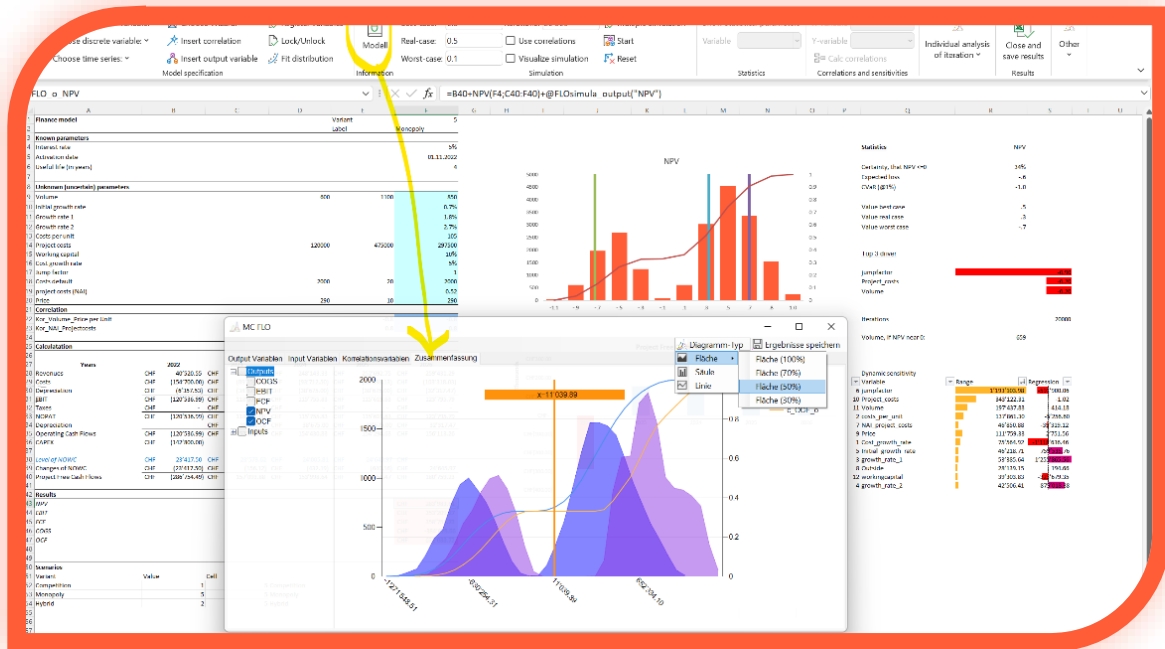
Info: You can reset the simulation with “Reset”. In this case all the results from the simulation will be erased from the computer memory.

Tip: To validate the model we recommend starting the simulation with a small number of iterations. From the results displayed, you determine whether the model needs to be adjusted before a exhaustive simulation (with thousands of iterations) should be performed.

Tip: You can visualize the input, the correlations and output variables in a preview simulation mode before the first simulation run. Proceed as follows: Under «Other -> “Seed» parameter enter a -1. Then go to a cell with an input, correlation or output variable and a new window will open which shows the shape of the variable according to the formula definition. You can leave the preview simulation mode by assigning the seed variable a value different to -1. Note that the preview simulation mode is not faster than the normal simulation mode and uses only random



variables for the input and correlation variables. You can extend visualization context to the entire model space and compare the results only graphically. Choose different style options (diagram type), if needed.

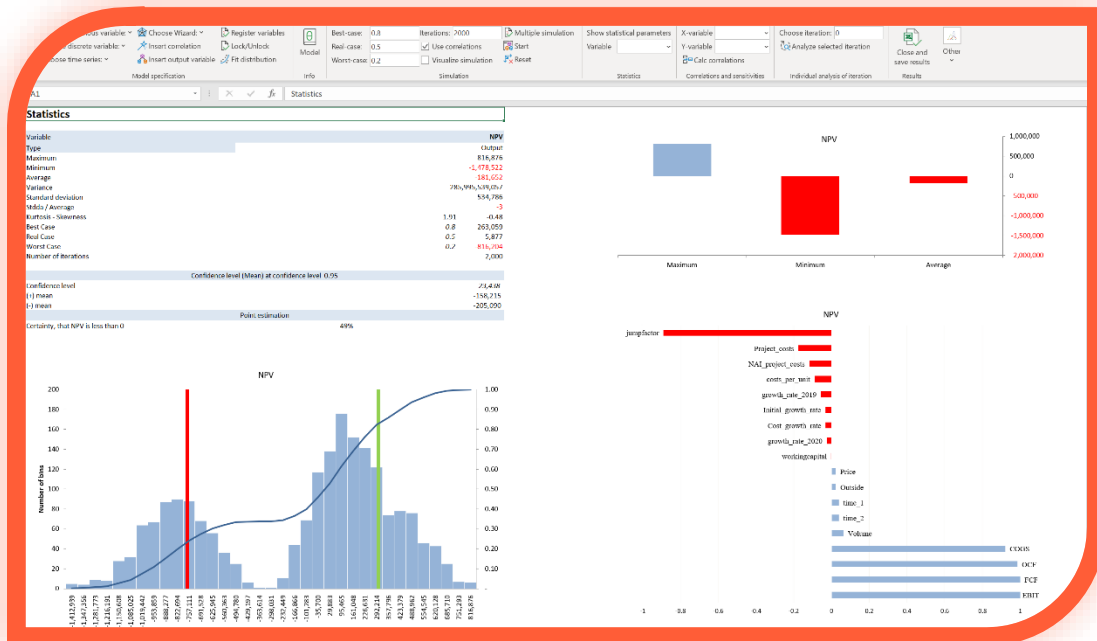


Note that a preview simulation mode is pre-build for the input variables without the need to set the seed parameter. For this select a cell containing a input variable and press “Start”. A new window with the possible results will be shown.

As mentioned before, the results can be stored by pressing “Store results” from the corresponding window of an output variable after a simulation run. The results will then be displayed in a new workbook, called the result workbook. This workbook will initially contain the following worksheets

- **Data:** Here all the values of the input variables and the output variables of each iteration are stored. You can verify the results with reference to the numbers stored. Info: The data in the worksheet "Data" is write-protected.
- **Summary:** All the input and output variables with the calculated maximum, minimum and average is value are displayed, followed by the corresponding variance, standard deviation and the results of the scenarios ("worst", "real" and "best"). The information is always compiled with the information stored on the Worksheet “Data”.

- Additionally, the skewness, kurtosis and all quantiles are represented in 5%-point steps, the extreme values (in 1%-point steps) and the +/- 3%-point deviations from the real case are displayed.
- **Statistics:** According to the selected input or output variable from the result workbook



the parameters of the summary worksheet are shown. Additionally, a confidence level of the mean (the standard is set with a confidence level of 0.9), a histogram with the best and worst scenario outlined by a vertical line, the limit value (typically the zero value) and a cumulative distribution curve is shown. The mentioned tornado graph reveals the relationship (correlation) between the selected variable and all other variables. Info: The Excel formula to determine the confidence level of the mean assumes a normal distribution. If the examined variable does not follow a normal distribution, the calculated confidence level of the mean is not meaningful and should not be used further.

