Best Practices and Lessons Learnt of Probabilistic Schedule Analysis for Oil and Gas Field Development Projects

Master Thesis
by
DI Matthias Katschnig

Submitted to the
Chair of Economic- and Business Management
Montanuniversitaet Leoben

Leoben, January 2012

"Whenever we can replace human judgement by a formula, we should at least consider it."

(Daniel Kahneman, 2011)

EIDESSTATTLICHE ERKLÄRUNG

Ich erkläre an Eides statt, dass ich diese Arbeit selbstständig verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt und mich auch sonst keiner unerlaubten Hilfsmittel bedient habe.

AFFIDAVIT

I declare in lieu of oath, that I wrote this thesis and performed the associated research myself, using only literature cited in this volume.

Ort/Datum

Unterschrift

Danksagung

Mein Dank gilt o.Univ.Prof.Dr. Hubert Biedermann für die Ermöglichung, diese Arbeit am Lehrstuhl für Wirtschafts- und Betriebswissenschaften zu schreiben.

Im Besonderen will ich Dr. Bernhard Brunnsteiner (AVL LIST GmbH), Dr. Stephan Staber (OMV) und Dr. Stefan Massl (OMV) danken, die mir jederzeit mit Rat und Tat zur Seiten standen.

Acknowledgements

At first I thank o.Univ.Prof.Dr. Hubert Biedermann for giving me the opportunity to write this diploma thesis at the Chair of Economic- and Business Management (Montanuniversitaet Leoben).

I thank in particular Dr. Bernhard Brunnsteiner (Institute for Economic- and Business Management), Dr. Stephan Staber (OMV) and Dr. Stefan Massl (OMV), who readily/willingly provided information and advice at any time.

Kurzfassung

Es ist ein Axiom des wirtschaftlichen Handelns, dass die Zukunft ungewiss ist. Diese Tatsache ist besonders der Ölindustrie bekannt, die mit vielen unbekannten Variablen und Risiken umgehen muss. Besonders die Projektplanung ist von diesem Problem betroffen, da es ihre Aufgabe ist, zukünftige Ereignisse zuverlässig vorherzusehen.
Daher ist die Analyse von Projektplänen ein wichtiger Bestandteil des Projektmanagements, um Risiken im Zeit- und Kostenplan angemessen zu begegnen. Dieser Vorgang wird als "Probabilistic Schedule Analysis" (PSA) bezeichnet. Es handelt sich hierbei um eine probabilistische Analyse der Zeitpläne, also eine Abschätzung möglicher Projektverläufe mit Hilfe der Wahrscheinlichkeitsrechnung. Schlüsselfragen sind die Schätzung der Projektlaufzeiten (und Kosten), das Bestimmen die wichtigsten Risiken und deren Einflussfaktoren und die Entwicklung von Ausweich- oder Milderungsmaßnahmen.
Der Fokus dieser Arbeit lag einerseits auf der theoretischen Herleitung (Literaturstudie, Interviews) von optimalen PSA-Strategien und andererseits auf der praktischen Erprobung der Strategien mit realen Projektplänen (MS Project) und der Risikoanalyse-Software @Risk.
Die Ergebnisse des theoretischen Teils besagen, dass die probabilistische Zeitplan-Analyse einem deterministischen Ansatz überlegen ist, weil sie Unsicherheit und damit reale Verhältnisse widerspiegelt. In vielen Fällen wurde und wird dabei die Monte-Carlo-Simulation angewendet. Neben Monte-Carlo ist die einfachere PERT-Analyse zwar eine rohe, aber schnelle Methode, um mögliche Zeitpläne abschätzen zu können. Weiters ist die Abschätzung valider Eingangsdaten wesentlich wichtiger als die Suche nach einer „magischen Eingangsverteilung“. Somit sollte die Abschätzung von Eingangsdaten mit einer elaborierten Methodik geschehen. Die Interviews mit Experten bei der OMV bestätigten im Grunde die Literaturrecherche. Zusätzlich wurde aber zusätzlich eine transparente und klar kommuniziert PSA-Richtlinie gefordert und PSA-Ergebnisse müssen für die Präsentation richtig aufbereitet werden.
Die Simulationsergebnisse des praktischen Teils ergaben, dass die Schlüsselfaktoren: Form der Eingangsverteilung, Mittelwert und Streuung der Eingangsverteilung, der zentrale Grenzwertsatz, Aktivitätseinschränkungen und Aktivitätskorrelationen für empfindliche Verschiebungen im Zeitplan verantwortlich sind. Diese Faktoren haben besondere Auswirkungen auf die Fertigstellung des Projekts und müssen sorgfältig während der Planung überprüft werden.
Alle wichtigen Erkenntnisse dieser Arbeit sind in einem Flussdiagramm (vgl. Abbildung 92) zusammengefasst, das eine adäquate PSA-Vorgangsweise beschreibt.
Insgesamt sagt die Arbeit aus, dass eine gut ausgeführte PSA eine Grundvorrausetzung ist, um reale Verhältnisse in die Projekt-Zeitpläne einzuarbeiten und so zuverlässige Daten für den Projektverlauf, die Projektkosten etc. zu bekommen. Nur so ist es möglich, Zeitverzögerungen und Kostentreiber wirkungsvoll zu identifizieren.

Abstract

It is an axiom of economically act that future is uncertain. This fact is known in particular by the oilfield industry which naturally is an industry that faces many uncertainties (unknown variables) and risks. Especially project scheduling is concerned by this problem, because its job is to make future events manageable.
Thus the analysis of project schedules is an important part of project management in order to manage project schedule risk adequately. This procedure is called probability schedule analysis (PSA). Key issues of PSA are estimating the project durations (and consecutively costs), finding the most critical risks and important impact factors and determining mitigating or avoiding strategies.
Therefore, the core of this study was to collect best practices of probabilistic schedule analysis from the theoretical (literature) view and conduct some PSA examples with @Risk and MS Project from the practical point of view. To do so this work has four base chapters (Chapter 1: theoretical probability schedule analysis, Chapter 2: input data estimation, Chapter 3: company survey, Chapter 4: applied probability schedule analysis) studying the most important fields of today`s PSA.
The literature outcome says that probabilistic schedule analysis is superior to a deterministic approach by taking uncertainty and therefore reality into account. In many cases Monte Carlo simulation was and is used. On the other hand the more straightforward PERT analysis can give you a raw and quick approach to possible values. Furthermore collecting the right input data is more crucial than finding a "magical input distribution". Input data estimation for your schedule analysis must include diverse groups of estimators (controlled by some estimation workshop) which are aware of black swans, cognitive biases, heuristics and the statistical concept of crowd wisdom. The company survey basically confirms the literature output, moreover a transparent and clearly communicated PSA guideline is demanded and PSA outcomes must be well documented for presentation.
On the practical side @Risk simulation results state that the main key factors are input distribution shape, input distribution mean and spread, the Central Limit Theorem, task constraints and task correlations. All of them have specific impacts on the completion date of the project and should be carefully revisited during scheduling.
All results mentioned above are implicated in a flowchart (cp. Figure 92) suggesting a way of doing PSA. Firstly an estimation workshop takes place where probabilistic input data is generated. Ideally this is combined with producing the deterministic baseline schedule, thus the same people can work on both parts. As result a probabilistic schedule is achieved that can enter the Monte Carlo simulation stage. Schedule risk drivers are now (hopefully) detected and schedule can be optimised. One important process step is filling, checking and maintaining a Monte Carlo input value database. This database builds the foundation for subsequent estimation workshop and will be an assessment reference in the following process steps.
Overall the true intent of PSA is encompassing all uncertainties to elicit confidence intervals in order to make better decisions and highlight important duration and cost drivers.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

CLT	Central Limit Theorem
Cp.	Compare
Etc.	Et cetera
MC	Monte Carlo (Method)
PSA	Probabilistic schedule analysis
→ *term*	You can look up this term in Chapter 7: Glossary.
CDF	Cumulative Density Function
PDF	Probability Density Function

Aims of this Study

In general, OMV is interested in improving its knowledge base on conducting a probability schedule analysis (PSA). To cope with those allowances PSA has to deliver the following results:

- Identifying sensitive tasks!
- Finding realistic project completion dates!
- Testing the robustness of existing deterministic schedules!

Therefore it was the central goal of this study to collect best practices of probabilistic schedule analysis from the theoretical and practical point of view. To do so the following sub goals were broken down:

- Investigation of PSA practices from relevant literature.
- Identification of best practices and lessons learnt.
- Practical analysis of real schedules of field development projects.
- Short guideline for conducting a PSA.

Furthermore, the basic structure of this work has a theory and a practical part:

Practical Part	**Theoretical Part**
Reality check: 2 genuine OMV schedules (based on MS Project 2010) and several dummy schedules will be observed and processed with the risk analysis software @Risk.	Literature check: a closer look at papers, books, articles dealing with PSA and estimation approaches in general.
Stakeholder opinions: company survey.	Finding critical points of PSA.
Knowledge management: short guideline for conducting a PSA.	Dealing with critical points of PSA.
Expected results: • Simulation results • Best practices • Lessons learnt • Important factors and specifics	Expected results: • PSA basics • Best practices • Lessons learnt • List of best literature

1 Theoretical Probabilistic Schedule Analysis

1.1 Basics of Deterministic and Probabilistic Scheduling

1.1.1 Principle of a Deterministic Schedule Analysis

Basic deterministic schedule estimating and analysis contains the subsequent points: construct a logical task network, determine a best estimate of every task duration, compare these estimates to find network`s critical path (→ *critical path*), sum all best estimates on this path and define the sum as the overall duration of the project.
The next chapter will describe two techniques that are very common to support the deterministic approach.

Critical Path Analysis and the PERT Approach

Critical Path Analysis (which is also called the Critical Path Method, CPM) and the additional PERT (Program Evaluation and Review Technique) were developed in the 1950s to control large defence projects in the United States and have been used routinely since then. Within a project it is likely that you will display your final project plan as a Gantt chart (for example using MS Project or other software). A Gantt chart is a bar chart to display project tasks, their duration, time locations and connections with each other.

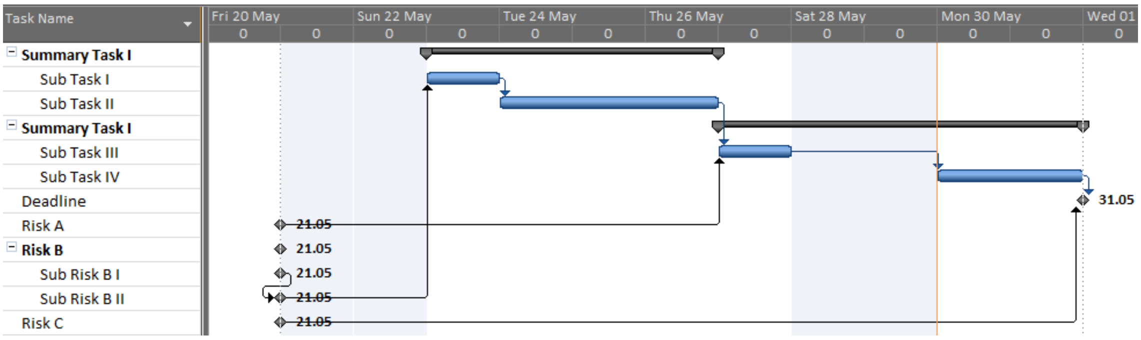


Figure 1: Example for a Gantt chart (MS Project).

The benefit of using CPM within scheduling and on base of a Gantt chart is helping to develop and test your plan ensuring robustness. Critical Path Analysis formally identifies tasks which must be completed on time for the whole project to be completed on time. These tasks form the critical path. CPM gives you therefore the minimum length of time needed to complete a project. It also identifies which tasks can be delayed if resource needs to be reallocated to catch up on overrunning tasks.[1]
PERT is a variation on CPM that takes a more sceptical view of time estimates made for each project stage and brings in uncertainty.
You can perform a basic PERT analysis to estimate task duration. After you specify the optimistic, pessimistic and expected durations of all tasks in your schedule, PERT (often implemented in project planning software) calculates a weighted average of the three durations. You can also use these durations separately to determine a shortest, longest and most likely project end date.[2]

1 Cp. Mind Tools Ltd (1996)
2 Cp. MS Office Support (2011)

1.1.2 How to use CPM?[3]

As with Gantt charts, the essential concept behind the Critical Path Method is that you cannot start some activities until others are finished. These activities need to be completed in a sequence, with each stage being more or less completed before the next stage can begin. These are called sequential tasks.
Other activities are not dependent on completion of any other tasks. You can do these at any time before or after a particular stage is reached. These are called parallel tasks.

Drawing a CPM Chart

The following simple example deals with a computer project. It gives an insight into fundamental drawing of a CPM chart. Normally this is made automatically by software, but to understand CPM basics it is worthwhile to do it "bottom up".
In general use the following steps to draw a CPM Chart:

- Step 1: List all activities in the plan

For each activity, show the earliest start date, estimated length of time it will take and whether it is parallel or sequential. If tasks are sequential, show which stage they depend on.

- Step 2: Plot the activities as a circle and arrow diagram

Critical Path Analysis is presented using circle and arrow diagrams. In these, circles show events within the project, such as the start and finish of tasks. The number shown in the left hand half of the circle allows you to identify each one easily. Circles are sometimes also known as nodes. An arrow running between two event circles shows the activity needed to complete that task. A description of the task is written underneath the arrow. The length of the task is shown above it. By convention, all arrows run left to right. Arrows are sometimes also called arcs. An example of a very simple diagram is shown in Figure 2. The computer project has a start event (circle 1) and a completion of the "High Level Analysis" task (circle 2). The arrow between them shows the activity of carrying out the High Level Analysis. This activity should take 1 week.

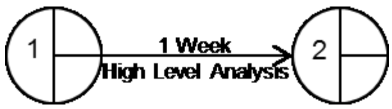


Figure 2: Task connection[4].

