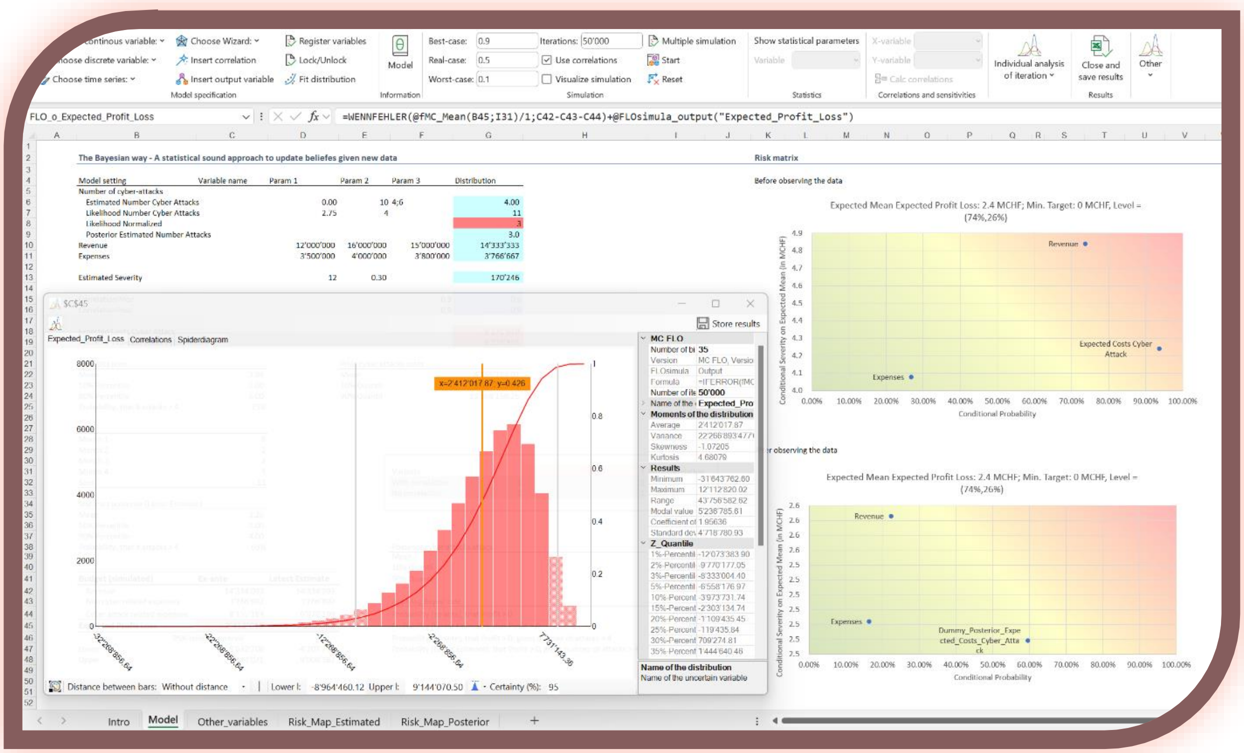
The Conditional Risk matrix –
Making better decisions. A Data Story.



support@mcflosim.ch

Foreword

Decisions are crucial. Some are quite complex in nature. Monte Carlo simulations in conjunction with a conditional risk-map maintain an overview in such situations, thereby facilitating better decisions. In the following, we will show you how to effortlessly implement better decisions within Microsoft Excel™ and the associated add-in MC FLO.

Introduction

Just as the layout of the keyboard arrangement of a typewriter aimed to break the flow of typing as little as possible, the task of the risk matrix is to represent the dangers within a very short time. Nevertheless, in the early days of the risk matrix, neither the technique of Monte Carlo simulation nor Bayesian statistics were widely prevalent as decision-making tools under uncertainty, or it took too much time to calculate all the "risks" consistently, thereby aggregated. According to economists, the continued widespread use of the risk matrix is due to path dependencies.

Critiques of the risk matrix are numerous (see also our blog post)[1], but the alternative quantitative representation using a tornado graph (which is based on a correlation analysis) is victim of the mentioned path dependency.