You can freely select the number of bars ("classes") to be displayed in the histogram. From the result workbook enter the number in the corresponding field "Set bins" and confirm with «Change histogram». Note that MC FLO sets the number of bars back to their original value each time a new result workbook is opened.

Info: To analyse time series and their respective correlations you should copy the data from the sheet "data" to a new sheet and insert a new diagram if necessary. You can now easily check the correlations visually between the data. Please note that the correlation may differ from the

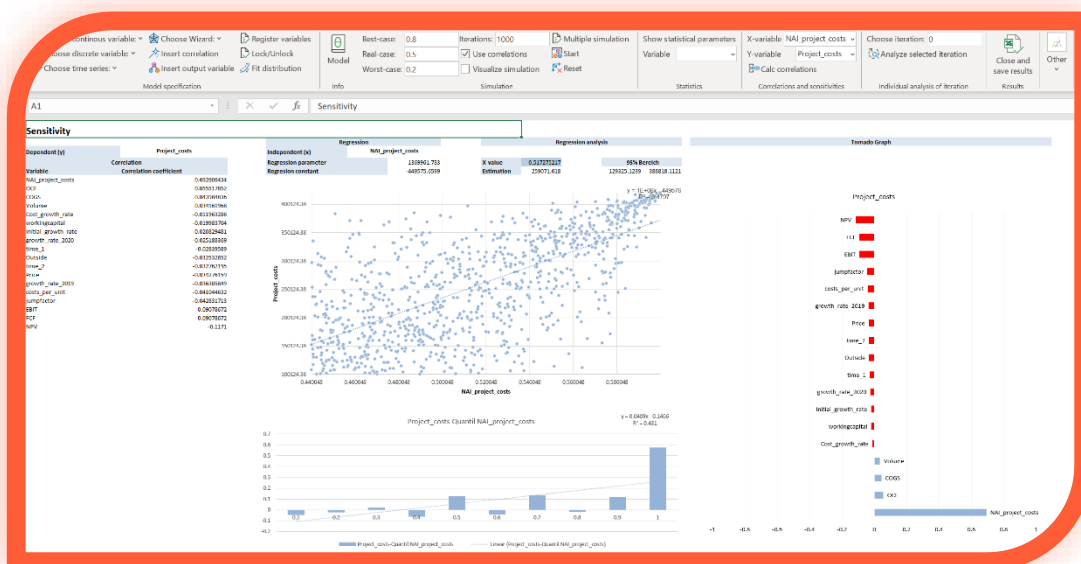
correlations defined in the model ("Spearman's Rho"). This corresponds to the fact, that correlations between time series are established upon the noise process.

Tip: You can change the label of the axes individually. The diagram can be integrated into other Microsoft Office products or you can share the spreadsheets with others. Change the colours, the size, add new information and so on. No additionally viewers are needed!

The close of the new workbook is completed with "Close and save results". The workbook is saved in the same folder where the model is stored. It can always be opened and analysed by MC FLO again.

Info: The new workbook is always saved with the date (day, month, year), the time (hour, minute, second), separated by "_" each, and the label "MC_FLO.xls" additionally to the name of the model specified. Info: if you have performed a simulation with more than one million iterations than only the iteration results will be stored in a "csv" file.

In addition to the analysis of individual variables, you can under "Correlation and sensitivities" examine the relationship between two variables, both graphically and numerically by regression analysis. Under the "X-variable" and from the result workbook ribbon fields you select the (independent) first variable and under "Y-variable" a (dependent) second variable. Confirm the



selection with "Calc correlations". In the result workbook a new sheet with the label "Sensitivity" is inserted and displayed. In the left side the correlations of all variables with respect to the selected "Y-variable" are shown. In the middle the simple linear regression analysis of the two selected variables is displayed. By entering new values in cell I5 you can calculate a forecast using simple linear regression. The corresponding approximated 95% forecast interval is presented in

cells K5 and L5, with an approximation based on the normal distribution and assuming a large sample (at least 300 iterations of a continuous or discrete distribution). Also note that the residuals must be independent and normally distributed if you use a linear regression. From line D52 you can read further statistical results of the regression analysis (hidden by default). You can perform further predictions using the linear regression analysis. Enter in cell I4 a new value for the independent variable and on cell I5 the estimated value is shown.

For the dependent variable, the correlation with the independent variable is displayed graphically in 10% -quantile steps. In this way, you immediately discover the area in which the correlation is particularly pronounced. A linear regression equation indicates the direction.

At the right the tornado graph from the summary worksheet is shown.

Info: If you analyse two variables that were set with a correlation relationship before the simulation run you will discover that the correlation result in Microsoft Excel does not coincide with the entered correlation coefficient. This is due to the method of calculation of MC FLO. MC FLO compiles the Spearman algorithm to determine the (rank-order) correlation coefficient, however, Microsoft Excel uses the linear Pearson correlation coefficient.

The screenshot displays the MC FLO software interface. On the left, the 'Simulation' tab is active, showing 'Finance model' with 'NPV' as the variable. The formula bar shows $FLO = NPV = \text{B}40 + \text{NRW}(F4; C40; F40) + @ FLO\text{Simula. output}("NPV")$. Below this, 'Unknown (uncertain) parameters' are listed with values for initial growth rate, discount rate, and various costs. The 'Calculation' section shows a table of values for years 2020 to 2024 across different categories like Revenues, Depreciation, and Operating Cash Flows. On the right, the 'Iteration results' table is visible, listing iterations 1 through 27 with columns for Name, Type, Output, Cost, Inv, ICF, NPV, DCF, Cost growth rate, Cash per unit, and Growth rate. A red box highlights the 'Results' section at the bottom left, which includes NPV, DCF, ICF, and COGS values.

MC FLO allows you to check the calculations individually. You can load the simulation results of the uncertain variables in the original model and compare the resulting results in Excel with the calculated data in the worksheet "Data". Choose under "Individual analysis of iteration" from the model workbook the selected iteration and confirm the selection with "Analyse selected iteration". As a result, the values of the input variables of the concerned iteration are loaded from

the result workbook into the original model, where they appear as numbers rather than as a formula. You can restore the formulas by closing the result workbook or by pressing any “Show” ribbon button from the model workbook.

It is easy with MC FLO to share the result with others. All the data of the calculations and diagrams are stored in native .xls files and can be analysed further with MC FLO. Independent advisors can check and proof the values by other instruments. Be one step closer with MC FLO and your success!

Tools

MC FLO also provides some functions via the formula editor of Excel or by the ribbon buttons provided by MC FLO. Note that some functions require a previously performed simulation, otherwise a value of -1 is displayed in the cell. In addition, we recommend that you do not use these functions in a visual simulation. Some functions can also be evaluated with reference to variants (multiple simulation) or using a visual simulation. To combine a visual simulation with a live creation of a histogram, we refer to the Excel program enclosed with MC FLO. Some of these formulas have a volatile extension – these formulas are then recalculated automatically with each change on the underlying data, otherwise there are by default volatile. As an example, the volatile extension of fMC_Pearson is fMC_PearsonV. fmc_Name on the other side has not a volatile counterpart and is therefore by default volatile. Note: "Volatile" features increase the level of automation but can severely affect performance. Especially for functions that perform extensive calculations, but are rarely used, a non-volatile variant is more efficient. Some formulas are controlled directly via the assignment of the corresponding data (e.g. “DataX”, “DataY”) instead the name of a variable. In this case the functions end with “Vector”; but the algorithm remains the same. For reasons of simplicity, this class of formulas will not be described in more detail below. Please note, however, that some formulas with the ending “Vector” require a previous simulation calculation, otherwise an error message will be displayed.