Where one activity cannot start until another has been completed, we start the arrow for the dependent activity at the completion event circle of the previous activity. You can see an example below (Figure 3):

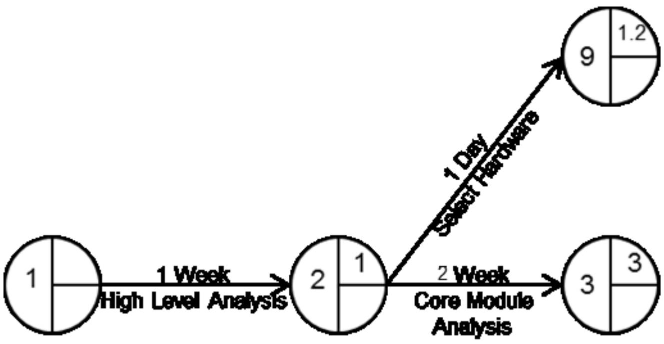


Figure 3: Tasks in sequential order[5].

3 Mind Tools Ltd. (1996)
4 Source: Mind Tools Ltd. (1996)

Here the activities of "Select Hardware" and "Core Module Analysis" cannot be started until "High Level Analysis" has been completed. You can see a second number in the top, right hand quadrant of each circle. This shows the earliest start time for the following activity. It is conventional to start at 0 (units: whole weeks).
A different case is shown below (Figure 4): Here activity 6 to 7 cannot start until the other four activities (11 to 6, 5 to 6, 4 to 6, and 8 to 6) have been completed.

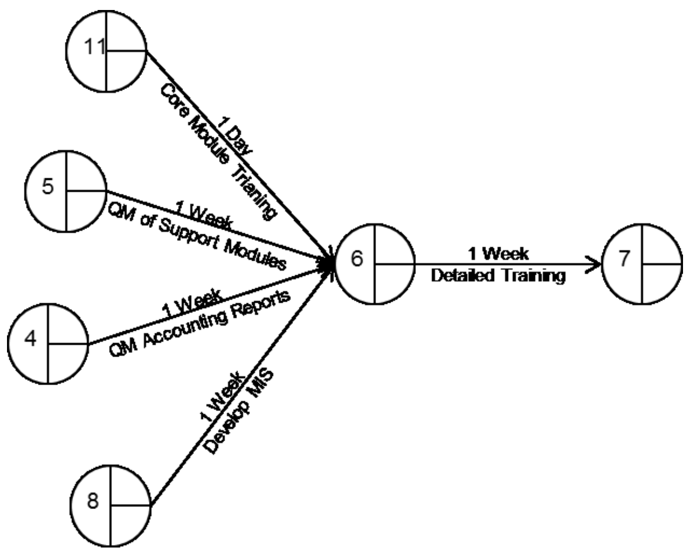


Figure 4: Dependency of task 6 to previous tasks[6].

Figure 5 shows all the activities that will take place as part of the project. Notice that each event circle also has a figure in the bottom, right hand quadrant. This shows the latest finish time that is permissible for the preceding activity if the project has to be completed in minimum time. You can calculate this by starting at the last event and working backwards.
Events 1 to 2, 2 to 3, 3 to 4, 4 to 5, 5 to 6 and 6 to 7 must be started and completed on time if the project has to be completed in 10 weeks. This is the critical path. So the latest finish time of the preceding event and the earliest start time of the following event will be the same for circles on the critical path. You have no slack (→ *slack*) possibilities on this path except you accept a change of the overall project duration. If jobs on the critical path slip, immediate action should be taken to get the project back on schedule.

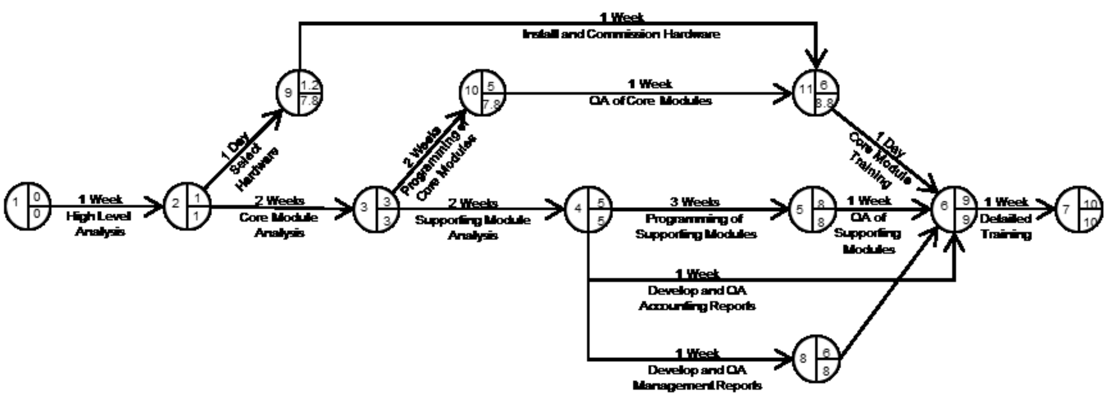


Figure 5: Whole computer project CPM network[7].

5 Source: Mind Tools Ltd. (1996)
6 Source: Mind Tools Ltd. (1996)

- Step 3: Schedule crashing

Frequently you want to complete a project earlier than your critical path says it is possible. In this case you need to re-schedule your project ("crashing").
As an example, it may be necessary to complete the computer project in Figure 5 in 8 weeks rather than 10 weeks. So you could look at using two analysts in activities 2 to 3 and 3 to 4. This would shorten the project by two weeks (but may raise the project cost).
In some situations, shortening the original critical path of a project can lead to a different series of activities becoming the critical path. For example, if activity 4 to 5 were reduced to 1 week, activities 4 to 8 and 8 to 6 would come onto critical path.
As with Gantt charts, in practice software tools like MS Project create CPM charts automatically. Not only this makes drawing them easier, they also make modifications of plans easier and provide facilities for monitoring progress against plans.

1.1.3 How to use PERT?

PERT is an enhancement to CPM due to taking a closer look on task duration estimation. There are generally two ways to define duration values: Firstly, you can apply a single value without any uncertainty (not very realistic) or you can define a range with a min value (shortest time), max value (longest time) and a most likely value (expected time). This range is a far more realistic approach. Hence, PERT is the first step into probabilistic schedule analysis. The underlying distribution (see Chapter 1.3 for further information on this topic) is a PERT (also called Beta) distribution.
The following formula (1) gives you the duration estimation for each task:

$$\text{Task Duration}=\frac{\text{Min}+4*\text{Most Likely}+\text{Max}}{6} \qquad 1$$

In a nutshell PERT is a simple way to keep you from too optimistic scheduling.

1.1.4 Principle of a Probabilistic Schedule Analysis

Uncertainty in Project Schedules

When you develop a project schedule that plans ahead for the future you make certain assumptions. These assumptions generally have to answer three questions:

- How long will it take to complete a certain task?
- How long will it take to complete the whole project?
- What main influences (risks) will have an impact on the first two estimations?

Therefore you have to glimpse into the future and the best you can do is to estimate the expected values. Not knowing with certainty what the actual value will be, but based on historical data, experiences or experts, you can draw an estimate. While this estimate is useful for a first appraisal, it contains some inherent uncertainty and risk. Therefrom it can be advantageous to estimate a range of values that includes all possible states of a value with high certainty.
For instance in a construction project, you might estimate the time it will take to complete a particular job. Based on some expert knowledge, you can also estimate the absolute maximum time it might take in the worst possible case and the absolute minimum time in the best possible case.
As already mentioned the benefit using a range of possible values instead of a single guess (remember PERT) is a more realistic picture of what might happen in the future. When a model is based on a range of estimates, the output of the model will also be a range. This

[7] Source: Mind Tools Ltd. (1996)

range will have the form of a distribution (→ *distribution*) where all output values have an associated possibility to occur (i.e. durations are random variables then). This is different from a normal forecasting model, in which you start with some fixed estimates (for example the time it will take to complete each task of a project) and get another value, the total time for the project.
So when each part has a minimum and maximum estimate, we can use those values to estimate the minimum, the maximum and the most likely time for the project.[8]

1.2 Monte Carlo Simulation

1.2.1 Basics

Monte Carlo Simulation is basically a forecasting method to estimate a process output involving uncertainty. The simulation is based on a mathematical model that describes how a process will likely turn out.
Named for Monte Carlo, the Monaco resort town renowned for its casinos, it was first used by scientists working on the atom bomb in the 1940s in Los Alamos. Monte Carlo is actually a general modelling technique that can be applied to any process where uncertainty is involved (for example life time assessments of products or physical processes). In our specific case it will be applied to project scheduling.

1.2.2 How does it exactly work?

Hence input data are probability distributions (→ *Probability Distribution*) of values, Monte Carlo calculates results over and over, each time using a different set of random values from the various probability distributions.
As already mentioned the simulation does not return a single answer but a range of possible answers and the probability that each answer will occur. All answers and the associated probabilities are combined then to an output distribution that is the final result of the simulation. A random number generator draws samples from various duration input distributions and calculates for each step a duration sum over the whole project. The result is an assembly histogram, which converts in a continuous distribution with more and more iterations (cp. Figure 6).

Figure 6: Fed with input distributions Monte Carlo simulation calculates an output distribution.

A Monte Carlo simulation could involve thousands to tens of thousands of recalculations before it is complete. Plainly spoken every simulation run is a possible project life cycle. Thus if you running through thousands of life cycles you eventually get the most likely project duration, the least likely ones and all between.[9]

[8] Cp. N.N., www.riskAMP.com (2011)

1.2.3 Monte Carlo Simulation and Project Management

A Monte Carlo model is in principle a project plan (Gantt chart) in which some tasks contain probability distributions rather than deterministic values.[10]

So creating your project schedule, you typically put together a series of tasks and estimate duration for each task. When you are finished, you look at the resulting timeline to see the estimated end date. Since uncertainty is associated with each step, additionally a Monte Carlo analysis can be performed in the following way:

First instead of just one duration estimate for an activity, we create three of them. From there we estimate the most likely duration and then we estimate the worst case and the best case. Note: we even can estimate in a more sophisticated manner a whole distribution on the base of the following parameters:

- A central value to anchor the distribution somewhere
- Two boundary values to confine the distribution
- Some values in between to shape the distribution

With each estimate, we assign what we think is a likely probability that it will occur.

Let's look at a small project with three tasks that must be worked on sequentially: [11]

Task A is likely to take two days (70 % probability), but it is possible that it could take one day (20 %) or three days (10 %).

Task B will likely take 5 days (60 %), but could take as few as 4 days (20 %) or as many as 8 days (20 %).

Task C will probably take four days (80 %), three days (5 %), or five days (15 %).

Now the question is: How long will this project take to complete?

The Monte Carlo analysis involves a series of random simulations on our little project. It is possible it would calculate 10 days (2 + 5 + 3) in the first run. The next time, it might calculate 11 days (3 + 5 + 3). Then it could calculate 10 days again (3 + 4 + 3) and so on.

Normally these simulations were run more than 1,000 times. By the time the simulation is completed, you can expect around 700 simulations in which task A took two days (70 %). Likewise, there should be around 150 simulations where task C took five days (15 %). When the Monte Carlo analysis is complete, you do not have a single end date. You have a probability curve showing expected outcomes and the probability of achieving each one. For the purposes of scheduling, we would look at a cumulative curve showing the probability of completing the project between the best case 7 days (1 + 4 + 3) and the worst case 16 days (3 + 8 + 5).

In general, the technique is used to provide safe end date estimates for far larger projects. You would not want to pick the end date that has a 50 % chance of success. The Monte Carlo analysis will tell you the date that you have an 80 % chance to achieve, or a 90 % chance, depending on how safe you need to be.

Note: incorrect input values are an important problem, because the best simulation model is worthless when fed with wrong data. Plainly spoken: garbage in - garbage out! Thus a right estimation setting (experts, historical data etc.) is crucial. This topic will be further treated in Chapter 2.

The following steps sum up the Monte Carlo process in brief:

Input:

- Schedule
- Probability Distributions (as example task duration)
- Simulation setting (like number of iteration)

Process:

10 Cp. Peterson (2005)

11 Cp. Mochal (2002)

- Monte Carlo Simulation

Output:

- P10, P50, and P90 value (→ P values) for the expected project completion time taken from a output distribution
- Confidence intervals (→ Confidence interval)
- Task sensitivities (→ Tornado chart)
- Critical indices (→ Critical index)

1.3 Distributions

1.3.1 Terms and Definitions

Knowing and understanding the basic terms of probabilistic analysis and distributions as backbone is crucial for an effective and efficient data drawing, simulation and result interpretation. Check Chapter 7 (Glossary) for further information on important terms and definitions.

1.3.2 Why Should I Use Distributions?

Task durations usually have a range of possible values and furthermore every value has its associated probability. To cover this fact you have to define an input distribution where values can have different probabilities of different outcomes occurring. As example Figure 7 shows a very simple Uniform distribution. All values have an equal chance of occurring, and one has to define simply the minimum and maximum value to confine the distribution. As example manufacturing costs are normally uniformly distributed.

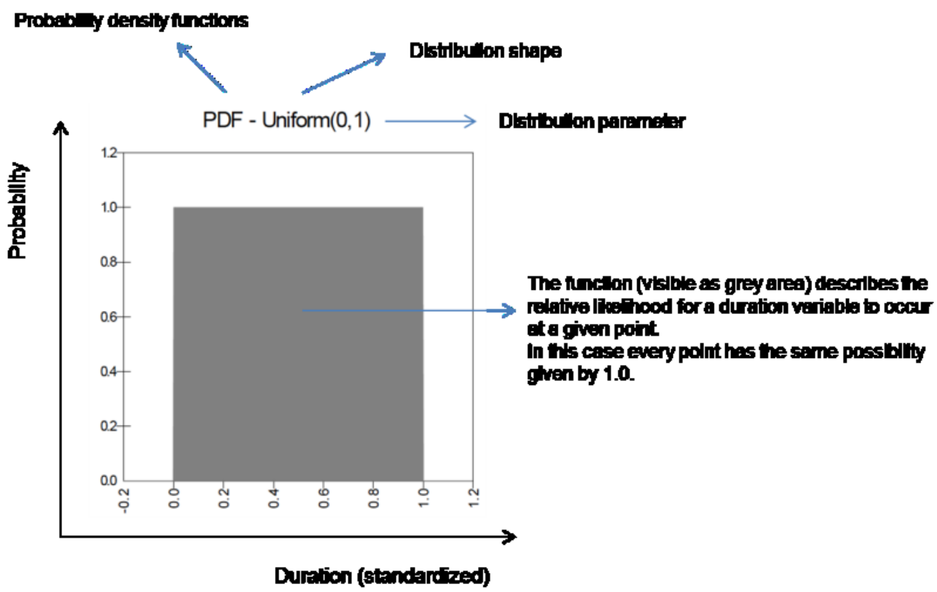


Figure 7: Example for a Uniform distribution.