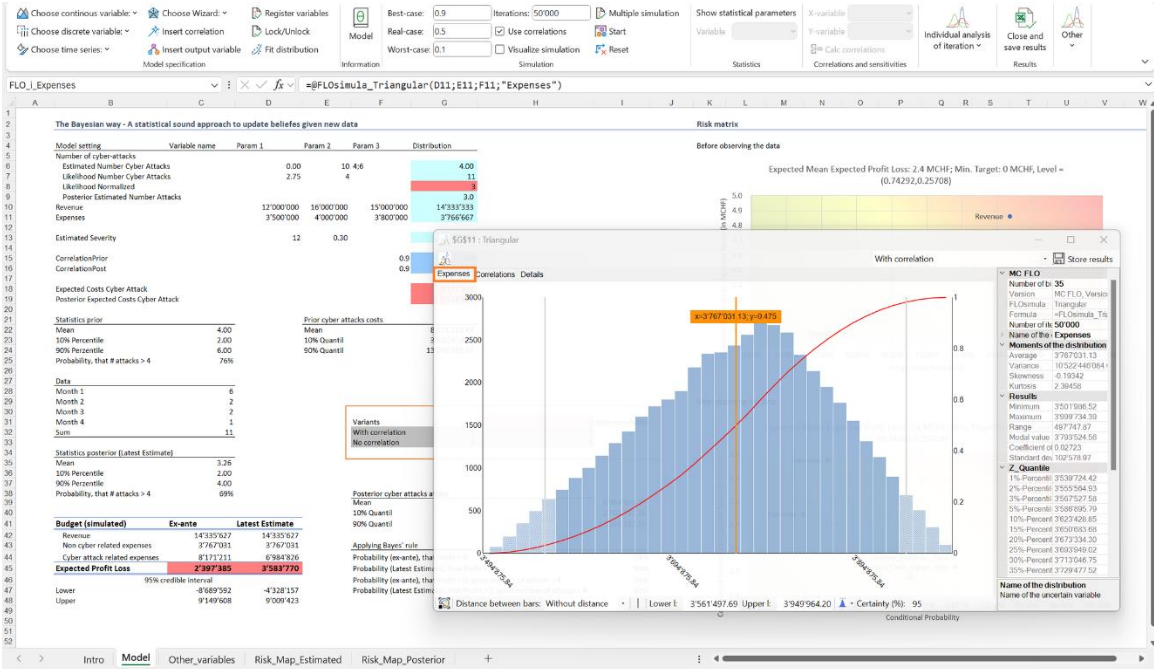
For those people who appreciate quantitative analysis, but still use risk matrices for communication, MC FLO in conjunction with Microsoft Excel™ can do this without effort.

Objectives

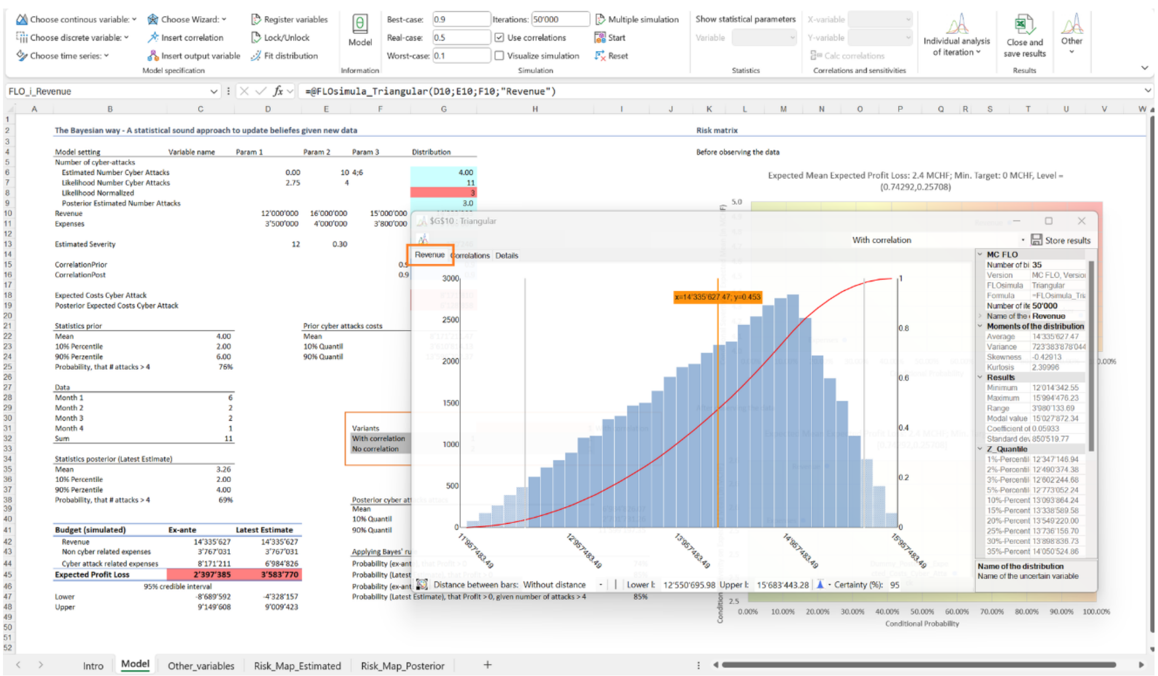
The aim of a quantitative analysis using Monte Carlo simulation and Bayesian statistics is to obtain an objectively justifiable decision recommendation, considering uncertainty. In the context of classic corporate planning, the primary goal ("objective") is to generate an economic profit, which is subsumed in the following example as the "Expected Profit Loss" or EBIT > 0.