Please also note that the formulas mentioned in connection with regressions currently require quantitative input variables.

Formulas

fmc_Name(Name): Returns the name of the unique MC FLO variable from the specified cell of the same workbook; can be used in conjunction with other functions, like fMC_Min(fMC_Name(Name)).

fMC_Spider (Name1; Name2; Value1; Value2): Calculates the linear correlation coefficient (Pearson) between the data points Value1 and Value2 of the variable Name1 with the corresponding data of the variable Name2.

fMC_ExtremeValue (Name1; Pos): Returns the value at the i-th position of the sorted list of the variable Name1.

fMC_PosValue (Name1; Pos): Returns the value at the i-th position of the variable Name1.

fMC_FitC (column1; column2): Performs a correlation adjustment of two series located in column1 and column2. The result is the adjusted correlation type as a number between 0 and 7 (see also [Define correlations between inputs](#)).

fMC_Pearson (column1, column2): Determines the correlation coefficient according to Pearson of two rows of data, which are in column1 and column2.

fMC_Spear (column1, column2): Determines the rank correlation coefficient according to Spearman of two rows of data, which are in column1 and column2.

fMC_Sort (column1, pos, [optional] Sort = 0,1): Returns the position (pos) of a sorted list whose numbers are displayed in a column. Sort is used to specify whether the list is sorted ascending (0, defined as default) or descending (1).

fMC_BSMVal(Basis, Time, Interest, Sigma): Returns the price of an underlying after "Time" periods, given the "Basis" price, a riskless interest rate "Interest" and a constant volatility "Sigma" according to the Black-Scholes-Merton" model.

fMC_Batch_Time (Matrix1, Number, ResultOrder, Selection, NumIterations, Percentil): Marks the column or row selected in Matrix1 as time series for a batch run. "Number" defines the number of data to be forecast, "ResultOrder" determines how the estimated historical and the forecast data should be displayed in Excel (0 = line by line, 1 = column by column), with the forecast function selected based on the "Selection" criterion (0 = Akaike Information Criterion, 1 = Sum of Residuals). "NumIterations" reflects the number of iterations and "Percentil" the resulting percentile to be considered.

fMC_Batch_HW_D_Time (Matrix1, Number, ResultOrder, Selection, NumIterations, Percentil): Marks the column or row selected under Matrix1 as a time series for a batch run using the Holt-Winters models (additive, multiplicative), with the data on a daily basis and a forecast for the next few days. "Number" defines the number of forecast data, "ResultOrder" determines how the estimated historical and new forecast data should be displayed (0 = line by line, 1 = column by column), with the forecast function selected based on the "Selection" criterion (0 = Akaike information criterion, 1 = sum of the residual squares). "NumIterations" defines the number of iterations and "Percentil" the percentile. With a "percentile" of 0.5, the standard forecast is performed.

fMC_Batch_HW_M_Time (Matrix1, Number, ResultOrder, Selection, NumIterations, Percentil): Marks the column or row selected under Matrix1 as a time series for a batch run using the Holt-Winters models (additive, multiplicative), with the data on a monthly basis and a forecast for the next few months. "Number" defines the number of forecast data, "ResultOrder" determines how the estimated historical and new forecast data should be displayed (0 = line by line, 1 = column

by column), with the forecast function selected based on the "Selection" criterion (0 = Akaide information criterion, 1 = sum of the residual squares). "NumIterations" defines the number of iterations and "Percentil" the percentile. With a "percentile" of 0.5, the standard forecast is performed.

fMC_Batch_HW_Q_Time (Matrix1, Number, ResultOrder, Selection, NumIterations, Percentil): Marks the column or row selected under Matrix1 as a time series for a batch run using the Holt-Winters models (additive, multiplicative), with the data on a quarterly basis and a forecast for the next few quarters. "Number" defines the number of forecast data, "ResultOrder" determines how the estimated historical and new forecast data should be displayed (0 = line by line, 1 = column by column), with the forecast function selected based on the "Selection" criterion (0 = Akaide information criterion, 1 = sum of the residual squares). "NumIterations" defines the number of iterations and "Percentil" the percentile. With a "percentile" of 0.5, the standard forecast is performed.

fMC_BootStrap (Matrix1, NumIt, Confidence, Percentile): Selects the column or row selected under Matrix1 and determines the confidence interval of the mean of that sample using NumIt of new samples (sampling with replacement). Result is a string with (mean value, lower range, upper range). If "Percentile" is specified, the confidence interval will be calculated with regard to the corresponding percentile.

fMC_BootStrapN(Name, NumIt, Confidence, Percentile): Selects the variable "Name" of a simulation and determines the confidence interval of the mean of that sample using NumIt of new samples (sampling with replacement). Result is a string with (mean value, lower range, upper range). If "Percentile" is specified, the confidence interval will be calculated with regard to the corresponding percentile.

fMC_BootStrapVector(Matrix1;NumIt;Percentile): Marks the column or row selected under Matrix1 as a sample and uses NumIt to determine new samples accordingly (drag and put back). If percentile is specified, the corresponding percentile of the respective sample is shown, otherwise the mean.

fMC_AD(Matrix1): Performs an Anderson-Darling test to check for normal distribution of the data selected under Matrix1. Returns True if the data is very likely to be normally distributed.

fMC_AD(Name»): Performs an Anderson-Darling test to check for normal distribution of the variable selected under "Name". Returns True if the data is very likely to be normally distributed.

fMC_Best(Optional Value): Returns the defined confidence level of the best case. If Value >= 0 and Value <= 1, the "best" case is determined dynamically based on Value.

fMC_Real(Optional Value): Returns the defined confidence level of the real case. If Value >= 0 and Value <= 1, the "real" case is determined dynamically based on Value.

fMC_Worst(Optional Value): Returns the defined confidence level of the worst case. If Value >= 0 and Value <= 1, the "worst" case is determined dynamically based on Value.

fMC_Iterations(): Returns the number of iterations

fMC_ExpGain(Name): Returns the expected average of all values above the expected gain of 0 from the distribution "Name".

fMC_ExpGainV(Name): Returns the variance of the values above the expected gain of 0 from the distribution "Name".

fMC_ExpLoss(Name): "Returns the expected average of all values below the expected gain of 0 from the distribution "Name".

fMC_ExpectedShortfall(Name, percentile): Calculates the expected shortfall from the distribution "Name" from the value-at-risk defined at "percentile".

fMC_Percentile(Name, Percentile): Returns the value at the k-percentile from the distribution "Name".

fMC_PercentileValue(Name, Value): Returns the percentile for a given "value" from the distribution "Name".

fMC_PercentileInterval(Name, Lower_Limit, Upper_Limit): Returns the confidence level (in %) from the distribution "Name", conditional on a lower ("Lower_Limit") and upper limit ("Upper_Limit").

fMC_CV(Name): Returns the coefficient of variation of the distribution "Name".

fMC_Semivariance(Name): Returns the semivariance of the distribution "Name".