As already mentioned probability distributions are a much more realistic way of describing uncertainty in your project schedule than single values.

1.3.3 Crucial Distributions in Project Scheduling

The following distributions were taken out of the @Risk database (in fact there are over 60 distributions stored there). They are all useful distributions in project scheduling.

Most common

- Uniform
- Triangular
- PERT (a special form of a Beta distribution)
- Normal/Lognormal

Additionally (in special cases):

- Exponential
- Weibull
- Rayleigh
- General
- Discrete

The selection has been done by literature study and a company survey at OMV. The target hereby was to create a useful catalogue for simply application in daily scheduling.

Distribution catalogue

Table 1 to 10 denote a more detailed view on the selected distributions. A short description, a guideline for application and some defining parameters are given in each case.

Table 1: The Uniform distribution.

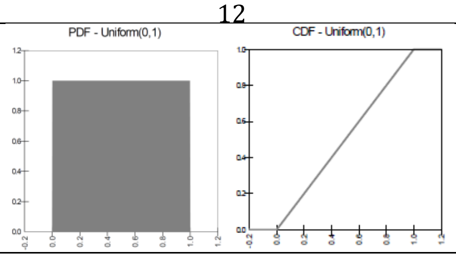
Distribution name	**Uniform** RiskUniform(Minimum,Maximum)
Description	All values have an equal chance of occurring and the user simply defines the minimum and maximum.
Guidelines	Uniform is sometimes referred to as “no knowledge” distribution. You have a base value but no clue, if the probability decreases moving away from that central value. Normally real world situations do not fall into this assumption; in many cases you can estimate a best guess (most likely value) and minimum and maximum values additionally. **Example: the position of a particular air molecule in a room.**
Parameters:	Continuous distribution
Range	$\mathrm{Min} \leq x \leq \mathrm{Max}$, continuous
Mean	$\frac{\mathrm{Max}+\mathrm{Min}}{2}$
Variance	$\frac{(\mathrm{Max}-\mathrm{Min})^2}{12}$
Example[12]	

Table 2: The Triangular distribution.

Distribution name	**Triangular** RiskTriang(Minimum,Most Likely,Maximum)
Description	3 points - minimum, most likely and maximum - define this

12 N.N., Guide to Using @RISK (2010), p. 576

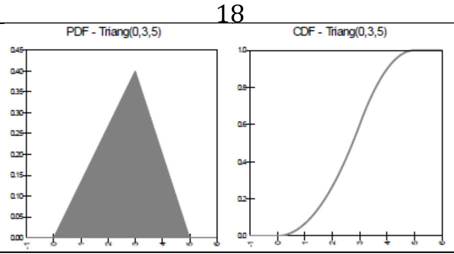
	distribution. It is a typical three point estimation, where the range is known and some central value ("inspired guess"). Skew direction is set by the relation of most likely to min and max.
Guidelines	If you have a best guess and a range (min-max), you can create a distribution that favours the most likely value in some way. The simplest distribution taking this into account is the triangular one. This distribution has a number of desirable properties, including a simple set of parameters and the use of a modal value for instance a most likely case. Furthermore values around the most likely are more likely to occur. There is no requirement that the distribution is symmetrical around the best guess, so you can model a variety of different circumstances. There are two main disadvantages of a Triangular distribution. First, when the parameters result in a skewed distribution, then there may be an over-emphasis of the outcomes in the direction of the skew. Second, the distribution is bounded on both sides, whereas many real-life processes are bounded on one side but unbounded on the other. **Example: inventory levels.**
Parameters:	Continuous distribution
Range	Min≤x≤Max, continuous
Mean	$\frac{\text{Max}+\text{Most Likely}+\text{Min}}{3}$
Mode	Most Likely (ML)
Variance	$\frac{\text{Min}^2 + \text{ML}^2 + \text{Max}^2 - \text{Max*ML} - \text{ML*Min} - \text{Max*Min}}{18}$
Example[13]	

Table 3: The PERT distribution.

Distribution name	**PERT (Beta)** RiskPert(Minimum, Most Likely, Maximum)
Description	It is a special Beta distribution with a min-max confinement and the shape parameters α_1 and α_2 that are calculated from the most likely value.
Guidelines	PERT is rather like the Triangular distribution (3 point estimation), but a more realistic approach and it could be seen a min-max confined Normal distribution. So assuming that many real world problems are normally distributed, you can take PERT as approximation without knowing the precise parameters of a Normal distribution. Like the triangular distribution, the PERT distribution emphasizes the most likely value over the minimum and maximum estimates. However values between the most likely and

[13] N.N., Guide to Using @RISK (2010), p. 573

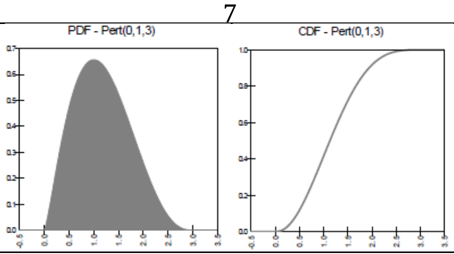
	extremes are more likely to occur than in triangular distributions and the extremes are not as emphasized. In practice, this means that we "trust" the estimate for the most likely value, and we believe that even if it is not exactly accurate, we have an expectation that the resulting value will be close to that estimate. Furthermore PERT is superior to Triangular, if skewness takes place as the smooth shape of the curve places less emphasis in the skew direction. PERT is heavily used in three point estimation techniques. Cons of PERT are bad capturing of extreme events and tails. **Example: Project costs.**
Parameters:	Continuous distribution
Range	$Min \leq x \leq Max$
Shape	$\alpha1 = 6\left[\frac{\mu - Min}{Max - Min}\right] \quad \alpha2 = 6\left[\frac{Max - \mu}{Max - Min}\right]$
Mean	$\mu = \frac{Min + 4 * ML + Max}{6}$
Mode	Most Likely (ML)
Variance	$\frac{(\mu - Min) * (Max - \mu)}{7}$
Example[14]	

Table 4: The Normal distribution.

Distribution name	**Normal** RiskNormal(Mean,Standard Deviation)
Description	2 parameters, mean and standard deviation, specify this well-known distribution. It is unbounded on both sides. Many data could be described by this "bell shaped" curve.
Guidelines	Generally the output of many models is approximately normally distributed, because they add a lot of uncertain sub-processes (→ Central Limit Theorem). The distribution can be used to represent the uncertainty of a model's input whenever it is believed that the input itself is the result of many other similar random processes acting together in an additive manner (but where it may be unnecessary, inefficient or impractical to model these detailed driving factors individually). Values in the middle near the mean are most likely to occur. **Example: total goal number in a soccer season.**
Parameters:	Continuous distribution
Range	$-\infty < x < +\infty$, continuous
Mean	μ, continuous location parameter
Mode	μ
Variance	σ^2, continuous spread parameter

[14] N.N., Guide to Using @RISK (2010), p. 560-561

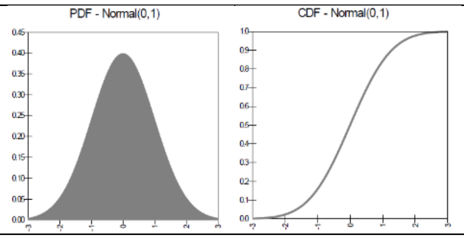
Example[15]	

Table 5: The Lognormal distribution.

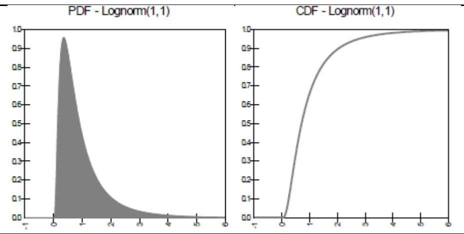
Distribution name	**Lognormal** RiskLognorm(Mean,Standard Deviation)
Description	2 parameters, mean and standard deviation, specify this distribution. Just as the Normal distribution results from adding many random processes, the Lognormal arises by multiplying many random processes (the logarithm of the product of random numbers is equal to the sum of the logarithms).
Guidelines	As the Normal distribution, the Lognormal has two parameters (μ, σ) corresponding to the mean and standard deviation. But in addition values are positively skewed, not symmetric. It is used to represent values which do not go below zero but have unlimited positive potential like stock prices. The distribution has a number of desirable properties of real world processes. These include that it is skewed and that it has a positive and one-side unbounded range i.e. it ranges from 0 to infinity. With this "tail" you can take seldom risks into account. **Example: oil reserves.**
Parameters:	Continuous distribution
Range	$0 \leq x < +\infty$
Mean	μ
Mode	$\frac{\mu^4}{(\mu^2+\sigma^2)^{\frac{3}{2}}}$
Variance	σ^2
Example[16]	

Table 6: The Exponential distribution.

Distribution name	**Exponential** RiskExpon(Beta)
Description	An exponential distribution is defined by the entered Beta value. The mean of the distribution equals Beta.
Guidelines	The exponential is often used to model the time between independent events that happen at a constant average rate. For example waiting for a train after the train before has

15 N.N., Guide to Using @RISK (2010), p. 545-546
16 N.N., Guide to Using @RISK (2010), p. 537

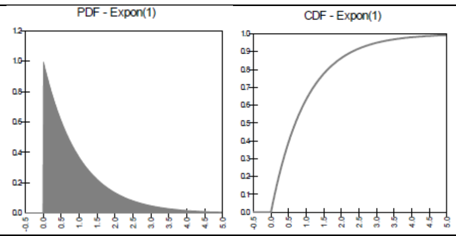
	passed. That can be calculated with the knowledge of a single variable called the expectation value (Beta). Because trains use to pass very regularly (with an expected value), say every 20 Minutes, it will be very unlikely to wait 3 hours for a train (except in Austria). The main disadvantage of an Exponential distribution is the assumption of constant event rates. If you don`t want this, take a Weibull distribution. **Example: incoming phone calls in a call center.**
Parameters:	Continuous distribution
Range	$0 \le x < +\infty$
Mean	β
Mode	0
Variance	β^2
Example[17]	

Table 7: The Weibull distribution.

Distribution name	**Weibull** RiskWeibull(Alpha,Beta)
Description	It is a more flexible version of an Exponential distribution. Weibull can take on the characteristics of other types of distributions, based on the value of the shape parameter β. α measures the time-dependent location and frequency of some event(s). So shape and scale are dependent to the parameters. α >1 Event rate increases over time. α =1 Constant event rate (random events, Exponential distribution). α <1 Event rate decreases over time.
Guidelines	This distribution is often used as a “distribution of time to first occurrence”, where it is desired to have a non-constant intensity of occurrence. This distribution is flexible enough to allow an implicit assumption of constant, increasing or decreasing intensity, according to the choice of its parameter α (α<1, =1, or >1). For example calculating life time of a mechanical forced product, one may choose to use α<1 to represent that the older something is, the more likely it fails. **Example: Material breakdown.**
Parameters:	Continuous distribution
Range	$Min \le x < +\infty$
Shape	$\alpha > 0$
Scale	$\beta > 0$
Mean	$\beta\Gamma\left(1 + \frac{1}{\alpha}\right)$ where Γ is the Gamma Function

[17] N.N., Guide to Using @RISK (2010), p. 499

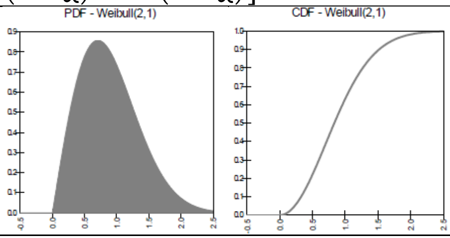
Mode	$\beta\left(1-\frac{1}{\propto}\right)^{\frac{1}{\propto}}$ for $\propto>1$ 0 for $\propto\leq 1$
Variance	$\beta^2\left[\Gamma\left(1+\frac{2}{\propto}\right)-\Gamma^2\left(1+\frac{1}{\propto}\right)\right]$ Γ ... Gamma Function
Example[18]	

Table 8: The Rayleigh distribution.

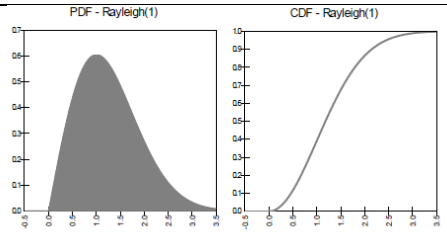
Distribution name	**Rayleigh** RiskRayleigh(Beta)
Description	The Rayleigh distribution is a Weibull distribution with a shape factor of 2.
Guideline	It could be used as an alternative to Normal but with a min-boundary.
Example[19]	

Table 9: A General distribution.