The classical planning process often separates drivers: The expenses and revenues which the company can directly control, as quantities produced and labor costs, are quantified and accounted in the planning tools. Indirect artefacts that are not directly under control, such as cyber-attacks, are at best modelled as a scenario, at worst delegated as part of a risk analysis using a risk matrix, and thus excluded from **an integrated quantitative** analysis. It gets even worse – because inefficient – when the drivers are accounted in both corporate planning and risk analysis without inter-action.

Let's take a closer look at the following planning problem. The company operates under uncertainty and therefore prepares the revenue and expenses stream using random variables. We assume that both quantities were generated using driver models (on an annual basis).

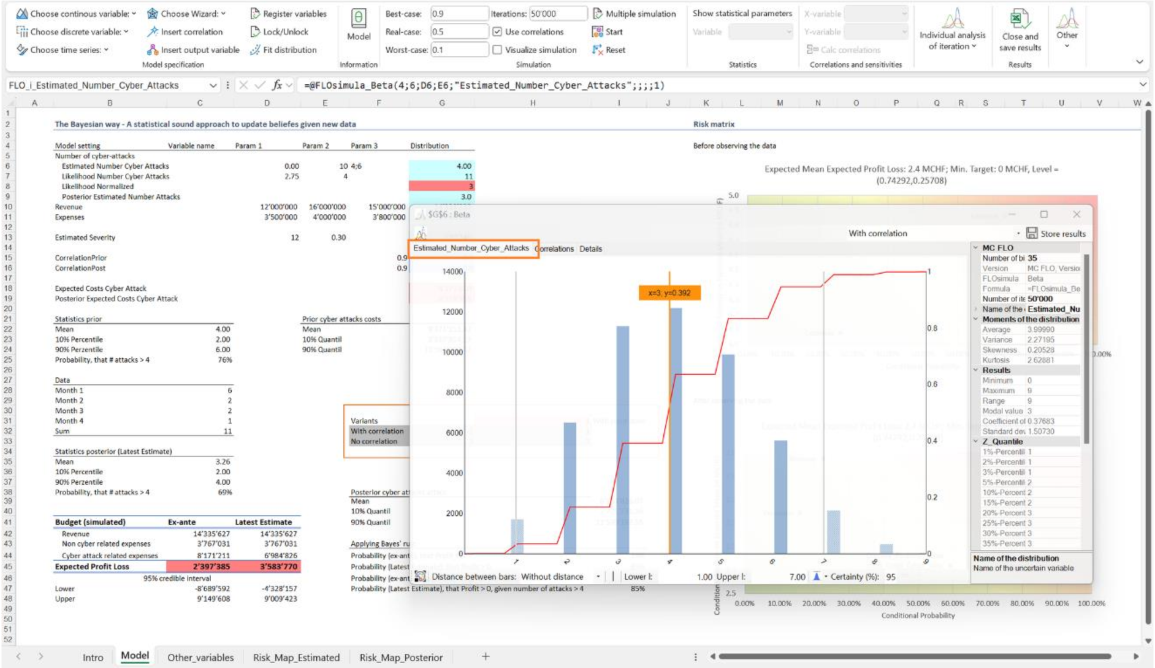


Picture 1: Uncertainty regarding expenses

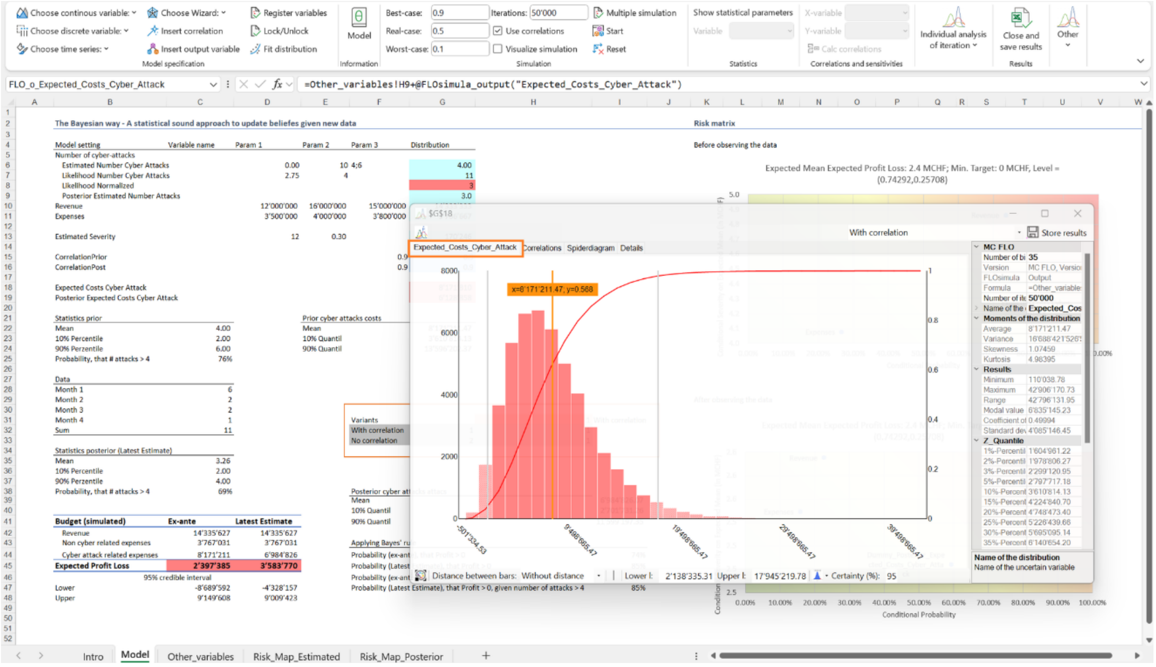


Picture 2: Uncertainty regarding revenue

For the following analysis, however, we will focus on the "risk" of a cyber-attack. To ensure consistency in the planning process with all other driver variables, the company additionally defined random variables concerning the number of cyber-attacks per month and the severity per cyber-attack. These values can be obtained through experience or, in the absence of own data, based on data from comparable companies. By combining severity with the number of cases per month, the possible expenditure per year can be inferred via the concept of convolution.