fMC_SemivarianceL(Name, Limit): Returns the semivariance of the distribution "Name", whereby only values below Limit are considered.

fMC_SemivarianceR(Name, Limit): Returns the semivariance of the distribution "Name", whereby only values above Limit are considered.

fMC_Skewness(Name): Returns the skewness of the distribution "Name".

fMC_Kurtosis(Name): Returns the kurtosis of the distribution "Name".

fMC_StdS(Name): Returns the of sample tandard deviation the distribution "Name".

fMC_Variance(Name): Returns the sample variance of the distribution "Name".

fMC_Average(Name) / fMC_Mean(Name): Returns the average of the distribution "Name".

fMC_Min(Name): Returns the minimum of the distribution "Name".

fMC_Max(Name): Returns the maximum of the distribution "Name".

fMC_Sum(Name): Returns the sum of the distribution "Name".

fMC_FinancialValue (Name1, Name2, Value2, Diff, Value1): Calculates the value of event A (Name1), given that event B (Name2) has occurred. For event B, the search area contains +/-% of the limit entered under Diff. The result is the percentile value "Value1" of "Name1".

fMC_FinancialDTrans ("Initial", "Periods", "ActualYear", "Amount"): Determines the proportional share per year (via "Actual Year") of the value "Amount", starting from "Initial" and "Periods" (in year) time.

fMC_Mode (Name, UsedBins): Determines the most frequent value of the Name distribution on the basis of a histogram with the number of UsedBins columns.

fMC_ConditionPosValue (Name1, Value): Returns the position in the list of Name1 which corresponds to the specified value Value.

fMC_DSensitivity(Name1, Name2, DSensotype, NumberBins): Returns the dynamic sensitivity between two variables, whereby Name1 corresponds to the input and Name2 corresponds to the output variable. DSensotype indicates, which value should be returned; 0 = range, 1 = minimum, 2 = maximum. NumberBins specifies the number of bin used to make the calculation.

fMC_GetInputName(Pos): Returns the input variable at position Pos.

fMC_GetUserDefinedInputName(Pos): Returns the output variable at position Pos.

fMC_GetOuputName(Pos): Returns the user defined input variable at position Pos.

fMC_Density (Name, Value): Determines the density of the simulated distribution at point 'Value'.

fMC_Bin (Value): Specifies the number of columns in a histogram via "Value" (globally valid).

fMC_SimulaOn(Value): Workbook-Simulation modus: Determines whether random values (TRUE) or the expected value of the distribution function (FALSE) should be displayed after the worksheet / workbook has been recalculated. Note that the simulation results can only be evaluated with the value FALSE.

fMC_RegressionX(Name1, Name2, UseRepeatedMedianRegression): Calculates the linear regression slope of two variables. Name1 is x-vector, Name2 is y-vector. If UseRepeatedMedianRegression = 1, instead of a linear least squares estimate, a regression based on the 50% percentile is carried out, in which the absolute deviations and not the squared deviations are minimized.

fMC_RegressionY(Name1, Name2, UseRepeatedMedianRegression): Calculates the linear regression intercept of two variables. Name1 is x-vector, Name2 is y-vector. If UseRepeatedMedianRegression = 1, instead of a linear least squares estimate, a regression based on the 50% percentile is carried out, in which the absolute deviations and not the squared deviations are minimized.

fMC_BootStrapVectorLinearRegressionX(DataY, DataX, NumberNewSamples): Determines the slope of a simple linear regression model for the dependent variable DataY and the independent variable DataX using bootstrapping, where NumberNewSamples indicates the number of new samples. The result is a matrix with the dimension NumberNewSamplesx1. In combination with the Excel function "Percentile*" you can reference the corresponding percentile.

fMC_BootStrapVectorLinearRegressionY(DataY, DataX, NumberNewSamples): Determines the intercept of a simple linear regression model for the dependent variable DataY and the independent variable DataX using bootstrapping, where NumberNewSamples indicates the number of new samples. The result is a matrix with the dimension NumberNewSamplesx1. In combination with the Excel function "Percentile*" you can reference the corresponding percentile.

fMC_BootStrapVectorLinearRegressionRE(DataY, DataX, NumberNewSamples): Determines the standard error of regression of a simple linear regression model for the dependent variable DataY and the independent variable DataX using bootstrapping, where NumberNewSamples indicates the number of new samples. The result is a matrix with the dimension NumberNewSamplesx1. In combination with the Excel function "Percentile*" you can reference the corresponding percentile.

fMC_PredictiveRegressionLine(Name1, Name2, Value, Percentile): Determines the typical value at the point Value using the slope parameter Name1 and the intercept Name2, where Percentile is the corresponding percentile value of the typical forecast distribution. If Percentile is specified as 0 (by default), the expected value of the typical value is calculated.

fMC_PredictiveRegressionLineValue(Name1, Name2, Name3, Value, Percentile): Determines the predictive value of the regression line at the point «Value» using the slope parameter Name1, the intercept Name2 and the standard error of regression Name3, where Percentile indicates the corresponding percentile value of the forecast distribution. If Percentile is specified as 0 (by default), the expected value is displayed. With a symmetrical estimation error, this corresponds to the expected value of the typical value.

Formula used for Bayesian statistics

fMC_Bayes (Name1; Comparison1; Value1; Name2; Comparison2; Value2): Calculates the conditional probability for event A (Name1), given that event B (Name2) has arrived. Example:

`fmc_Bayes('NPV'; '>='; 1000; 'Project costs'; '>='; 100)` calculates the probability that NPV is greater (or equal) than 1000, given that the project costs are greater (or equal) than 100.

`fMC_BayesV (Name1; Comparison1; Value1; Name2; Comparison2; Value2; Optional Percentile)`: Calculates the average value induced by event A (Name1), given that event B (Name2) has occurred; If percentile is given, then the percentile value of the result is taken into account. Example: `fmc_BayesV ('NPV'; '>='; 1000; 'Project costs'; '>='; 100; 0.1)` calculates the 10% percentile value of NPV \geq 1000, given that the project costs are over 100. If Percentile is not specified, the average value of NPV above 1000 is calculated.

`fMC_Bayes4(NameX, CompareX, ValueX...)`: Calculates the probability for event A (Name1), given that event B, C, D to E (Name2, Name3, Name4, Name5) has occurred, see `fmc_Bayes`.

`fMC_Bayes4V (NameX, CompareX, ValueX..., Percentil)`: Calculates the value induced by event A (Name1), given that event B, C, D to E (Name2, Name3, Name4, Name5) has occurred, see `fmc_BayesV`.

`fmc_BayesClassifier (FeatureData, ClassData, PredictData, Testproportion, ModelSelection, ReturnConfusionMatrixTestSet, ReturnPredictionInterval, PredictionInterval, Iterations, HisBins, LowerLimit, UpperLimit, FeatureImpact, FeatureDescription, kfolds)`: Performs a classification of data based on Bayes' theorem. The starting point is pre-classified data (ClassData), which is described using the characteristics from FeatureData. The data to be reclassified is available in PredictData. The FeatureData are divided into a training set and a test set based on Testproportion. The classifier can be further specified via ModelSelection, with a distinction being made between the Normal/Bernoulli classifier or the empirical distribution - determined using kernel density estimators - or an automatic selection. If ReturnConfusionMatrixTestSet is specified as TRUE, the output is a truth matrix related to the test data. If FALSE is specified, the PredictData will be classified based on the entire data set (training and test set). If ReturnPredictionInterval is specified as TRUE, the credible interval is returned at PredictionInterval instead of a point forecast. Alternatively, the influence of the explanatory variables on the classification based on the test data can also be returned (FeatureImpact = TRUE). For the explanatory variables, the name can be taken from FeatureDescription, otherwise the name "Variable_" is added in front.