Distribution name	**General** RiskGeneral(Min, Max, $\{X_1, X_2,\ldots, X_n\}$,$\{p_1, p_2,\ldots, p_n\}$)
Description	This generalized probability distribution is based on a density curve created using the specified (X,p) pairs. Each pair has a value X and a probability weight p that specifies the relative height of the probability curve at that X value.
Guidelines	This distribution is a try to fit a general distribution to some approximated values: For instance:[20] RiskGeneral(0,10,{2,5,7,9},{1,2,3,1}) specifies a general probability density function with four points. The distribution ranges from 0 to 10 with four points 2,5,7,9 specified on the curve. The height of the curve at 2 is 1, at 5 is 2, at 7 is 3 and at 9 is 1. The curve intersects the X-axis at 0 and 10. **Example: no specific example.**
Parameters:	Continuous distribution
Range	No closed form
Mean	No closed form
Mode	No closed form
Variance	No closed form

18 N.N., Guide to Using @RISK (2010), p. 580
19 N.N., Guide to Using @RISK (2010), p. 565
20 N.N., Guide to Using @RISK (2010), p. 507

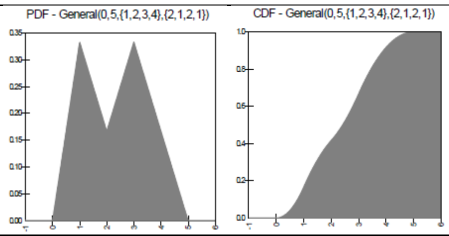
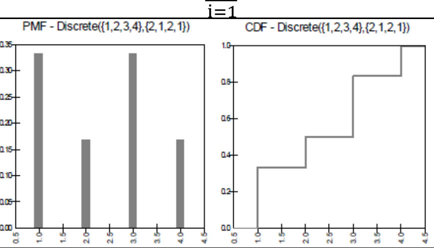
Example[21]	

Table 10: A Discrete distribution.

Distribution name	**Discrete** RiskDiscrete({X1,X2,...,Xn},{p1,p2,...,pn})
Description	Each outcome has a value X and exactly one weight p (probability weight) which specifies the outcome's probability of occurrence.
Guidelines	The user defines specific values that may occur and the likelihood of each. For instance think of the results of a lawsuit: 20 % chance of positive verdict, 30 % change of negative verdict, 40 % chance of settlement and 10 % chance of mistrial. **Example: dicing.**
Parameters:	Discrete distribution
Domain	$X \epsilon \{X_n\}$
Mean	$\mu \equiv \sum_{i=1}^{N} x_i p_i$
Mode	The X value corresponding to the highest p value.
Variance	$\sigma^2 \equiv \sum_{i=1}^{N} x_i p_i$
Example[22]	

1.3.4 Selecting a Distribution

Generally it is wise to ensure that each input distribution has a range/parameters corresponding to realistic input data. For example there could be technical limits like a min value of 0 for well production rates and a max value that is certainly not infinity.
Furthermore one should favour simple distributions with logical story behind each parameter. If you have a small data set and no most likely points, then take a uniform distribution. Where some evidence of modes/means/most likelies and fixed min-max values is arising, take a simple triangular distribution as first approach. Do you expect some extreme events, take a long tailored distribution to cover that fact and so on.
The general conclusion may be that the most common mistake in choosing distributions is time waste debating the choice of shape (Central Limit Theorem, Chapter 1.4.1) rather than trying to get realistic distribution parameters.

21 N.N., Guide to Using @RISK (2010), p. 508
22 N.N., Guide to Using @RISK (2010), p. 490

1.4 PSA Key Factors Given by Literature

1.4.1 Central Limit Theorem (CLT)

Two important statistical facts can take place, if you have a large collection of input distributions with different parameters sampled. These two outcomes summarize under the name Central Limit Theorem (also called CLT). The sum of independent input distributions of any shape will (1) tend to a normal output distribution with (2) diminishing standard deviation.
David Vose states in his book *Risk Analysis*[23]*:* "The Central Limit theorem [sic!] is probably the most important theorem for risk analysis modelling."[24]

Normal shaping of input distributions

When you have an overall summation of a sufficiently large number of uncorrelated input distributions by Monte Carlo simulation, you will get an output distribution that is approximately Gaussian. Amazingly it is irrelevant what types of input distributions are summed (cp. Figure 8).

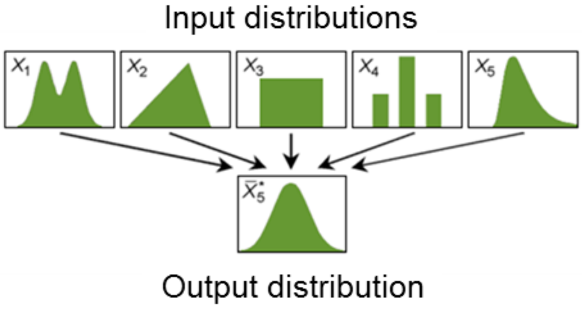


Figure 8: The output distribution of five independent random variables is still approximately normal despite the fact that the five random variables have very different distributions. This is one main aspect of CLT.

For example various natural phenomena, such as the heights of individuals, are approximately normal distributed. These phenomena are sums of a large number of independent random variables and hence approximately normally distributed by the Central Limit Theorem.
Additionally some side-effects take place:

- Output mean will be close to the sum of means of input distributions.

However, because Gaussian distribution is symmetric, mean=median=mode holds. Therefor the sum of all task duration medians/modes is normally smaller than the total duration median/mode (especially due to positively skewed input distributions).[25]

- Output variance will be close to the sum of variances of input distributions.

Both effects are also true for a multiplication of input variables except for the output distribution shape which is lognormal.

Narrowing output distribution

Furthermore increasing the sample size the output distribution becomes narrower as given by Equation 2:

$$\sigma_o = \frac{\sigma_p}{\sqrt{n}} \qquad 2$$

σ_o = standard deviation of output distribution

[23] Vose (2003)
[24] Vose (2003)
[25] Book (2002)

σ_p = standard deviation of population (=sum of all input standard deviations)
n = sample size
In plain words: with more and more tasks the spread of the result will reduce until the results are clearly in conflict with our experience.[26]
We can counter this effect by[27]

- reducing the number of distributions,
- avoid to narrow input ranges,
- introducing correlations between tasks[28]
- and implementing seldom risk events and their impacts.

Summarizing (1) and (2) you can say that output distributions of a large number of input variables depend more on the input means and variances and less on their individual shapes. Furthermore schedules should focus on limited set of essential task durations.[29,30]

1.4.2 Constraints

Constraints impose restrictions on the start and finish dates of tasks. There are different types of constraints in MS Project[31]:

Flexible constraints

Such constraints like "As Soon As Possible (ASAP)" or "As Late As Possible (ALAP)" do not have specific dates associated with them.

Semiflexible constraints

A semiflexible constraint requires an associated date that controls the earliest or latest start or finish date for a task. It allows a task to finish at any time, as long as it meets the start or finish deadline.

Inflexible constraints

Such constraints like "Must Start On (MSO)" and "Must Finish On (MFO)" have an associated date, which controls the start or finish date of the task.
For optimal scheduling flexibility, the MS Office team recommends that you allow MS Project to use flexible constraints to calculate the start and finish dates for tasks based on the durations and task dependencies you enter. Despite schedulers often use inflexible constraints it could be useful to eliminate them, so the project can overrun in the simulation and not in reality. Inflexible constraints will hide risks and give an unrealistic simulation result such as too early completion dates.

1.4.3 Correlations

For example Holtz[32] and Murtha[33] have emphasized the importance of taking into account correlations between uncertainty variables when conducting Monte Carlo simulation.
Correlation tells us thereby how two task durations are related to each other. It makes the durations move together (in same directions if positive correlated, in other directions if negative correlated). Both tasks will take more or less time together. Therefore correlations increase the risk of unexpected and/or extreme completion dates. Correlation strength is

26 Cp. Murtha (2002)
27 Cp. Peterson (2005)
28 Cp. McIntosh (2004)
29 Cp. Saleri (1996)
30 Cp. Murtha (1987)
31 Cp. N.N., MS Office Support: About Constraints (2011)
32 Cp. Holtz (1993)
33 Cp. Murtha (1993)

measured by correlation coefficients. Correlation coefficients if existing have to be well estimated and implied in the schedule.

1.4.4 Monte Carlo Analysis as Analysing Tool

Monte Carlo Simulation is nowadays a widely used analysing tool. Since Hertz`s widely read article[34], there have been interesting articles[35,36] and books[37,38,39] dealing with the topic. Monte Carlo is arguably better than other methods, especially PERT, due to the following points:

- PERT only uses Beta distributions, but Monte Carlo can use all.
- PERT does not recognize a changing critical path, whereas Monte Carlo calculates the actual critical path (→ *critical path*) within every simulation run.
- PERT becomes very time-consuming with big schedules. Monte Carlo gives you many "case studies" within a very short time.

Monte Carlo offers additional information like critical indices (→ *critical indices*) or sensitivity charts based on a correlation analysis.
But MC Simulation has some limitations too:
Schedule must be complete and correct. All tasks are correctly tied in and lags (→ *lag time*) are appropriate. Be careful of negative lags (→ *lead time*), because they are not easy to handle.[40] Secondly estimation of input data must be appropriate (very often experts are not very experienced in estimating min and max values). Collection of real data for comparison will be beneficial.[41] Thirdly, it is difficult to represent correlation between tasks, so often approximations are developed to simplify simulation. The effects of approximations are not detected precisely.[42]

1.4.5 Precise and Accurate Input Values

A correct and well defined stochastic modelling is worthless, when it is fed with incorrect data or in other words "garbage in - garbage out". There has to be a strong emphasize on gathering correct input data (see also Chapter 2 Input Data Estimation). Also a regular update of input data and associated distributions using performance measurement data is very useful. Secondly correlations if of existence must be defined and selected for further simulation.

- **Input values** are the **fuel** of good scheduling.

1.4.6 Well Defined Schedule

If input data is the fuel of every simulation, the basic schedule is the backbone. A non-logical, wrongly linked and inflexible (constraints) Gantt chart kills every simulation. Another problem associated with Monte Carlo simulations is that, if a project slips, project managers usually perform certain actions. So the base schedule is changing without simulation control. As already mentioned in Chapter 1.4.5, a regular schedule and simulation revision is needed.

- A **well-defined schedule** is the **backbone** of good scheduling.

34 Hertz (1964)
35 Megill (1985)
36 Cronquist (1991)
37 McCray (1975)
38 Megill (1977)
39 Murtha (1993)
40 Cp. McCabe (2003), p. 1564
41 Cp. McCabe (2003), p. 1564
42 Cp. McCabe (2003), p. 1564

1.4.7 The Difficulty of Assessing Uncertainty - Objective and Subjective Estimation

Uncertainty estimation can be objective, subjective or both.[43] Objective data are historical data sets etc., subjective data is normally generated by experts (cp. 1.4.8 Expert Judgements). However, Capen[44] and subsequently Rose[45] has confirmed with their work, that experts are not precise estimators.
Maybe the combination of both approaches is the best way for data assessment, so the following discussion is divided in two parts.

Objective Data Estimation

As already mentioned the basis of objective uncertainty assessment is historical data. Databases could be generated by evaluating older projects, comparing similar projects or buying it from third-party suppliers (for example International Project Management Association; IPMA).
Once you have some data, you can analyse it by finding important parameters like mean, standard deviation etc. You can also do a distribution fitting with the use of commercial software like Best Fit (Palisade Corporation).

Subjective Data Estimation

As commonly known interviewing different experts can produce very different data sets. Inherent biases, experiences, knowledge and so on cause these differences. Chapter 1.4.8 and in deep Chapter 2 deal with that phenomenon.

In summary it is the combination of both, objective and subjective data, that will improve data input for schedule estimation.[46]

1.4.8 Expert Judgements

Probabilistic Schedule Analysis can be realized either with large historical data sets or expert judgement of actual data. Latter approach is a valid way to generate estimations, but there some potential pitfalls and limitations worth thinking of.[47]
For example experts could be influenced by negative or positive experiences or far too certain regarding their own ability. A very often-referenced paper by Capen[48] provides a good insight in this topic. Key results of this work are:[49]

- Experts are often using too narrow min-max ranges. Basic rule of thumb says that when you feel right about your estimation you are probably too narrow.
- Use as many experts as possible. This will give you feedback, peer reviews and a broader range of estimation (also cp. Chapter 2.4 Estimation and the Wisdom of Crowds).

1.5 Literature Outcome

Probabilistic schedule analysis is superior to a deterministic approach by taking uncertainty and therefor reality into account.
In many cases Monte Carlo simulation was and is used. On the other hand PERT analysis can give you a raw and quick approach to possible values.

43 Cp. Hawkins (2002)
44 Cp. Capen (1995)
45 Cp. Rose (1987)
46 Cp. Hawkins (2002)
47 Cp. Akins (2005)
48 Capen (1976)
49 Cp. Akins (2005), p. 4

Probability distributions represent the relative likelihood of each of the possible task durations. Single values (deterministic approach) instead are not realistic and will lead to non-feasible completion times. The typical distribution for task durations is continuous. This means an activity can take all durations within the range. Additionally most distributions have a single most likely value and diminishing probability toward the optimistic and pessimistic duration limits. Finally distributions do not have to be symmetrical, because possible durations tend to be asymmetrical distributed between the max and the min value.[50]
All in all literature say the effort in examining and selecting the right distributions is greater than the payback. Time will be better spent in finding the right input data than finding a "magical distribution". If this statement is true for a small number of tasks (minor CLT impact) or correlated input values was not mentioned. This will be a point of departure for applied probabilistic schedule analysis (cp. Chapter 4). Moreover it is not prudent to say that one shape, say the normal, fits all cases. Rather, the ten most common distributions given in Chapter 1.3.3 seem to be sufficiently diverse to represent uncertainty. Choosing between them should be done by fitting the special characteristics of a task with the distribution catalogue (1.3.3).[51]

1.6 Literature List

This is a selection of best literature regarding PSA. In fact all literature in Chapter 6 (References) is worthy to read.

An Innovative Tool on a Probabilistic Approach Related to the Well Construction Costs and Time Estimation (A. Merlo et al., Eni, 2009)

- Describes a software tool and its basis for cost and time estimation in well construction requested by Eni E&P.