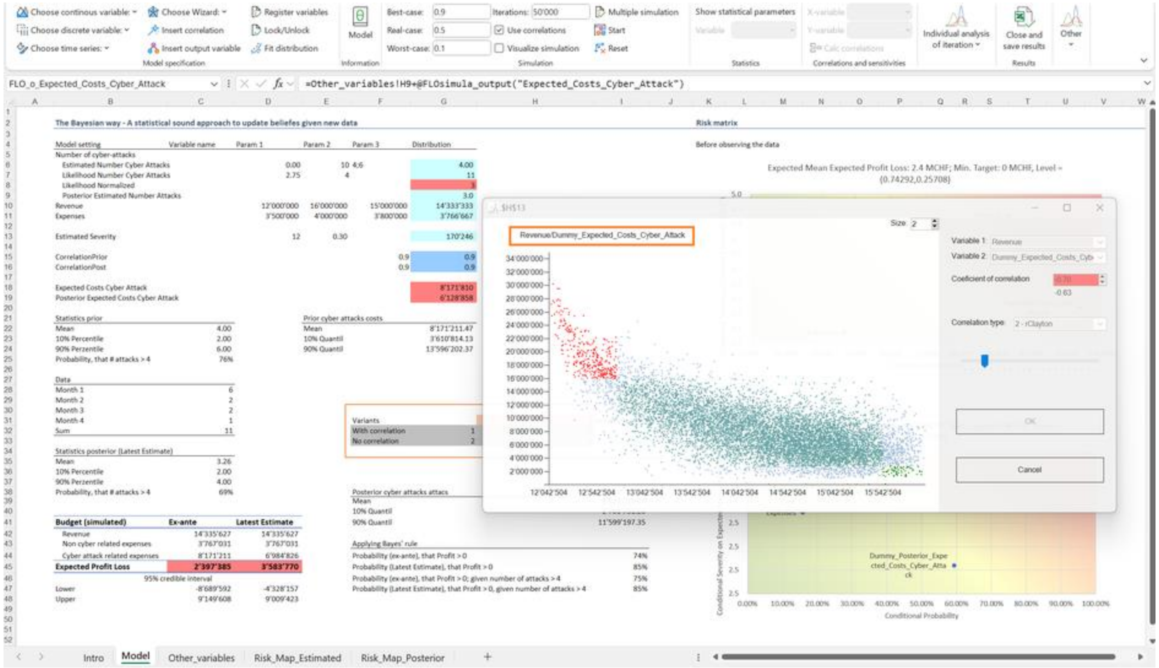


Picture 3: Distribution of number of cyber-attacks per month



Picture 4: Distribution of compound severity

The driver variables (such as number of attacks, severity) are usually not independent, the degree of dependency is expressed in the simplest case by a correlation measure [2]. Among other things, the model assumes that there is a negative correlation between the severity caused by cyber-attacks and the revenue. The structure of the correlation is described by a Clayton copula.



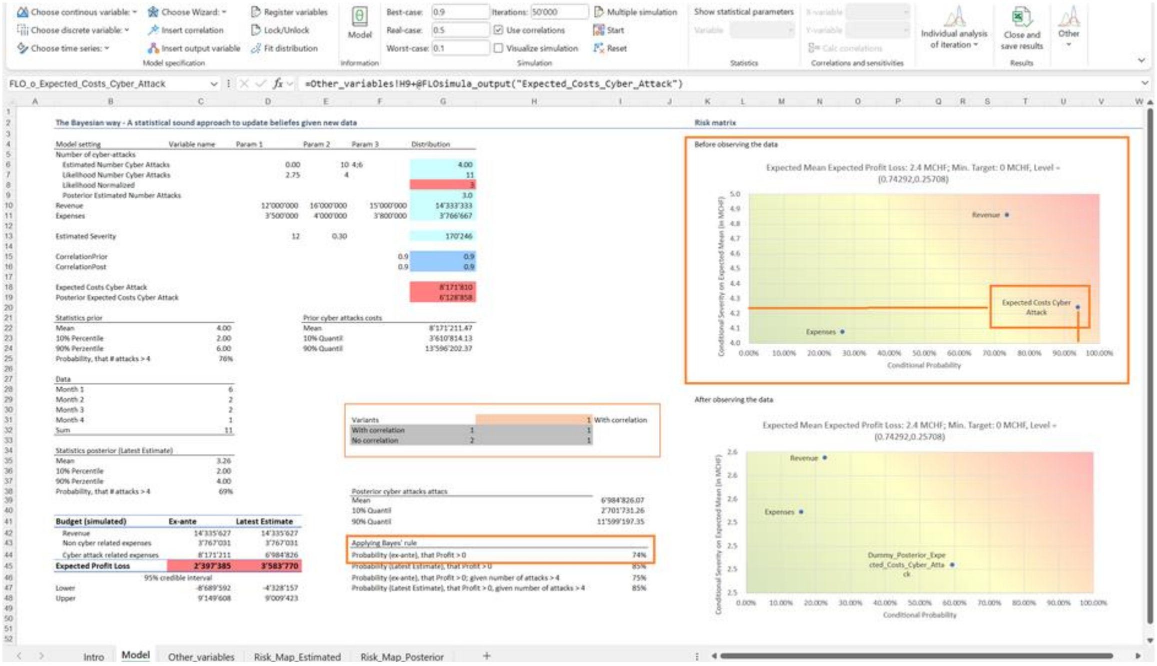
Picture 5: Modelling inter-relationships

Once the model has been set up and the simulation has been carried out, it is relevant for the decision-makers to determine which drivers have the highest impact on the objective (in this case EBIT, "Expected Profit Loss", with an estimated value of 2.4 MCHF).