The number of samples from the respective distribution is determined using Iterations; HisBins, LowerLimit and UpperLimit specify the kernel density estimator in more detail. kfolds determines the number of attempts that are applied when ReturnConfusionMatrixTestSet = TRUE.

`fMC_BayesPredictionRegressionX (DataY, DataX, UseLogForY, NumNewSample, Percentile, UseRepeatedMedianRegression, kfolds)`: Determines the slope of a Bayesian simple linear regression model for the dependent variable DataY and the independent variable DataX using bootstrapping, where NumberNewSamples indicates the number of new samples (if NumNewSamples is not specified, 1,000 new samples are taken). With Percentile you enter the

desired percentile of the resulting distribution (if not specified, the mean value is returned). Using `UseRepeatedMedianRegression` you specify whether the least squares estimator (default = 0) or the absolute deviations (with `UseRepeatedMedianRegression` = 1) is applied based on the median (0.5 quantile) estimator. With `kfolds` (default = 1) you specify how many independent samples from the data set `DataY` and `DataX` (in pairs) are selected to limit the sample variability. The slope parameter is derived using a Bayesian method; for this purpose, a small sample is drawn from `DataY` and `DataX` (in pairs) to determine the prior knowledge of the regression parameter. If you select `UseLogForY` = 1 (default = 0), the data from `DataY` is taken into the calculation in logarithmic form.

`fMC_BayesPredictionRegressionY(DataY,DataX, UseLogForY, NumNewSample, Percentile, UseRepeatedMedianRegression, kfolds)`: Determines the intercept of a Bayesian simple linear regression model for the dependent variable `DataY` and the independent variable `DataX` using bootstrapping, where `NumberNewSamples` indicates the number of new samples (if `NumNewSamples` is not specified, 1,000 new samples are taken). With `Percentile` you enter the desired percentile of the resulting distribution (if not specified, the mean value is returned). Using `UseRepeatedMedianRegression` you specify whether the least squares estimator (default = 0) or the absolute deviations (with `UseRepeatedMedianRegression` = 1) is applied based on the median (0.5 quantile) estimator. With `kfolds` (default = 1) you specify how many independent samples from the data set `DataY` and `DataX` (in pairs) are selected to limit the sample variability. The intercept parameter is derived using a Bayesian method; for this purpose, a small sample is drawn from `DataY` and `DataX` (in pairs) to determine the prior knowledge of the regression parameter. If you select `UseLogForY` = 1 (default = 0), the data from `DataY` is taken into the calculation in logarithmic form.

`fMC_BayesPredictionRegressionRE(DataY,DataX, UseLogForY, NumNewSample, Percentile, UseRepeatedMedianRegression, kfolds)`: Determines the standard error of regression of a Bayesian simple linear regression model for the dependent variable `DataY` and the independent variable `DataX` using bootstrapping, where `NumberNewSamples` indicates the number of new samples (if `NumNewSamples` is not specified, 1,000 new samples are taken). With `Percentile` you enter the desired percentile of the resulting distribution (if not specified, the mean value is returned). Using `UseRepeatedMedianRegression` you specify whether the least squares estimator (default = 0) or the absolute deviations (with `UseRepeatedMedianRegression` = 1) is applied based on the median (0.5 quantile) estimator. If `UseRepeatedMedianRegression` = 1 is selected, the deviations between the result of the regression line and the measured value, otherwise the squared deviations are used as the standard error of regression ($\sqrt{\text{sum of the squared deviations} / (n-2)}$).

`fMC_BayesPredictionRegression(DataY, DataX, Value, Percentile, UseLogForY, NumNewSample, UseRepeatedMedianRegression, NumNewSample, kfolds)`: Determines a forecast using a linear Bayesian regression model and based on the data `DataY` for the dependent variable and `DataX` for

the independent variable. With Percentile you specify the desired percentile of the resulting distribution that includes the forecast (if not specified, the mean is returned), whereby the number NumNewSample is used to generate the forecast distribution. If you select UseLogForY = 1 (default = 0), the data from DataY is taken into the calculation in logarithmic form. Using UseRepeatedMedianRegression you specify whether the least squares estimator (default = 0) or the absolute deviations (with UseRepeatedMedianRegression = 1, only for simple linear regression) is applied based on the median (0.5 quantile) estimator. With kfolds (default = 1) you specify how many independent samples from the data set DataY and DataX (in pairs) are selected to limit the sample variability.

Note: To determine the “a-priori” distribution of the regression parameters (e.g. slope, intercept), a kernel density estimate is carried out using a random small sample of DataY and DataX (in pairs) in order to stabilize the regression model. The likelihood and its distribution are derived using the bootstrap method and processed using a kernel density estimator.

With kfolds (default = 1) you specify how many independent samples from the data set DataY and DataX (in pairs) are selected to limit the sample variability. The standard error of regression parameter is derived using a Bayesian method; for this purpose, a small sample is drawn from DataY and DataX (in pairs) to determine the prior knowledge of the regression parameter. If you select UseLogForY = 1 (default = 0), the data from DataY is taken into the calculation in logarithmic form.

fMC_BayesPredictionRegressionCIP(TrainDataY, TrainDataX, TestDataY, TestDataX, Percentile, UseLogForY, NumNewSample, UseRepeatedMedianRegression, NumNewSample, kfolds): Determines the coverage rate of newly measured data TestDataY and TestDataX using a Bayesian linear regression model based on the training data TrainDataY and TrainDataX. With Percentile you specify the coverage area you are looking for, which is used to determine the coverage rate (if not specified, the 95% percentile is used).

If you select UseLogForY = 1 (default = 0), the data from TrainDataY is taken into the calculation in logarithmic form. Using UseRepeatedMedianRegression you specify whether the least squares estimator (default = 0) or the absolute deviations (with UseRepeatedMedianRegression = 1, only for simple linear regression) is applied based on the median (0.5 quantile) estimator. With kfolds (default = 1) you specify how many independent samples from the data set DataY and DataX (in pairs) are selected to limit the sample variability.

Interpretation: You test a model with a percentile of 95% on 100 test data. The result is a 1 (100%); i.e. in this case the model correctly identified 95 of 100 test data within the prediction interval, 5 test cases fall outside the respective prediction interval. The result is therefore 95 (correctly recognized) / 95 (% percentile) = 1. If, however, all of 100 test data fell within the respective forecast interval, the result would be 100 / 0.95 \approx 1.05 (105%). If, with a percentile of 50% of 100 test data, 50 are recognized correctly, this results in a coverage rate of 1 (100%; 50 (correctly

recognized) / 50 (% percentile)). A good model should have a coverage rate close to 100% in the respective percentile. Coverage rates of less than 100% indicate that the regression parameters determined from the training set do not correctly anticipate the test data. This can be the case when the dispersion of the data increases. With coverage rates of over 100%, it can be assumed a priori that the dispersion will decrease compared to the training data.