Monte Carlo Techniques Applied to Well Forecasting: Some Pitfalls (H.S. Williamson et al., SPE, BP, 2006)

- Contrasts deterministic and stochastic approach and highlights the potential benefits of the latter. Main part is the application of Monte Carlo simulation to time and cost estimation of single wells.

Judgment in Probabilistic Analysis (D.C. Purvis, The Strickland Group Inc., 2003)

- Provides a general instruction to probabilistic analysis and deals with some examples of pitfalls and common mistakes.

Decision Making in the Oil and Gas Industry: From Blissful Ignorance to Uncertainty-Induced Confusion (J. Eric Bickel et al., Texas A&M University, SPE, 2007)

- Tries to combine decision making tools and uncertainty analysis and recommends that uncertainty quantification should be made decision-focused and with an iterative modeling.

A Probabilistic Approach for Drilling Cost Engineering and Management Case Study: Hassi-Messaoud Oil Field (M. Saibi, Sonatrach, 2007)

- Introduces a method for cost estimation, focal points are data distribution sampling, Monte Carlo simulation, correlations etc.

Judgement in Probabilistic Analysis (D.C. Purvis, The Strickland Group Inc., 2003)

- Gives a general introduction to probabilistic analysis and some examples of pitfalls and common mistakes.

[50] Cp. Hulett (2009), p. 11-16

Risk and Uncertainty Management: Best Practices and Misapplications for Cost and Schedule Estimates (S.K: Peterson et al., J Murtha Assocs., SPE, 2005)

- Presents best practices for applying risk management, risk analysis and uncertainty analysis to capital expenditure cost and schedule estimates. Some current misapplications as potential barriers are highlighted.

A New Tool To Evaluate the Feasibility of Petroleum Exploration Projects Using a Combination of Deterministic and Probabilistic Methods (A. Al-Thawadi, Schlumberger, SPE, 2007)

- Gives a comparison between (deterministic) decision trees and (probabilistic) Monte Carlo simulation.

2 Input Data Estimation

2.1 Overview

Without precise input values every simulation must fail. So it is a crucial part in the beginning of a probabilistic schedule analysis (or better in times of starting scheduling) to find a proper way to estimate the right values with the right people and the right surroundings.
To do so the main characteristics of estimation must be examined and correctly implemented in an estimation workshop. The following chapters should give the basis for an estimation workshop. Chapter 2.2 examines psychological biases, Chapter 2,3, 2.4 and 2.5 are based on recently published books dealing with our capability of (right) estimation and our awareness of risks.

2.2 Estimation and Biases

Estimations made by human beings cannot be objective and have an unavoidable subject specific component. Humans are basically not accurate estimators.[52]
They are exposed to many different biases and other cognitive effects. Knowing cause and effect of various so-called biases is essential for realistic estimation.
Based on previous work of Stephan Staber (OMV) and the ground-breaking work of Kahneman[53] and Tversky a list of cognitive effects was compiled which can help to be aware of potential disadvantageous influences on estimating task duration values.

2.2.1 Cognitive Effects on Duration Estimation[54]

There are 2 types of biases, motivational and cognitive ones. Motivational biases are basically the discrepancy related to personal situation and/or reward. Cognitive biases are a subconscious difference in the way mind processes information.[55]
Many papers[56,57,58] say that the two most frequent biases are cognitive:

Anchoring and Adjustment

Estimations of durations are linked to a starting point (anchoring) like a best guess and then are stepwise adjusted (adjustment). Therefore the estimation of uncertain durations is dependent on the anchor as a subjective reference point.[59]

Availability Bias

Human beings estimate the probability of possible durations, the easier and faster they are able to imagine examples for the underlying tasks or they recall past examples. Therefore emotionality, experience and recent occurrence lead to an overestimation of the probability and in consequence to biased estimates.

Some other less frequent but also important biases are:

52 Cp. Hawkins (2002)
53 Cp. Kahneman (2011)
54 Cp. Staber (2008)
55 Cp. Hawkins (2002)
56 Cp. Tversky (1974)
57 Cp. Kahneman (1979)
58 Cp. Hogarth (1987)
59 Cp. Purvis (2003)

Inexpert Expert

The person nominated (wrongly) as being able to provide the best estimate occasionally has very little idea.

Culture of Organization

The environment people work in may sometimes impact on their estimating. Project manager, for example, will often provide too optimistic estimates of future project completion dates because of the eager working culture.

Conflicting Agendas

Sometimes the expert will have strong interest in the values that are submitted to a schedule.

Unwillingness to Consider Extremes

Estimators will frequently find it difficult or be unwilling to cognize circumstances that would cause a variable to be extremely low or high. The data analyst will often have to encourage the development of such extreme scenarios in order to force an opinion that realistically covers the entire possible range. This can be done by the analyst making up some examples of extreme circumstances and discussing them with the estimator.

Eagerness to Say the Right Thing

Occasionally, the estimator will be trying to provide the answer she/he thinks the workshop moderator wants to hear. For this reason, it is important not to ask questions that are leading or to offer a value for the expert to comment on.

People Too Busy

A time-intensive estimation workshop may not be very welcome. Obvious symptoms in returned estimations are when the expert offers over-simplistic estimate or minimum, most likely and maximum values that are equally spaced for all estimated variables.

Belief that the Estimator Is Quite Certain

It may be perceived by the expert that assigning a large uncertainty to a parameter would indicate a lack of knowledge and thereby undermine his/her reputation. The expert may need to be reassured that this is not the case. An expert should have a more precise understanding of a parameter's true uncertainty and may, in fact, appreciate that the uncertainty could be greater than expected.

Halo Effect

When estimating a subject or object one single aspect - which is very well developed - is outshining all other aspects. In other words the decision maker is dazzled by positive or negative aspects. For example persons who have done well in the past, benefit from the halo effect when being evaluated in the present.

Illusion of Control

We tend to believe we have more control over certain tasks than we actually have. This leads to an under- or overestimation of probabilities.

Representativeness

One type of bias is the erroneous belief that the large scale nature of uncertainty is reflected in small scale sampling. This means one usually believes that the output sample drawn from a population has the same characteristics as the population itself. For example in National Lotteries many would say they had no chance of winning if they selected the consecutive numbers 16, 17, 18, 19, 20 and 21. The lottery numbers are randomly picked each week so it is believed that the winning numbers should also exhibit a random pattern, e.g. 3, 11, 15, 21, 29 and 41. Of course, both sets of numbers are equally likely.

Need for Confirmation

We tend to gather facts that support certain conclusions but disregard other facts that support different conclusions.

Wishful Thinking

We tend to see things in a positive light and this can distort realistic estimation.

Group Thinking

There can be peer pressure to conform to the opinions held by the group. So task duration estimation in a group is frequently biased.

Source Credibility Bias

We reject something if we have a bias against the person, organization or group to which the person belongs. On the other hand we tend to accept a statement by someone we like.

2.3 Estimation and Black Swans[60]

A Black Swan is a metaphor of something so rare that it seems not to exist. For a long time it was widely believed in Europe that all swans were white. Then a Dutch explorer in Australia sighted a black swan. One contrary observation overturned years of shared understanding. This is called the problem of induction. In his works beginning in the 1930`s the famous philosopher Karl Popper delivered a solution for this "black swan problem". He proposed empirical falsification (basically try to detect a black swan) instead of verification to find out, if the statement: I have only seen white swans so all swans are white is true or not.

Circa 70 years later this is again the leading topic of a book called *The Black Swan* by Nassim Nicholas Taleb published in 2007. Taleb defines Black Swans as a metaphor for surprising events (to the observer) with major impact. His thesis is that not only Black Swans are new information or events entirely outside our expectations or predictions and more common than we think they are, they are far more influential as well. Additionally he argues that we naturally expect tomorrow to be like yesterday, build narrative explanations that establish a sense of continuity and predictability and have great difficulty wrapping our minds around the possibility of the unexpected. But such aspects of modern life as highly interconnected financial structures, hugely increased access to information and increases in world population are making Black Swans more and more likely.

In fact we drastically overestimate risks with spectacular narratives attached, drastically underestimate other risks (particularly ones embedded in day-to-day life) and do a bad job at analysing risk trade-offs around improbable events. Taleb argues that we naturally and through education expect most aspects of the world to follow something like a bell curve, where the outliers drop off so quickly that they can be safely ignored. One example he uses is human height: if you fill a stadium full of people and then want to measure their average height, missing a person or two would not matter. No single person can be so tall or so short as to throw off your measurement very far. Height is a property that we are good at

60 Cp. Taleb (2007)

reasoning about, as are many natural facts that we have lived with since we evolved from apes. Wealth, on the other hand, is one of Taleb's examples of a more modern, information-driven, and Black Swan-susceptible measurement. If you are attempting to measure the average wealth of that same stadium full of people, it suddenly matters whether you account for everyone or not. If Bill Gates is in the stadium, he could easily have more money than everyone else in the stadium put together, so he has a significant effect on the outcome. Most of *The Black Swan* is devoted to an inventory of all the ways in which our reasoning breaks on such quantities and how much we are in denial about that. So basically it is a book about the limits of knowledge.[61]

For project scheduling the benefit of including Black Swans is to widen the estimation perspective and more specifically to use long tailored distributions to cover these extreme events. Consecutively it is necessary to imply the Black Swan concept in systematic estimation workshops.

2.4 Estimation and the Wisdom of Crowds[62]

We tend to think, that if we make the individuals smarter, we make the group smarter and a team more effective. But it can be shown that if we make the individuals more diverse, we get even better teams, smarter groups and wiser crowds.

Of course, there are some conditions, which must be fulfilled to get wise crowds working. For instance, if someone requires open-heart surgery, he certainly does not want a collection of butchers, bakers and carpenters to open the chest cavity. He would prefer a trained heart surgeon and for good reason. But in other circumstances, such as designing a physics experiment, cracking a secret code or evaluating post-heart attack treatment, diversity can be very helpful.

Understanding when and why wise crowds prove beneficial and how it could be used in an estimation workshop is the purpose of this chapter.

In 2004, a book called *The wisdom of crowds: why the many are smarter than the few and how collective wisdom shapes business, economies, societies and nations* by James Surowiecki was published and instantly had a big success. Its central thesis was that a diverse collection of independently deciding individuals is likely to make better decisions and predictions than any single member of that group or even experts. For Surowiecki consecutively three types of problems could be easier solved by involving crowds. The first are **cognition problems**, those have definitive solutions. For example how many products a firm will sell during the next year. The second is called a **coordination problem**, such as driving safely in heavy traffic or buyers and sellers finding each other and trading at fair prices. The final one is called a **cooperation problem**, which focuses as example on getting mainly self-interested people to work together paying less taxes or dealing with pollution.

On the other hand there are many cases, where crowds are acting very irrational like riots or stock-market people creating bubbles. So there must be some conditions for a wise crowd.

2.4.1 Requirements for a Wise Crowd[63]

There are some key factors according to Surowiecki separating wise crowds from unreasonable ones.

- Diversity of opinion and knowledge

Each member of the group should have private information. This could also be a very individualistic interpretation of known facts.

61 Cp. Allbery (2008)
62 Cp. Surowiecki (2005)
63 Cp. N.N., Wikipedia (2009-2011)

- Independence

One`s opinion should not be influenced by surrounding opinions.

- Decentralization

This means, an individual is able to draw on local knowledge and surroundings. New perspectives and a broader knowledge basement are reached consecutively.

- Aggregation

Some mechanism is needed in the end to turn individual judgements into collective decisions. This is very important, because decentralisation without aggregation is a recipe for disaster.

- Incentives

The final condition is to have incentives. The basic idea is a reward, if you are right, and a penalty, if you are wrong. The payoffs can be monetary (e.g. stock markets), but do not have to be. For example a reputation payoff or something else can be installed. Incentives have several impacts:[64]

- Improved prediction accuracy.
- Encouragement to greater diversity (payoff is higher if your bet is away from the crowd).

2.4.2 Types of Crowd Wisdom

Three types of crowd wisdom can occur regarding Surowiecki. These types are correlating to the already mentioned problems for wise crowds in Chapter 2.4:

- Cognition (=insight into something)

As a good example market judgement, which can be seen as based on crowd wisdom, is regularly much faster, more reliable and less influenced by political forces than decisions of experts or even expert committees.

- Coordination (=a reasonable interaction or teamwork between individuals)

Not colliding in moving traffic flows or pavement flows is a fine example for the ability of crowds to coordinate their behaviour. Surowiecki examines in his book how common understanding within a society or culture, ergo a group, allows accurate judgements about specific reactions of other group members.

- Cooperation (=trade-off between individual targets)

A free market is basically a network of trust without a central station controlling the behaviour of people in that network. Also their compliance is not directly forced. Therefore it is a self-organised example of decentralised teamwork.

2.4.3 What Can Go Wrong with Crowd Wisdom?

In Surowiecki`s opinion there are several situations, in which the crowd produces very bad judgements (like rational bubbles) and argues that in these types of situation their cognition or cooperation fails because the group members are too conscious of the opinions of others and began to conform rather than think in a different way. Now the benefit of individual judgement and private information is lost and the crowd can only do as well as its smartest member rather than perform better.
There are several ways this could happen:

- Homogenity

Same background, same culture, a certain bias within a crowd cannot ensure enough diversity and therefore less variance in approach.

- Centralization

64 Cp. Mauboussin (2007)

Everyone is part of the same rigid hierarchical structure or has the same authority, so important parts of crowd are isolated.

- Emotions

Herd phenomena, group pressure, in extreme cases collective hysteria, all based on a strong feeling of belonging can destroy a sensible crowd.

- Imitation

Where choices are visible and in sequence, an information cascade can occur: once first decisioners have made their choice, this choice is sufficiently informative, and it pays for later decision makers to simply copy those. You lose diversity.

- Division

Free flow of information is needed, so everyone can choose what information is important to him. Divisions, let`s say expert sub divisions as an example, are knowledge walls and lead to isolation of important data.