This brings us to the conditional risk matrix. All drivers are placed in the known grid (probability of occurrence and severity) but - according to Bayes' theorem - **conditional probabilities and conditional severities regarding the impact on objectives** (target) are used. But first things first.

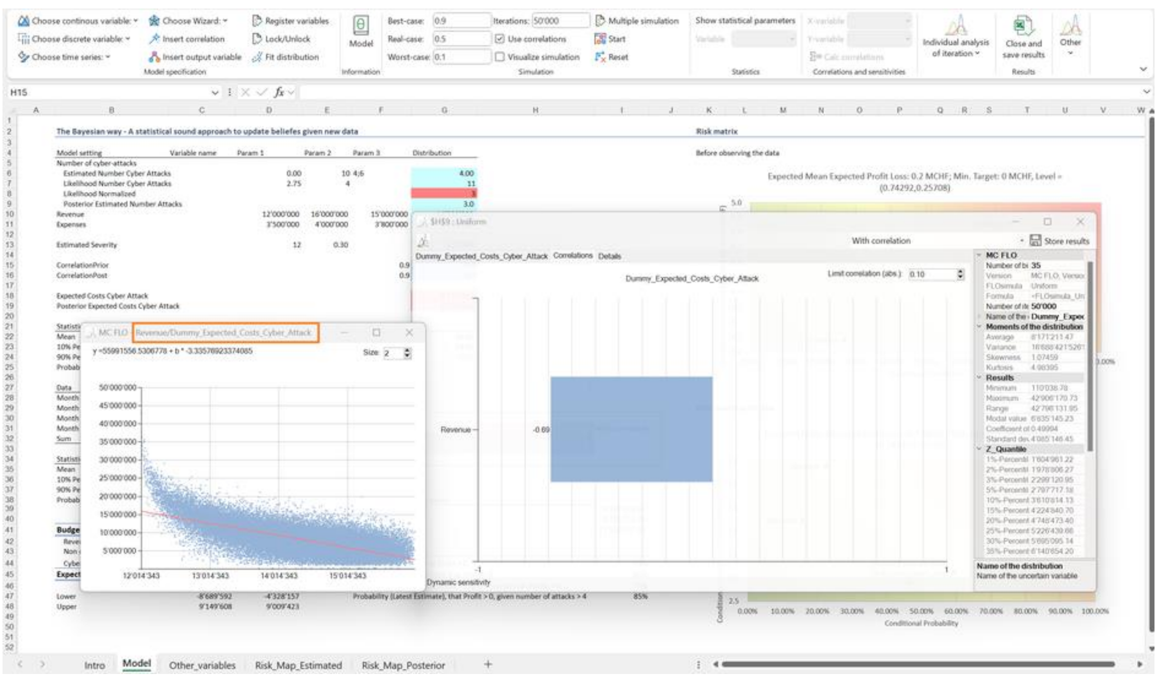
In the long run a target of EBIT >= 0 is mandatory, so drivers that induce an EBIT of < 0 are to be regarded as risky. The simulation reveals that in about 26% of all simulated cases an EBIT < 0 is obtained. This 26% threshold is the "value-at-risk" at EBIT = 0. With the built-in Bayes formulas in MC FLO a conditional risk-matrix using the traditional charting capabilities of Microsoft Excel™ can be built from scratch. On the x-axis of the conditional risk matrix the conditional probability on the target variable at the "value-at-risk" level is displayed. It answers the following question: "What is the probability that EBIT is < 0, **given** that the driver variable "a" surpasses the "value-at-risk" threshold?". The severity can be seen on the y-axis of the risk matrix. It answers the question: "What is the average loss of the target variable (EBIT < 0), **given** that the driver variable "a" surpasses the "value-at-risk" threshold?".