`fMC_BayesPredictionRegressionMAE(TrainDataY, TrainDataX, TestDataY, TestDataX, UseLogForY, NumNewSample, UseRepeatedMedianRegression, NumNewSample, kfolds)`: Determines mean absolute error of newly measured data `TestDataY` and `TestDataX` using a Bayesian linear regression model based on the training data `TrainDataY` and `TrainDataX`.

If you select `UseLogForY = 1` (default = 0), the data from `TrainDataY` is taken into the calculation in logarithmic form. Using `UseRepeatedMedianRegression` you specify whether the least squares estimator (default = 0) or the absolute deviations (with `UseRepeatedMedianRegression = 1`, only for simple linear regression) is applied based on the median (0.5 quantile) estimator. With `kfolds` (default = 1) you specify how many independent samples from the data set `DataY` and `DataX` (in pairs) are selected to limit the sample variability.

Interpretation: The regression model, or rather the regression line, should provide forecasts that are as “good” as possible in terms of expected values. The greater the predicted value deviates from the actual value based on the test data, the larger the standard error and thus the mean absolute error. One reason may be that the model specification differs from the actual conditional distribution, or the regression model depends on many other variables that are not taken into account in the model. A large standard error in turn means that the credible interval for a point estimate becomes very wide and therefore the regression model becomes less suitable as a prediction model.

`fMC_BayesPredictionRegressionParameters(DataY, DataX, CI, FeatureDescription, UseLogForY, NumNewSample, UseRepeatedMedianRegression, NumNewSample, kfolds)`: Calculates the regression parameters a linear Bayesian regression model and based on the data `DataY` for the dependent variable and `DataX` for the independent variable. With `CI` you specify the desired credible interval of the parameters (if not specified, the mean is returned), whereby the number `NumNewSample` is used to generate the forecast distribution. `Feature Description` is a reference to the labels of the features (without intercept). If you select `UseLogForY = 1` (default = 0), the data from `DataY` is taken into the calculation in logarithmic form. Using `UseRepeatedMedianRegression` you specify whether the least squares estimator (default = 0) or the absolute deviations (with `UseRepeatedMedianRegression = 1`, only for simple linear regression) is applied based on the median (0.5 quantile) estimator. With `kfolds` (default = 1) you specify how many independent samples from the data set `DataY` and `DataX` (in pairs) are selected to limit the sample variability. Function returns a $(m+2) \times 4$ matrix, where `m` corresponds to the number of features used, additionally the intercept and the regression error are displayed.

Note: To determine the “a-priori” distribution of the regression parameters (e.g. slope, intercept), a kernel density estimate is carried out using a random small sample of DataY and DataX (in pairs) in order to stabilize the regression model. The likelihood and its distribution are derived using the bootstrap method and processed using a kernel density estimator.

With kfolds (default = 1) you specify how many independent samples from the data set DataY and DataX (in pairs) are selected to limit the sample variability. The standard error of regression parameter is derived using a Bayesian method; for this purpose, a small sample is drawn from DataY and DataX (in pairs) to determine the prior knowledge of the regression parameter. If you select UseLogForY = 1 (default = 0), the data from DataY is taken into the calculation in logarithmic form.

fMC_BayesPredictionRegressionResiduals(TrainDataY, TrainDataX, TestDataY, TestDataX, UseLogForY, NumNewSample, UseRepeatedMedianRegression, NumNewSample, kfolds): Calculates the residuals newly measured data TestDataY and TestDataX using a Bayesian linear regression model based on the training data TrainDataY and TrainDataX.

If you select UseLogForY = 1 (default = 0), the data from TrainDataY is taken into the calculation in logarithmic form. Using UseRepeatedMedianRegression you specify whether the least squares estimator (default = 0) or the absolute deviations (with UseRepeatedMedianRegression = 1, only for simple linear regression) is applied based on the median (0.5 quantile) estimator. With kfolds (default = 1) you specify how many independent samples from the data set DataY and DataX (in pairs) are selected to limit the sample variability.

Final notes: Bayesian statistics can be seen as a universal tool when “knowledge” or “assessments” are to be processed into a consistent conclusion using measured data.

Matrix-functions

fMC_PortfolioCombn(AValue): Calculates the number of portfolio combinations (without repetition, without order requirements) of n (“AValue”) titles.

fMC_PortfolioMatrix(aMatrix, AsInput, Comparex, Comparex, Comparevx): Determines the number of possible portfolios (matrix notation), where n titles are used. The result is a n x 2 matrix, where the first column contains the MC FLO variables and the second one the corresponding names of the portfolio. If AsInput is true (1), the variables are interpreted as input, otherwise as output variables.

fMC_PortfolioManualMatrix (aMatrix, cMatrix, isInput): Determines a portfolio based on a 0-1 matrix. Objects with a 1 are added to the portfolio. AMatrix corresponds to the names of the titles (as nx1 matrix) from which the portfolio is determined and cMatrix is an nxm matrix with 0-1 numbers. Make sure that the number of lines in the aMatrix matches the number of lines in the

cMatrix matrix. If isInput is true (1), the variables are interpreted as input, otherwise as output variables.

fMC_PortfolioCorrMatrix((aMatrix, FillRandom): Determines the correlation matrix of a portfolio of n titles. Result is a q X 1 Matrix. If FillRandom is true (1), then every correlation association is prebuild with a correlation coefficient between -0.3 and +0.3 (random number), otherwise the correlation coefficient is 0.

fMC_PortfolioCorrMatrixF (aMatrix, CorrelT): Determines all correlation combinations and their correlation coefficients from n variables of aMatrix. The result is a matrix with $n \times (n-1) / 2$ rows; CorrelT: Specifies the correlation coefficient: 0 = Pearson, otherwise Spearman.

fMC_SimpleMarkovChain(aMatrix, n): Model of a Markov chain to predict the final state after n periods, parting from a n x m matrix.

fMC_NBT(aMatrix, UseSummary): Checks whether data is Newcomb-Benford distributed using the Chi-squared test. Output is a 18x3 matrix, where the first 9 lines represent the frequencies of the numbers 1 to 9 (column 1 = measured; column 2 = Newcomb-Benford distribution).

fMC_LDT(aMatrix, UseSummary): Checks whether the last digit before the decimal separator is uniform distributed using the Chi-squared test. Output is a 19x3 matrix, where the first 10 lines represent the frequencies of the numbers 0 to 9 (column 1 = measured; column 2 = Uniform distribution).

fMC_BayesCombine (NumberX, ShareX, ReturnData, NumBin): Combines up to five different priori probability distributions into an aggregated priori distribution. ReturnData: Specifies whether the data of the combined distribution should be expressed directly as an nx1 matrix (TRUE) or as an nx3 histogram (FALSE). If as a histogram, the number of columns is used via NumBin.

fMC_BayesMCMC (Name1, Name2, BurnIn, Thinning, NumberIterations, ReturnData, NumBin, Number_binsDivide): Combines the priori probability distribution (Name1) with new data of the same type (Likelihood, Name2) for the posterior distribution using Markov-Chain Monte-Carlo (MCMC). "Number_binsDivide" specifies the number of bins used to calculate the MCMC, "BurnIn" corresponds to the number of initial hits that are not taken into account in the output matrix (burn in rate). "Thinning" indicates whether every (= 1) permissible hit or only every x hit (> 1) should be considered. With "NumberIterations" the sample size of the target distribution is specified. ReturnData indicates whether the data of the target distribution (posteriori distribution) should be expressed directly as an nx1 matrix (TRUE) or as an nx3 histogram (FALSE). If as a histogram, the number of columns is used via NumBin. As the proposal distribution a Gaussian distribution is used.