2.4.4 Applications

Applications of crowd wisdom exist in three general categories: Prediction markets, Delphi methods and extended traditional opinion polls.
Prediction- or decision markets are speculative and virtual markets based on a betting structure. They ask questions like “Who do you think will win the election?” and predict outcomes rather well, in many cases even better than opinion polls. Several companies offer prediction market places to predict project completion dates or the market potential for new ideas. An adapted prediction market is also used in some project management software such as Yanomo (http://www.yanomo.com) to let team members predict a project's "real" deadline and budget. Consecutively Yanomo manages projects, collaborates across teams and tracks activities. It bases on the principle of team sourcing and therefore on the wisdom of the crowd, which states that the entire team knows more about the project than the manager alone.
Another good example of crowdsourcing is the online game “Fold it!” (http://fold.it). Online gamers are motivated to re-structure proteins in 3D ensuring energy minimum. Bio-Scientists can use the results to solve "real-world" problems.
The Delphi method is important for an estimation workshop, so it will be treated more deeply in the following chapter:

<u>The Delphi Method - a Way to Catch the Wisdom of the Crowd</u>

The Delphi method (see Figure 9) is a systematic, interactive prediction method, which is built on a panel of independent experts.
The accurately selected experts answer questionnaires in two or more rounds. After each round, a facilitator provides an anonymous summary of the experts’ forecasts from the previous round as well as the reasons they offer for their judgments.
Therefore participants are encouraged to revise their earlier answers in light of the replies of other members of the group.
During the process the range of the answers will decrease and the group will converge towards the most correct answer. Many of the consensus predictions have proven to be more accurate than forecasts made by individuals.

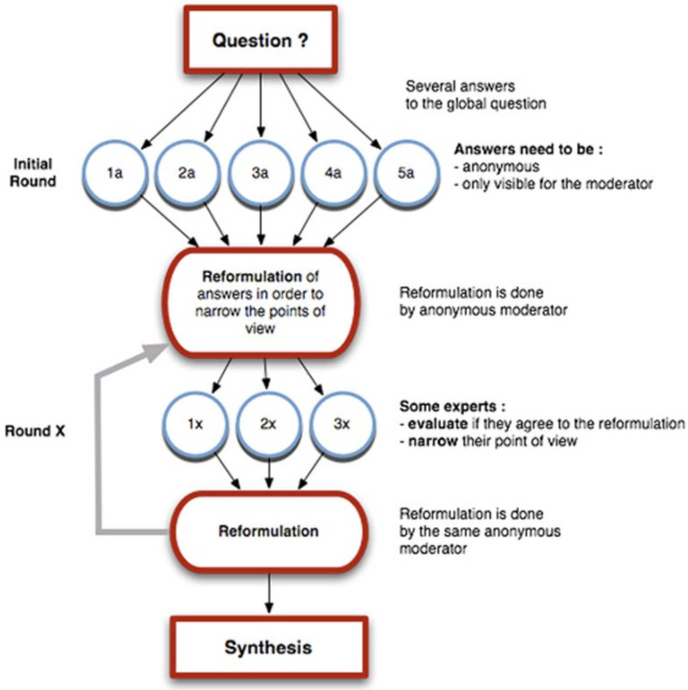


Figure 9: The principle of a Delphi method.

The key characteristics of this method are:

- Feedback loops force intensive examination of the own estimations
- Independence of the participants (no herding effect)
- Decentralization of participants
- Diverse structuring of group
- Controlling and aggregation of information flow

For an estimation workshop there are two practical ways:

- Large Circuit: Email-Loops etc.
- Small Circuit: Estimate-Talk with participant-Estimate

2.5 Estimation and the Power of Diversity[65]

The payoff of this chapter will be some theorems that should explain the centre of crowd wisdom, diversity. These theorems are mathematical facts and show, why crowds must be wiser then people within the crowd and why in particular the crowd does so well. Furthermore crowds of experts (intentionally they may predict best) and the effect of incentives on prediction accuracy will be investigated. This is a more detailed approach then Surowiecki did.

So why can teams of people find better solutions than brilliant individuals working alone? The answer lies in diversity. It is not the difference in what we look like outside, but what we look like within, in other words our distinct tools and abilities.

A book called *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools and Societies* by Scott E. Page examines this topic and reveals that progress and innovation may depend less on lone thinkers with enormous IQs than on diverse people working together and using their individuality. In particular and different to other published books or papers this publication uses mathematical modelling and case studies to show how variety in staffing produces organizational strength. Furthermore Page shows how groups that display a

[65] Cp. Page (2007)

range of perspectives outperform groups of like-minded experts. Diversity yields superior outcomes, whether you are talking about citizens in a democracy or scientists in the laboratory. Also he examines practical ways to apply diversity's logic to a bunch of problems.[66]

2.5.1 Expert or Crowd?

The answer to that question depends on the underlying problem. For example if your plumbing is damaged, you will need a plumber rather than a social worker, a German teacher and a politician working together. But diversity becomes more important when the problem is complex and given rules cannot solve it.

It is very common to say “let`s form an expert group!” to solve emerging problems, but experts are not the only game in town. This notion is perfectly introduced by Philipp E. Tetlock. His primary finding is that (political) experts are poor forecasters. He demonstrated this with a large sample of forecasts and with comparison to reasonable alternatives.[67] Furthermore he stated that the best way is to combine “foxes” (=flexible non-experts) and “hedgehogs” (=inflexible experts) notions to get more accuracy. In other words: his results lead to accumulate diverse information sources.

On the other hand crowds based on a majority of non-experts can do very well. This will be shown later on (in Chapter 2.5.3) by three distinct problems, which give an insight in the power of crowd wisdom. The first one is called needle-in-the-haystack problem. Here some people in the crowd know the answer while many, if not most, do not (e.g. “Who wants to be a millionaire?”). The second one is an estimation problem, where one person knows the correct answer, but the crowd does not (e.g. estimation of an ox-weight). Third problem turns out to be a prediction, where no one knows the correct answer, because it has to be revealed in the future (e.g. duration estimation)[68]. To discuss these three problems we will need a so-called toolbox of diversity at first (see next chapter).

2.5.2 The Toolbox of Diversity

More often than not, when organizations discuss diversity, they refer to social identity diversity, for example gender, race, religion, age, etc. Social identity diversity and cognitive diversity are certainly not the same. The ultimate goal is nevertheless cognitive diversity.[69]

Page unpacks diversity into four frameworks in his book, which he calls the toolbox of diversity.[70] Plainly told these frameworks are explaining how we differ cognitively.

<u>Perspectives: Ways of Seeing Things and Solving Problems.</u>

A perspective switch and a new way of solving a problem are often very effective. A simple example is the change from Cartesian coordinates to Polar coordinates for area calculation of rotation-symmetrical 2D shapes.

<u>Heuristics: Ways of Generating Solutions to Problems.</u>[71]

All in all it is obvious that more heads have more different solution tools in mind. A good example hereby is finding the correct rules for uncompleted number rows often found in IQ tests:[72]

a) 1 2 3 5 _ 13

Solution: **8** (Fibonacci), the underlying rule: $x_{i+2} - x_{i+1} = x$ (differences)

66 Cp. N.N. (Book review: "The difference…" by Scott E. Page, 2010)
67 Cp. Tschoegl et.al. (2007)
68 Cp. Mauboussin (2007)
69 Cp. Mauboussin (2007)
70 Cp. Page (2007)
71 Cp. Page (2007)
72 Cp. Page (2007)

b) 1 4 _ 16 25 36
Solution: **9**, the underlying rule: $x_i^2 = x$ (squares)

c) 1 2 6 _ 1806
Solution: **42**, the underlying rule: $x_{i+2} - x_{i+1} = x_i^2$ (differences of squares)

The last one is just a combination of the first two heuristics. So if you have one person with one heuristic (subtraction) in mind and a second person with a second heuristic (squaring down) in mind, you may get for free a third heuristic (subtract the squares!). You can see some sort of superadditivity effect. This is very important to understand, because some people tend to treat diversity like a diverse portfolio in the stock market: You get a specific problem, so you need a diverse group to find the specific expert. But in fact if you have a problem, you need the superadditivity of tools or heuristics to solve it faster and easier.

Interpretations: Ways of Appropriate Categorizing the World (e.g. your Perspective).[73]

Or easy said: "we lump to live!" (→ *lumping*). If we treat every experienced event as an individual, idiosyncratic one, we will basically struggle because of data overflow.
For instance placing the following food items in piles:

Table 11: Food items.

Broccoli	Canned Beets	Ahi Tuna	Fennel
Fresh Salmon	Arugula	Sea Bass	Niman Pork
Spam	Canned Posole	Carrots	Canned Salmon

A BOBO (BOurgeois BOhemian) may sort them in three categories:

Table 12: BOBO selection.

Veggies	**Fish & Meat**	**Canned Stuff**
Broccoli	Fresh Salmon	Canned Salmon
Carrots	Ahi Tuna	Spam
Arugula	Niman Pork	Canned Beets
Fennel	Sea Bass	Canned Posole

A Hillbilly may sort them in three other categories:

Table 13: Hillbilly selection.

Veggies	**Fish & Meat**	**Weird Stuff**
Broccoli	Fresh Salmon	Arugula
Carrots	Canned Salmon	Fennel
Canned Beets	Spam	Ahi Tuna
	Niman Pork	Canned Posole
		Sea Bass

So the different piles are nothing else than interpretations of the world. The point is that people categorize the world differently and so they have different predictions of what will come next. This circumstance leads us to the last point:

73 Cp. Page (2007)

Predictive Models: Ways of Figuring Out What Happens Next.[74]

Predictive models make inferences from the categories we create. If we lump differently, then we are likely to predict differently. For instance one might say: "Facebook is pure recreational activity", another one might say "Facebook is an effective election campaigning tool". The former one is possible still working at Starbucks, the latter one could be President of the United States.

2.5.3 An Insight in Diversity Power: 3 Problems[75]

Problem One: Who Wants to Be a Millionaire?

The first problem is a needle-in-a-haystack problem. There is an answer and some members of the crowd know what it is. The value of diversity is easy to see here. The remarkable fact is it does not take many people knowing the answer (or even having a better than random chance to guess the right answer) for the correct answer to emerge.
To illustrate the point, I borrow Page's example from *The Difference*.[76] He hypothetically presents the following question to a crowd:

Which person from the following list was not a member of the Monkees (a 1960s pop band)?
(A) Peter Tork
(B) Davy Jones
(C) Roger Noll
(D) Michael Nesmith

The non-Monkee is Roger Noll, a Stanford economist.
Now imagine a crowd of 100 people with knowledge distributed as follows:

- 7 know all 3 of the Monkees.
- 10 know 2 of the Monkees.
- 15 know 1 of the Monkees.
- 68 have no clue.

In other words, less than 10 per cent of the crowd knows the answer, and over two-thirds do not know the Monkees at all. As assumption we can say that individuals without the right answer vote randomly. As next step you can make the following breakdown:

- The 7 who know all the Monkees vote for Noll.
- 5 of the 10 who know 2 of the Monkees will vote for Noll.
- 5 of the 15 who know 1 of the Monkees will vote for Noll.
- 17 of the 68 clueless will vote for Noll.

Therefore Noll will get 34 votes versus 22 votes for each of the other choices. The crowd easily identifies the non-Monkee. We could add even more clueless people without violating the result: while the percentage margin by which Noll wins would decline, he would be the selection nonetheless.
Two variables are important: the percentage of the crowd who know the answer and the degree of randomness in the answers. Hereby randomness is more important than accuracy. A surprisingly small percentage of the population can know the answer and the population itself will be right with high randomness. Deviations from randomness will create less-than-perfect crowd answers, but still very good ones.
A popular problem, which is the same type as the above one, is simply asking a World Wide Web search engine for something. The engine uses rankings based on the wisdom of

[74] Cp. Page (2007)
[75] Cp. Mauboussin (2007)
[76] Cp. Page (2007)

crowds and answers most routine questions like the correct names of the Monkees very quickly.

Problem Two: Guessing the Right Number of Jelly Beans[77]

The second problem deals with estimating a state. Hence only one person (the questioner) knows the answer and none of the problem solvers do. A classic example of this problem is asking for a guess concerning the number of jelly beans in a jar. The core of this experiment is handled in Page`s book and it is called the Diversity Prediction Theorem.
Basically we want to define how far a crowd, an expert, a non-expert etc. is away from reality. Secondly we want a comparison, who does better. So we have to define some parameter to measure these differences. The solution is to build squared errors. Simple subtraction would be not enough, because there will be a cancelling out, if you have positive and negative values. In plain language: if one dart is thrown 5 cm left to the bull and one dart 5 cm right to the bull, you certainly do not hit bull`s eye. Equation 3 gives an example.

$$(\text{predicted value} - \text{actual value})^2 = \text{squared error of prediction} \qquad 3$$

Squared errors as a measure of accuracy are the mathematical foundation for the theorem. The bigger the error the worse estimating is.
The following simply example is applying this squared errors technique. Imagine some sort of low temperature prediction (in °F) for different American cities. You have two expert estimations, NOAA (National Oceanic and Atmospheric Administration) and WEDC (weather.com). The average (Ave) of these two experts will be the crowds guessing. The actual temperature is given in the last row.

Table 14: Predicted and actual temperatures in different cities.

	New York	Chicago	Los Angeles
NOAA	16	6	40
WEDC	10	14	46
Ave (Crowd)	13	10	43
Actual	18	16	39

Firstly the average individual error, that combines the squared errors of all of the participants, is calculated.