For the driver variable "Expected Costs Cyber Attack", the probability is approximately 94% based on Bayes' theorem. There is a 94% probability that EBIT will be < 0, **given** that the driver variable "Expected Costs Cyber Attack" has surpassed the "value-at-risk" threshold at 26%. In this case, EBIT will be around -4.2 MCHF.



Picture 6: Conditional risk matrix before observing new data

As can be seen, the variable "Expected Costs Cyber Attack" is the decisive driver variable, and not the expenses from the operating business. Also critical is the revenue stream, which follows from the modelled dependence between revenue and cyber-attack severity.



Picture 7: Correlation between cyber-attack costs and revenue

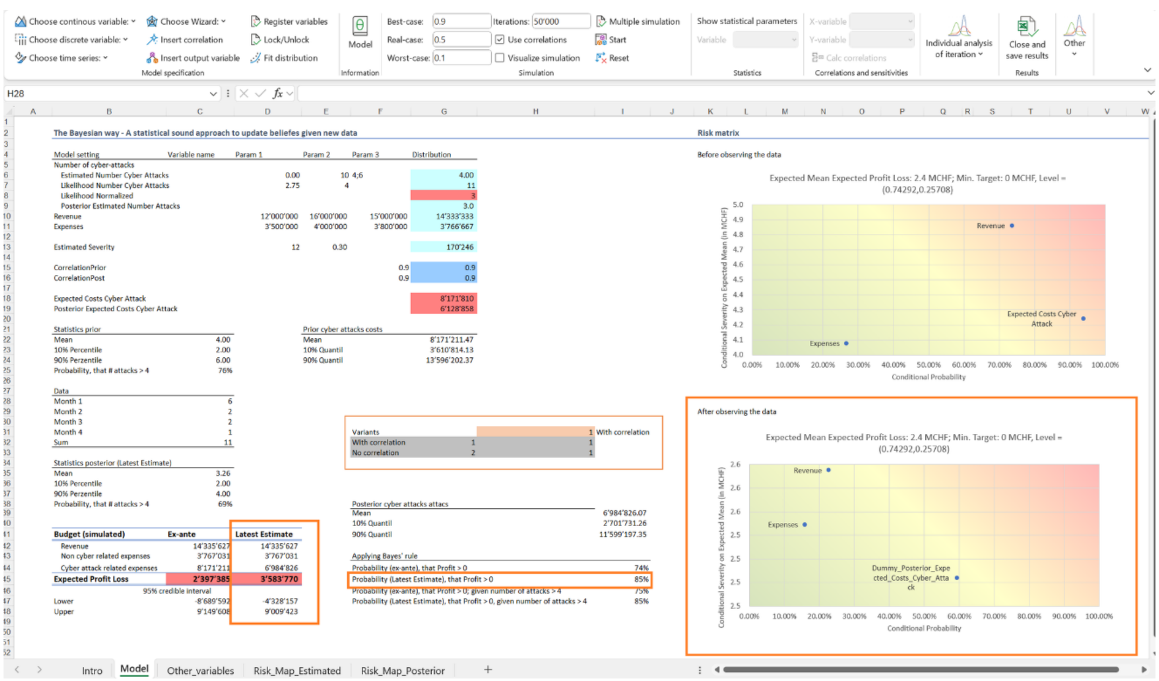
Given the results from the conditional risk matrix, the company can now derive measures and, for example, consider a variant that would increase investment in the IT infrastructure in combination with a decreased expectation of the number of cyber-attacks. After all, the company can accept the risks described above and follow a "wait-and-see" strategy.