The `fMC_BayesMCMCA ()` function can be used in the same way. In this case, an asymmetrical uniform distribution is used as the proposal distribution.

The volatile versions `fMC_BayesMCMCV` and `fMC_BayesMCMCAV` can be used to enable a highly automated calculation of the posterior distribution. They ensure that whenever the model changes, the corresponding distributions are re-simulated. Use the volatile functions only to perform tests or when you know that the inputs are congruent.

`fMC_BayesPrediction (Name1, Name2, Count, Attempts, NumberIterations, ReturnData, NumBin)`: Determines a prediction based on the posterior distribution ("Name1") and the corresponding likelihood ("Name2"). "Count" corresponds to the number of occurrences, "Attempts" to the number of tolerable unsuccessful attempts. With "NumberIterations" the sample size of the target distribution is specified. ReturnData specifies whether the data of the target distribution (predict distribution) should be expressed directly as an nx1 matrix (TRUE) or as an nx3 histogram (FALSE). If as a histogram, the number of columns is used via NumBin.

`fMC_Kernel (Name, AdaptRate, Bins, Bandwith)`: Carries out an automatic kernel density estimation for variable "Name", based on a Gaussian kernel. The result is a nx1 matrix, where n corresponds to the number of iterations of the current simulation. Bins corresponds to the number of columns in the (internally used) histogram and AdaptRate to the adaptation rate (in %). It indicates by how many % the lowest value should be adjusted downwards, and the highest value should be adjusted upwards for the kernel density estimate. "Bandwith" corresponds to the smoothing parameter.

`fMC_Kernel2 (Data, LowerLimit, UpperLimit, Bandwith, Bins)`: Carries out a kernel density estimate based on a Gaussian kernel for the data selected in Excel. The result is nx2 matrix (first column contains the reference value and the second column the associated density), where n corresponds to the number of classes in the histogram. "LowerLimit" and "UpperLimit" indicate the range limits, "Bandwith" the bandwidth and "Bins" the number of columns of the histogram.

`fmc_BayesClassifier («FeatureData»; «ClassData»; «PredictData»; «Testproportion»; «ModelSelection»; «ReturnConfusionMatrixTestSet»; «ReturnPredictionInterval»; «PredictionInterval»; «Iterations»; «HisBins»; «LowerLimit»; «UpperLimit», «FeatureImpact», «FeatureDescription»)`: Performs a classification of data based on Bayes' theorem. The starting point is pre-classified data (ClassData), which is described based on the characteristics from FeatureData. The data to be reclassified is available in PredictData. The FeatureDates are divided into a training set and a test set based on TestProportion. Model selection can be used to further specify the classifier, distinguishing between the Normal/Bernoulli classifier, the empirical distribution – determined using kernel density estimators – or automatic selection. If ReturnConfusionMatrixTestSet is specified as TRUE, the output is a confusion matrix with respect to the test data. If FALSE, the PredictData are classified based on the entire data set (training and test set). If ReturnPredictionInterval is specified as TRUE, the credible interval at

PredictionInterval is returned instead of a point prediction. Alternatively, the influence of the explanatory variable on the classification based on the test data can be shown (FeatureImpact = TRUE). For the explanatory variables, the name can be taken from FeatureDescription, otherwise the designation "Variable_" is preceded.

Iterations are used to determine the number of samples from the respective distribution, HisBins, LowerLimit and UpperLimit specify the kernel density estimator in more detail. kfolds specifies the number of cross-validation attempts, which is applied if ReturnConfusionMatrixTestSet = TRUE.

Ribbon functions

You can select the data from a column and use the Ribbon button "Sort asc" (found under "Others") to display the numbers to the right of the selected column in ascending order. Alternatively, you can use the "Sort des" function to map a descending list. Please note that existing numbers and / or formulas are overwritten in both cases.

With "Excel-Check" your model will be analysed in a similar fashion as the detective of Excel does. However, MC FLO stores all the results in a separate workbook for easy comprehension.

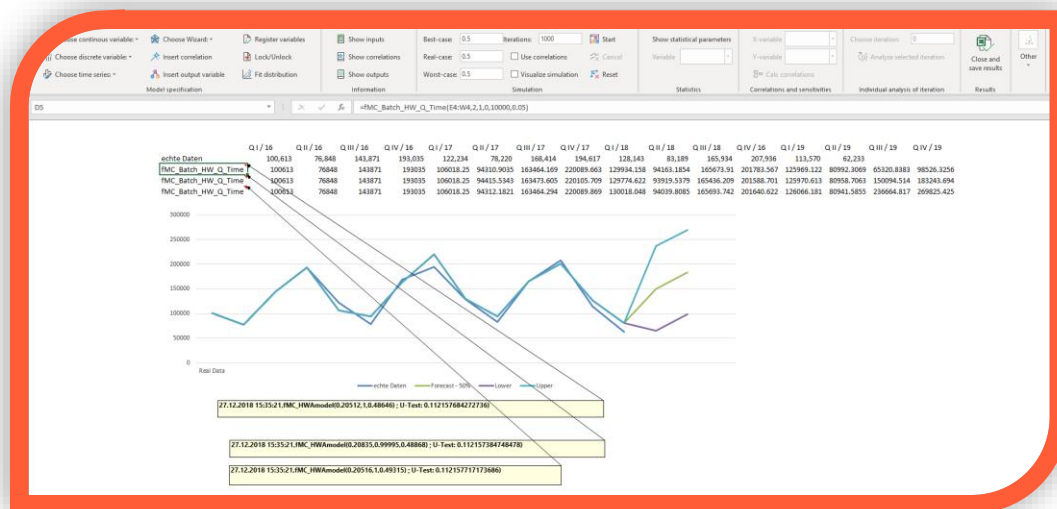
When activating «Use DIP», all output variables in the model are saved as an array when the results are exported. This data can then be used in Excel independently of a simulation. When a result workbook is opened, the output variables contained there are also imported into the current workbook as an array and stored using the notion "FLO_r_". Note that the variables are saved directly in the Excel workbook, the reference to the external file is retained. After reopening the file, an Excel message window may appear to update the file reference (for further details see: <https://youtu.be/qM9-OpwpXsA>).

Forecasting using time-series

Make forecasts for prices or volumes with just a few clicks – Select under “Other”, “Forecast-Tool” the appropriate time series model and introduce the relevant parameters. We recommend using previously our “Fit distribution” wizard to assign the historic data an appropriate time series function. Under “Number” you define the forecast periods (if “1” is selected, only one value will be forecasted). Under “ResultOrder” you can define if the values are displayed in rows or columns. Please note that existing values will be overridden.

By using the "Start Batch" function in your workbook, you can assign series (like revenues, costs etc.) defined in advance with the MC FLO formula "fMC_Batch_Time" and/or "fMC_Batch_HW*" to a time series process with one click and automatically generate a forecast, taking the current figures into account. The result is a series of numbers, which includes the estimated historical data and the forecast, and an entry as a comment. The comment contains the date of the batch start and a U-test based on the estimated historical data.