NOAA: $(16-18)^2 + (6-16)^2 + (40-39)^2 = 105$
WEDC: $(10-18)^2 + (14-16)^2 + (46-39)^2 = 117$
Average: $(105 + 117)/2 = 111$

Plainly said it captures the average accuracy of the individual guesses. So this is a measure how smart people are and how well they can guess. We have to calculate the crowd error:

Ave (Crowd): $(13-18)^2 + (10-16)^2 + (43-39)^2 = 77$

Firstly you can see the crowd is better than the average individual guess. Secondly to get the variance of the predictions as comparison between experts and crowd, you have to calculate another difference:

$(\text{NOAA - Crowd})^2 = (16-13)^2 + (6-10)^2 + (40-43)^2 = 34$

77 Cp. Page (2007)

$(\text{WEDC - Crowd})^2 = (10\text{-}13)^2 + (14 - 10)^2 + (46\text{-}43)^2 = 34$

These results are called "Prediction Diversity". Prediction diversity combines the squared difference between the individuals guess and the crowd guess. It reflects the dispersion of guesses or how different they are. In other words: It is a parameter how the estimators differ in mind and how disperse the estimations are.
What Page found out is that you can calculate the collective ability or the crowd error (that is simply the difference between the actual value and the crowd guess) as follows by equation 4:

crowd error = actual value – crowd guess = average individual error – prediction diversity	4

This equation is called Diversity Prediction Theorem (DPT) and has some important implications. At first a crowds guess is depended on individual ability and collective difference (=diversity) in equal parts. You can reduce crowd error by either increasing accuracy by a unit or by increasing diversity by a unit. That is maybe surprising for a society that is tending to praise ability. Secondly a diverse crowd will always predict more accurately than the average individual. Consecutively the crowd always predicts better than the people in it. Finally, while not a formal implication of the theorem, it is true that the collective is often better than even the best of the individuals. A diverse crowd always beats the average individual and frequently beats everyone. And the individuals, who do beat the crowd, generally change, suggesting they are more of a statistical remainder than super-smart people.
To illustrate the Prediction Diversity Theorem the following experiment was done at Columbia Business School in 2007:

Experiment 1: Jelly Beans

73 students had to guess independently the number of jelly beans in a jar. As incentive/punishment there was a $20 money reward for best guessing and a $5 penalty for worst guessing. In Figure 10 Column A shows the individual estimations. The crowd`s guess that is calculated by the mean of all guesses was 1.151 beans. The actual bean number was 1.116. So the crowd was off by 35 beans or 3,1 %. Column B-E present differences and squared errors to calculate the DPT.
I want to run through an example with Student 1 to exemplify how this works. Her guess was 250 beans. Hence the actual value was 1.116 beans, the difference to reality was -866 beans. Then we have to square this difference and get 749,9. Next step is to do this for each student and take the average for the whole class. This is the average individual error with the value of 490,9 (remember: the more accurate each student estimates the smaller this error is).
Now we can compare Student 1`s guess (250) with the class` average guess (1.151). The difference is -901 and squared 811,8 (column E). Again we do this for each student. Thus we get the prediction diversity with the value of 489,7 (remember: the more disperse the estimations are the larger the prediction diversity is).

Finally we bring all together and calculate the crowd error or collective error by means of the DPT:

crowd error = average individual error – prediction diversity = 490,9 – 489,7 = 1,2

You can see that the crowd is guessing more accurate than the average individual. As a result check you can calculate the square root of the collective error:

$\sqrt{1.258} \sim 35$ beans.

35 beans are the difference between the crowd`s guess and the actual bean number in the jar. Only two estimators were better than that (both voted for 1.120 beans, so they were +4 beans off).
Again a diverse crowd has beaten the average individual, but as already mentioned not always could beat everyone. The individuals, who do beat the crowd, generally change. They seem to be more a random phenomenon than some sort of super-estimator. Therefore you cannot predict the person or persons who may beat the crowd. The crowd`s guess is the nearest assured value to reality you can detect.

Problem Three: "And the Oscar Goes To…".[78]

The last problem deals with a prediction, where the answer is unknown and will be revealed in the future. This problem is almost like a combination of the first two. As with the Monkees` example (problem one), some people probably have better predictive models than others (i.e. they know the film industry), hence allowing the most likely answer to arise.
To get this answer surely we combine a lot of diversity (some must have the right predictive models) with little predictive accuracy (most estimators know not much about films). This step is concerning problem two.
The following example is an experiment conducted at Columbia Business School with students and will show how the DPT also works on this type of problem.

Experiment 2: Oscar Favourites

Prior to the Academy Awards Ceremony 2007 a two-sided sheet with some known and some less known prize categories was distributed. The goal was to predict the winner in those categories. The students could contribute 1 dollar to a pot (with the winner getting the proceeds as incentive) and then select their favourite. The goal was to win the pot, not to choose the emotional favourites.
Front page are the six most popular Academy Awards categories:

- Best actor
- Best actress
- Best supporting actor
- Best supporting actress
- Best film
- Best director

On the back are six less known categories:

- Best adapted screenplay
- Best cinematography
- Best film editing
- Best music (original score)
- Best documentary
- Best art direction

Figure 11 shows the categories and a 1 for each right guess and a 0 for the wrong ones. Furthermore all differences (highest possible value is 12, lowest is 0) and squared errors are inscribed.

[78] Cp. Mauboussin (2007)

The results show the Diversity Prediction Theorem at work. The crowd guess defined as the most likely value in each category got 11 of 12 winners correct, including all 6 lesser known categories. Two students win the pot for the most correct answer (9 of 12 correct) and the average student (= average individual error) got just 5 of 12 right.
This result illustrates again how accuracy and diversity combine to produce an individual-beating answer.
While the logic of diversity certainly provides some important insights and useful models, we by no means fully understand the wisdom of crowds. But the given models and examples provide a concrete step in the right direction.

Finally Page has summarized the power of diversity in a very short and impressive manner:

"These theorems that when solving problems diversity can trump ability and that when making predictions diversity matters just as much as ability are not political statements. They are mathematical truths."[79]

[79] Page (2007)

Student	COLUMN A Guess	COLUMN B "Difference From Actual"	COLUMN C Squared "Difference From Actual"	COLUMN D "Difference From Average"	COLUMN E Squared "Difference From Average"
1	250	-866	749,956	-901	811,801
2	315	-801	641,601	-836	698,896
3	399	-717	514,089	-752	565,504
4	400	-716	512,656	-751	564,001
5	420	-696	484,416	-731	534,361
6	437	-679	461,041	-714	509,796
7	479	-637	405,769	-672	451,584
8	500	-616	379,456	-651	423,801
9	540	-576	331,776	-611	373,321
10	585	-531	281,961	-566	320,356
11	600	-516	266,256	-551	303,601
12	600	-516	266,256	-551	303,601
13	604	-512	262,144	-547	299,209
14	616	-500	250,000	-535	286,225
15	624	-492	242,064	-527	277,729
16	632	-484	234,256	-519	269,361
17	645	-471	221,841	-506	256,036
18	650	-466	217,156	-501	251,001
19	651	-465	216,225	-500	250,000
20	699	-417	173,889	-452	204,304
21	721	-395	156,025	-430	184,900
22	723	-393	154,449	-428	183,184
23	734	-382	145,924	-417	173,889
24	750	-366	133,956	-401	160,801
25	750	-366	133,956	-401	160,801
26	750	-366	133,956	-401	160,801
27	750	-366	133,956	-401	160,801
28	768	-348	121,104	-383	146,689
29	780	-336	112,896	-371	137,641
30	800	-316	99,856	-351	123,201
31	800	-316	99,856	-351	123,201
32	820	-296	87,616	-331	109,561
33	850	-266	70,756	-301	90,601
34	874	-242	58,564	-277	76,729
35	876	-240	57,600	-275	75,625
36	900	-216	46,656	-251	63,001
37	900	-216	46,656	-251	63,001
38	900	-216	46,656	-251	63,001
39	1,000	-116	13,456	-151	22,801
40	1,000	-116	13,456	-151	22,801
41	1,008	-108	11,664	-143	20,449
42	1,120	4	16	-31	961
43	1,120	4	16	-31	961
44	1,152	36	1,296	1	1
45	1,234	118	13,924	83	6,889
46	1,234	118	13,924	83	6,889
47	1,250	134	17,956	99	9,801
48	1,250	134	17,956	99	9,801
49	1,260	144	20,736	109	11,881
50	1,288	172	29,584	137	18,769
51	1,300	184	33,856	149	22,201
52	1,400	284	80,656	249	62,001
53	1,500	384	147,456	349	121,801
54	1,500	384	147,456	349	121,801
55	1,500	384	147,456	349	121,801
56	1,523	407	165,649	372	138,384
57	1,564	448	200,704	413	170,569
58	1,575	459	210,681	424	179,776
59	1,580	464	215,296	429	184,041
60	1,583	467	218,089	432	186,624
61	1,588	472	222,784	437	190,969
62	1,700	584	341,056	549	301,401
63	1,732	616	379,456	581	337,561
64	1,872	756	571,536	721	519,841
65	1,896	780	608,400	745	555,025
66	1,899	783	613,089	748	559,504
67	1,963	847	717,409	812	659,344
68	2,000	884	781,456	849	720,801
69	2,250	1,134	1,285,956	1,099	1,207,801
70	3,000	1,884	3,549,456	1,849	3,418,801
71	3,000	1,884	3,549,456	1,849	3,418,801
72	3,024	1,908	3,640,464	1,873	3,508,129
73	4,100	2,984	8,904,256	2,949	8,696,601
Average	1,151		490,949		489,692

"Actual" # of Jelly Beans	1,116
"Average Guess" of # Jelly Beans	1,151
Average Individual Error (Average of Column C)	490,949
Prediction Diversity (Average of Column E)	489,692
Collective Error ("Average of Column A"-"Actual")^2	1,258

CHECKS:

Collective Error = Average Individual Error - Prediction Diversity

1,258 = 490,949 - 489,692

√Collective Error = ABS ["Average Guess" - "Actual"]

√1,258 = ABS [1,151 - 1,116] ≈ 35

Figure 10: Statistical approach to the Jelly Beans Problem.[80]

[80] Source: Mauboussin (2007)

	COLUMN A (Guesses)												COLUMN B	COLUMN C
Student	Lead Actor	Lead Actress	Support Actor	Support Actress	Mot. Pic.	Directing	Screenplay	Cinematography	Film Editing	Music	Documentry	Art Direction	Squared "Difference From Actual"	Squared "Difference From Average"
1	0	0	1	1	1	0	0	0	0	0	1	1	7.0	3.0
2	1	1	0	1	0	0	0	1	0	0	1	0	7.0	2.1
3	0	1	0	1	1	1	1	1	0	1	1	1	3.0	3.4
4	1	0	1	0	0	0	0	0	0	0	1	0	9.0	2.4
5	0	1	0	1	1	1	0	1	1	1	1	1	3.0	3.5
6	1	1	0	1	0	1	0	1	0	1	1	0	5.0	2.6
7	0	1	0	0	0	0	0	0	1	0	1	0	9.0	2.2
8	0	1	0	1	0	0	0	0	0	0	0	0	10.0	2.1
9	0	0	0	1	1	1	1	0	1	0	0	0	7.0	3.2
10	1	1	0	0	1	0	0	1	0	0	1	0	7.0	2.2
11	1	1	1	0	0	0	1	1	0	0	0	0	7.0	3.1
12	0	0	0	1	0	1	0	1	0	0	1	0	8.0	2.3
13	1	1	0	1	0	1	1	1	0	0	1	1	4.0	2.8
14	0	1	0	1	0	0	1	1	1	1	0	1	5.0	4.0
15	1	1	1	1	1	1	0	0	0	0	1	1	4.0	2.9
16	1	1	0	1	0	0	1	1	0	0	1	1	5.0	2.8
17	1	1	0	1	0	1	1	0	0	1	0	0	6.0	2.9
18	1	1	0	1	0	0	1	0	1	1	1	0	5.0	3.1
19	0	0	1	0	0	0	0	0	0	1	0	0	10.0	3.2
20	1	1	0	0	0	1	0	1	0	0	1	1	6.0	2.4
21	1	0	0	0	0	1	1	0	0	0	1	0	8.0	2.3
22	0	0	0	0	0	0	0	0	0	1	1	0	10.0	2.4
23	0	1	0	1	1	0	1	0	1	0	1	0	6.0	2.7
24	1	1	0	1	1	1	1	0	0	0	0	0	6.0	2.6
25	1	1	0	1	0	0	0	1	0	0	0	1	7.0	2.8
26	0	0	0	0	1	0	0	1	0	0	1	1	8.0	2.7
27	1	1	0	0	1	1	0	0	0	0	1	0	7.0	1.9
28	1	1	0	0	1	0	0	0	0	1	1	0	7.0	2.4
29	1	1	0	1	1	1	1	0	0	0	1	0	5.0	2.3
30	0	1	1	0	0	0	0	1	0	0	1	1	7.0	2.9
31	0	0	1	1	0	0	1	0	0	0	1	1	7.0	3.1
32	0	0	1	0	1	0	1	0	0	0	1	0	8.0	2.9
33	1	1	1	0	1	1	0	0	1	0	0	0	6.0	3.1
34	0	1	0	0	1	1	1	0	1	0	1	0	6.0	2.6
35	0	0	0	0	1	0	0	1	0	0	0	0	10.0	2.6
36	0	0	0	0	0	1	1	0	0	0	1	0	9.0	2.2
37	0	0	1	0	0	1	0	0	0	0	0	1	9.0	3.1
38	1	0	0	1	0	0	0	0	0	0	0	0	10.0	2.3
39	1	1	0	1	1	0	0	1	0	0	1	0	6.0	2.3
40	0	0	0	0	0	1	0	0	1	1	1	0	8.0	2.9
41	1	0	1	0	0	0	0	0	0	1	1	0	8.0	2.9
42	1	1	0	1	0	0	1	0	0	1	1	0	6.0	2.6
43	1	1	0	0	0	0	0	0	1	0	0	0	9.0	2.5
44	1	0	0	1	1	1	1	0	1	1	1	0	4.0	3.4
45	0	0	0	0	1	0	0	0	0	0	0	0	11.0	2.4
46	1	1	0	0	0	1	1	1	0	1	1	1	4.0	3.2
47	0	1	1	0	0	1	0	1	0	0	1	1	6.0	2.9
48	0	1	0	0	0	1	0	1	0	0	1	1	7.0	2.4
49	0	0	1	0	0	1	0	0	1	0	0	0	9.0	3.1
50	1	1	0	1	1	1	0	0	1	0	1	0	5.0	2.4
51	0	1	1	0	0	1	0	0	1	0	0	1	7.0	3.3
52	0	0	0	0	1	0	0	0	1	0	0	0	10.0	2.8
53	0	0	0	0	1	1	0	0	0	0	0	0	10.0	2.4
54	0	0	0	1	1	1	0	1	0	0	0	0	8.0	2.7
Avg.	0.48	0.59	0.26	0.46	0.43	0.50	0.35	0.37	0.28	0.26	0.65	0.31	7.1	7.0