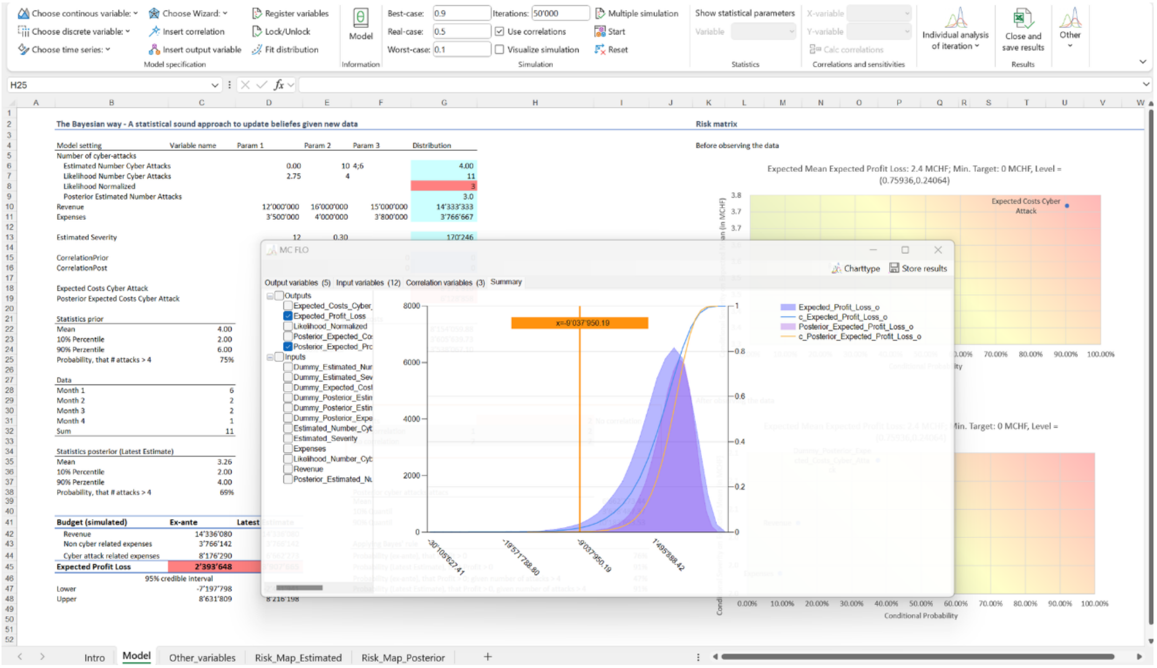
Despite the advantages of a simulation over mere point estimates and the coherent presentation of the outcome in the conditional risk matrix, every model must measure itself against reality and learn from it. This is where the Bayesian view of statistics helps decision-makers to refine their beliefs. To put it simply, let's assume that the uncertainty on sales and expense, the severities in the event of a cyber-attack as well as the assumed correlations are unchanged compared to the status before the data was measured, however, the number of cyber-attacks has to be adapted on the basis of the observed new data (see cells C28:C31).

If it is assumed that every cyber-attack follows a Poisson distribution. The posterior distribution of the cyber-attacks weights the initial beliefs – using Bayes' theorem - with the new data to obtain a new forecast for the end of the year ("Latest Estimate"). While the assumed number of cyber-attacks per month was 4 before the new data was measured, this drops to about 3 attacks per month based on the measured data from January-April. As the observed data is now part of the forecast, uncertainty decreases.

Ultimately, the probability of measuring EBIT < 0 at the end of the year has fallen to around 15%, resulting in a new estimated EBIT of 3.6 MCHF per EoY. As a result, the driver variables take an adjusted position in the conditional risk matrix. Given the new observed data, the conditional risk matrix can show the relevant driver variables and thus guide the decision makers.

With each new data set, the prediction can be refined and adjusted consistently. It is therefore the instrument that puts the combination of initial beliefs and data on a scientific basis.





[1] See also [The Risk Matrix Approach: Strengths and Limitations (garp.org)]

[2] To determine the correlation, the (linear) Pearson correlation coefficient can be used, alternatively the (non-linear) Spearman Rank correlation coefficient. Copulas are used to describe the structure of dependency, such as the Gauss Copula or those named Clayton, Frank, Gumbel and others. A completely On the other hand, *transinformation*, which draws on entropy, is not yet sufficiently anchored.

Remark: the calculations were carried out on Microsoft Excel for Office and with the Monte Carlo Add-In MC FLO. MC FLO is available in German, English and Spanish.

Training video of MC FLO and Bayesian analysis: https://youtu.be/dNz9YTqyK14

Please let us know if you have any questions.

www.mcflosim.ch

LinkedIn: https://www.linkedin.com/company/14018968/