We recommend using the Holt-Winters model for ad-hoc forecasts before performing an exhaustive stochastic time series process.



Info: If you want to map the 90% confidence interval as a bound, define two "fMC_Batch_Time" or "fMC_Batch_HW *" batch functions, where the lower bound is defined in the first batch function using a percentile of 0.05 and the upper bound percentile is defined in second batch function using a percentile of 0.95. Note that MC FLO calculates in any case the empirical confidence level as measurement, using 50,000 iterations for each time point.

Under "Hide formulas" you can remove all MC FLO formulas (inputs, correlations, outputs) from the model and replace them with the corresponding numbers. This allows external people who have not installed MC FLO to view the model without restrictions. With "Load formulas" the

original formulas are reloaded. We recommend that you map the logic of the output variables in separate cells so that they remain comprehensible for third parties.

Frequently Asked Questions

What exactly does MC FLO and how is this implemented?

MC FLO provides a set of functions to determine the possible future states using simulation. Each simulation consists of several iterations. In each iteration, an allowable value is determined for pre-defined uncertain variables using a random generator. The functions are completely embedded in Microsoft Excel and can be used like any other function of Excel.

What differentiates MC FLO from other Excel-based simulation programs?

MC FLO is consistently designed for easy operation, lightweight, and is 100% -Excel using the .net technology. All diagrams and the statistics of the result workbook are created using pure Excel functions and can therefore be distributed to third parties which do not have MC FLO installed.

MC FLO allows you to load the individual iterations into the original model and thus to trace the logic and accuracy of the calculations. All input and output values are stored and displayed for maximum transparency!

You can load the MC FLO Excel-Add as you desire. The complete Excel integration makes MC FLO fast and secure.

Perform easily forecasts in your model with one click: all marked series (revenues, costs etc.) will be assigned automatically to a time series process and a forecast will be done considering the required confidence level.

After a calculation with more than 1,000 variables (input or output variables), it is not possible to examine an individual variable from the result workbook. Why not?

MC FLO uses Ribbon controls to display the results. These are currently limited to 1,000 entries. However, you can use the usual Excel functions to access all the results. Some customers want a higher number of supported variables, but they refrain from using the analysis tools implemented in MC FLO.

On some models, the calculation with several thousand iterations is much faster than using another model, although the number of variables and the size are comparable. Sometimes the difference can be several minutes. How is this possible?

Calculation speed depends on the used Excel formulas. If you use "volatile" functions of Excel, such as "VLOOKUP" or "Index", the processing speed of Excel is throttled dramatically. Do not use these functions if a quick result is important.

When estimating time series, it is noticeable that the estimated data points and the real data are almost congruent and therefore a U-test near zero results. This is hardly realistic. What is MC FLO doing here?

In assigning the random components to a time series process, a minimization of the residual sum of squares is carried out with the "nearest neighbour" method. If you find significant differences between the forecasted and actual historical time series data, this can be taken as an indication that the time series process identified does not reflect the actual course. Then try to use a different time series. Note that minimizing the sum of squares can also lead to false conclusions. When examining a time series with trend, you may be presented with a "correct" estimate. Make sure that only in a Wiener process a trend may occur in the model, otherwise false predictions will result. We recommend the Holt-Winters models to make forecasts for non-stationary series.

In the graphical environment it is not possible to show very large numbers (+/- 1e96), why not?

MC FLO uses Microsoft's .net library to show the results. In the current .net specification it is not possible to display values greater than about + - / 1e96. However, both in the Excel cells and in the result file the values will continue to be displayed appropriately.

When using Bayesian statistics, the combination of "priori" knowledge and the data, i.e. the so-called "a posterior" distribution, appears coarse-grained and less precise in comparison to other programs, such as the statistical program "R". What is the reason?

Statistics programs very often use analytical - exact - methods to derive the results. MC FLO, on the other hand, is consistently geared towards simulations. Parameters such as mean value or density of a distribution at point "x" are therefore always derived only with reference to the simulated results. As a result, the Monte Carlo Markov Chain method used by MC FLO for determining the «a posterior» distribution is less accurate than the analytical programs. For most practical questions, however, MC FLO should derive a sufficiently precise approximation. To increase the accuracy, we recommend increasing the number of iterations of a simulation.

Specification of MC FLO

Number of input variables:	Up to 65 536
Number of correlations:	Up to 32 768
Number of output variables:	Up to 65 536
Number of worksheets supported within the workbook	Only constrained by Excel
Number of supported iterations:	Only constrained by memory
Supported continuous distributions	
Uniform distribution	
Triangle distribution	
PERT distribution	
Normal distribution	
Lognormal distribution	
Beta distribution	
Gumbel distribution	
Gamma distribution	
Exponential distribution	
Truncated Normal distribution	
Weibull distribution	
Chi-2 distribution	
Student-T distribution	
Pareto distribution	
F distribution	
ArcSin distribution	
Cauchy distribution	
Fréchet distribution	
Logistic distribution	
Custom defined	
Supported discrete distributions	
Uniform distribution	
Bernoulli distribution	
Binomial distribution	
Poisson distribution	
Pascal distribution (Negative Binomial distribution)	
Geometric distribution	
Hyper geometric distribution	
Custom defined	
Supported time series	
AR(1) – Autoregressive model	
MA(1) – Moving average model	
ARMA(1,1) – Autoregressive-moving-average model	
ARCH(1) - Autoregressive conditional heteroskedasticity	
WP – Wiener process (geometric Brownian motion)	

License Agreement

1. Scope

1. The following license and usage conditions apply to the installation and use of the present simulation software MC FLO (hereinafter referred to as "Software"). For all orders and other business transactions in connection with our products and services, our General Terms and Conditions of Business apply in their current version. MC FLOsim is the sole owner of the software and grants you a license to use the software. If you do not agree to or violate the terms, the license expires, and you are no longer entitled to use the software.

2. License and Terms of Use

1. The software and any patents, copyrights, trademarks, trade names, know-how and developments relating to it are the property of MC FLOsim and are protected by applicable law. As a result, you are only entitled to use the software, which is expressly permitted by this license and its usage conditions. You are not authorized to modify, reverse engineer, develop, decompile, disassemble, distribute, lend or lease the Software.
2. We grant you a revocable, personal license (using a separate license key) for the use of the software in accordance with these license and usage conditions. The license is granted for a limited period (according to the order) or for an unlimited period.
3. The software is provided to you in its current state. We do not warrant that the software will meet all your requirements and expectations, nor that the execution of simulation calculations using the software will be uninterrupted and error-free.
4. We disclaim all liability for loss, loss of profits, tangible and intangible losses and for indirect damages such as lost or destroyed data, non-observing schedules, third party claims or other special, incidental or consequential damages resulting from the installation and use of the software whether or not it was appropriate.
5. MC FLOsim is entitled to amend or supplement these license and usage conditions at any time and to notify you of this in a suitable way. By continuing to use the software, you agree that the changes and additions will be binding upon you as soon as you have received them. Otherwise, the license expires, and you are no longer entitled to use the software.
6. The present license and usage conditions are subject to Swiss law.