Average Individual Error (Average of Column B)	7.056
Prediction Diversity (Average of Column C)	2.724
CHECK: Collective Error = Average Individual Error - Prediction Diversity Collective Error = 7.056 - 2.724	4.332

Figure 11: Statistical examination for the Oscar Problem for all categories.[81]

[81] Source: Mauboussin (2007)

The Academy Awards Experiment—One Category

Category: "Lead Actor" Student	COLUMN A Guess	COLUMN B Squared "Difference From Actual"	COLUMN C Squared "Difference From Average"
1	0	1	0.232
2	1	0	0.269
3	0	1	0.232
4	1	0	0.269
5	0	1	0.232
6	1	0	0.269
7	0	1	0.232
8	0	1	0.232
9	0	1	0.232
10	1	0	0.269
11	1	0	0.269
12	0	1	0.232
13	1	0	0.269
14	0	1	0.232
15	1	0	0.269
16	1	0	0.269
17	1	0	0.269
18	1	0	0.269
19	0	1	0.232
20	1	0	0.269
21	1	0	0.269
22	0	1	0.232
23	0	1	0.232
24	1	0	0.269
25	1	0	0.269
26	0	1	0.232
27	1	0	0.269
28	1	0	0.269
29	1	0	0.269
30	0	1	0.232
31	0	1	0.232
32	0	1	0.232
33	1	0	0.269
34	0	1	0.232
35	0	1	0.232
36	0	1	0.232
37	0	1	0.232
38	1	0	0.269
39	1	0	0.269
40	0	1	0.232
41	1	0	0.269
42	1	0	0.269
43	1	0	0.269
44	1	0	0.269
45	0	1	0.232
46	1	0	0.269
47	0	1	0.232
48	0	1	0.232
49	0	1	0.232
50	1	0	0.269
51	0	1	0.232
52	0	1	0.232
53	0	1	0.232
54	0	1	0.232
Average	0.4814815	0.2688615	0.250

"Actual" Winner	1
"Average Guess" of Winner	.4818
Average Individual Error (Average of Column B)	0.518519
Prediction Diversity (Average of Column C)	0.249657
Collective Error ("Average of Column A"-"Actual")^2	0.268861

CHECKS:

Collective Error = Average Individual Error - Prediction Diversity
.2689 = .5185 - .2497

√Collective Error = ABS ["Average Guess" - "Actual"]
√.2689 = ABS [.4815 - 1] ≈ .5185

Figure 12: Statistical examination for the Oscar Problem for one category[82].

[82] Source: Mauboussin (2007)

3 Company Survey

3.1 Objective

PSA is already applied in various companies, so at OMV as well. This means some related knowledge must be accumulated over time and it is reasonable to do some data mining.
The main target was to detect the right experts and the right way to grab their know-how in terms of conducting PSA practically.

3.2 Approach

3.2.1 Information Sources in a Company

Overall there are two different types of information sources:[83]

- Primary information sources (counting, measuring, asking etc.)
- Secondary information sources (interpreting of and deducing from information)

In this work an interrogation as primary source was chosen. That could be conducted in different ways showed in Figure 13:

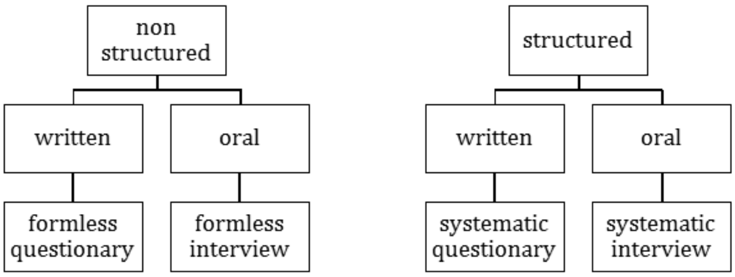


Figure 13: Interrogation concepts.

In our case a structured, oral form (telephone interview based on a questionary) was preferred. The reason for that was the possibility of standardization and comparability. Secondly the problem of misunderstanding some questions in written form like e-mails can occur. An oral and structured interview should avoid all these difficulties.
In order to do so OMV compiled a catalogue with contact data of experts, who offer great knowledge concerning PSA. This catalogue was used to make telephone interviews. Table 15 shows the way all interrogations were done.

Table 15: Guideline for telephone interviewing.

Task (in chronological order)	**Duration (approximation)**	**Tools**
Working out questionary and contact person list with expert panel	1 week	E-mail
Making appointments with contact persons	2 days	E-mail, Telephone
Interviewing	30 Min. to 60 Min. (1 Interview). Pay attention: to have all	Telephone

[83] Cp. Klaus; Krieger (1998)

	persons interviewed could take some time due to appointment spreading.	
Info wrap-up und conclusions	1 day	MS Word

3.3 Outcome and Lessons Learnt from Interviews

In this chapter the list of asked questions is presented and to every question a summary of given answers below.

<u>Have you ever done a PSA before (PERT, Monte Carlo Simulation etc.) und if so, what are your experiences with?</u>

PSA is very important to all interview partners and has become a widely used tool in practical project management. A deterministic approach is too optimistic or inexact and has its benefits only in the concept phase to make rough schedule drafts. An important fact is sometimes the lack of "native mathematics" in the project. So it is a common agreement that PSA should be transparent and well documented.
Furthermore all biases on the input value side (for example a too optimistic estimator) do have a big influence on the output. This should be captured in an estimation workshop to make estimators more aware of that.
Moreover constraints are an important factor. They are necessary for deadlines that cannot float in the schedule. But constraints should be kept to a minimum, because they could not be captured by probabilistic analysis due to their deterministic character. Negative constraints should be totally avoided.
In all cases MC simulation was and is used. If you are uncertain about your duration estimations, a PERT analysis can give you a raw, but often realistic approach. A well-thought logic diagram of your schedule (sequence of tasks etc.) is necessary too.
Mostly a top-down principle is applied to get your final schedule. This means you generate a target schedule at first, very rough and with most likely values. Then uncertainty is taken into account and as result you get a mean schedule, which should be regularly revised and communicated.
Basic project management tools are MS Project[84] and Oracle Primavera[85]. Primavera is industry standard and for more complex projects, MS Project has its credits in creating fast and less complex schedules.
For modelling uncertainty @Risk[86] and Futurenova are in use. Different to @Risk the latter is designed to be in „live action". This can help a team to evolve a comprehensive overview of a situation or project in terms of time and budget.
In general there should be no distinction between cost and schedule analysis, both base on the same principles and should take action simultaneously. Perhaps there will be a merging in the future to get synergy effects. Secondly you should work out your schedule with great foresight and intensive information gathering. This will pay off later, because working in a "no surprise area" is more efficient.

<u>If you have done a PSA before: what were typical pitfalls and the main difficulties you faced respectively?</u>

Employees should have statistical and practical understanding in doing PSA and a profound knowledge about distributions. For example: A new compressor has to be delivered

84 www.microsoft.com
85 www.oracle.com
86 www.palisade.com

in 2 years. Engineers with "triangular bias" estimate gently on the right side of the distribution (let`s say +10 months). Engineers with a tendency to tailored distributions estimate boldly, also including Black Swans (for example +4 years, it is not very likely, but nevertheless not impossible).
There are also different estimation individuals:

- realistic,
- cautious,
- management affine and
- bold.

That aspect of diversity should be balanced.
A well-planned and continuous communication of PSA is needed. It is not only a working tool, but also a way of thinking or in other words: PSA does not start in MS Project, but in the employee`s mind.
As often mentioned: "garbage in - garbage out!". Correct and extensive input information are crucial. Otherwise you have too many assumptions and this can harm your results.
Furthermore a too optimistic approach can take place. On the one hand you can detect this in a very small range between the P10 and P90 value (some risks are denied or not seen), on the other hand the use of a P10 to P30 value as mean value (should lie between P30 to P50) often occurs, but is not realistic at all. Reasons for over-optimistic statistics are a not well-defined project and the ambition and/or ignorance of basic PSA theory. If you have a very optimistic target schedule to motivate your co-workers, you must communicate this well.
A not often considered pitfall is the time loss in the beginning of the project, because a lack of motivation (as example: first oil is far away, no milestone pressure) is occurring.
Another aspect is that people tend to handle PSA information as sacrosanct, because these values are coming out of a black box they will never understand. But in fact it is more a working platform that should be used as reference system, where you can rely on for further planning. PSA is therefore another example for the need of a PDCA (Plan, Do, Check and Act!) system.
Moreover realistic values should come first followed by corresponding distributions, not the other way round. People tend to favour "convenient" distributions without looking at the data basement.
When taking a look at the project management side: the target and the scope of the project must be very clear. People have to buy-in. Intermediate milestones are necessary to keep a "motivation tension".

<u>In your opinion, what are the main factors that can cause uncertainties in the project schedule?</u>

Typical for all interview partners was some uncertainty according to partner approval. This means sometimes financial burdens are allocated between project partners, but approval from your own company is not sure (for example: board is on holidays).
If you take a look at BP: 378 risks concerning budget/schedule/quality are stored in a database and as suggestion it could be worthy to combine different risk matrices to find common risks. Consequently you could create an uncertainty/risk ranking, which leads to priorities and postpone actions later on.
Moreover one important uncertainty factor is the company itself, as an example: no proper project management.
Of course political surroundings can have a big influence. Typically been treated as Black Swans, they are not necessarily unlikely, because many oil projects are located in politically unstable countries.

All in all there are very different causes of uncertainty, let`s say contracting, financing, authorisation (bureaucracy), board decisions, delivery delays. Every project is different and therefore the main uncertainty drivers are changing every time.

For which projects does PSA make sense (budget, complexity)? Can you quantify it?

There is an overall agreement that PSA does make sense for all projects. For example recalling a rental contract from the day an office movement is finished. PSA is done for projects in a financial range of 1 Million to 1 Billion Euros and for schedule task numbers up to 10 000.

In which project phases (concept, pre-feasibility, feasibility and execution) would a PSA make sense?

PSA in the concept phase does not make sense, because the level of details is too low and there are often more versions of the same schedule considered. If you take a look at OMV: from Tollgate 5B (= defined general company milestone) on it is necessary.
In general a PSA should be started in the pre-feasibility phase and it has to be started at feasibility level. Furthermore you can say: the earlier the better, because you have more alternatives at the front-end. Mostly PSA is performed at the end of a phase. The best way would be in the middle, because you already have important information and enough time contingencies to go against.

How do the interview partners usually gather the input data (minimum, mode, maximum etc.), during the risk workshop or in advance? How many different opinions are considered?

Look at BP: a special one man-office is installed risking all the time. In addition a planning team runs a database, so many considerations, opinions and hints are stored. This system has its advantages, because risk workshops do not offer a proper data updating and it is further a problem, that in workshops sometimes the deterministic base schedule is discussed too long.
Look at OMV: here you can find a risk workshop, where ca. 12-15 (head quarter) experts are discussing the appropriate values. Additionally a basic and raw schedule is determined and then uncertainty data is gathered.
Overall a risk workshop/expert panel with a multi-disciplinary group is common. Major data sources are historical data (most accurate, database!) and expert judgement. Uncertainty data should be allocated in advance, when global schedule data is collected.
If you want to rely on expert judgement, sometimes single expert interviews (on a simplified Delphi method base) are preferred. This could counteract herd effects.

What is the usual number of tasks considered in the PSA? All tasks? Is lumping of tasks performed? Suggestions?

It normally depends on the planning stage:

- concept phase: ca. 100 Tasks (several scenarios),
- prefeasibility phase: ca. 250 Tasks (ca. 2 scenarios) and
- Tollgate 5B: ca. 1000 and more (can go up to 10.000 tasks).

But there were also notions that the schedule should have a maximum task number of 200 and you better pay attention to tied networks with no logically gaps. Otherwise you can get problems with complexity and logically breakdowns during simulation.
At the feasibility stage all tasks should be included in a PSA. In the execution phase you should have one man-one work package structures.
The impact of lumping is generally not known, there was no exact opinion on this topic.

Which distributions are in use? How are they selected? Suggestions?

It heavily depends on the situation. Frequently used distributions are:

- Triangular
- Normal
- Long-tailored like Lognormal
- Uniform
- Trapezoid
- Yes-No (discrete, only for risks)
- Beta (constraints in the beginning of a task)

Some say that you should take a Triangular one, where it is possible. This approximates real data best. If you have absolutely no information, take a Normal distribution.
Distribution selection is based on experience, software settings and trial and error approach. Literature has less impact. But again it is very important to get some uncertainty key data as distribution parameters first before you look out for the right distributions.
Sometimes the curves are not expanded enough (too small range) and therefore to optimistic. A counter strategy could be asking the question "In one of hundred cases, what could happen?" Then you get an estimation of fringe values.
Overall this is a complicated topic and a preference is to take more simple and understandable distributions.

How the result of the PSA should be implemented in the further course of the project? Actual situation? Suggestions?

At BP you will find a PDCA loop, so everyone can have his project evaluated all the time. This means you get for example a quality-quantity-link: anchor chain of a production ship is broken (quality), how does this influence my whole schedule and/or single tasks (quantity)? Consecutively you can achieve time line scenarios (with/without mitigation etc.).
Generally there is a PSA update two times a year, for this some single talks with persons responsible for the work package or a small work team is sufficient to gather information.

What "end dates" for the projects are communicated to the board motion after having performed PSA? P50? Deterministic value? Deterministic value + contingency?

In general risk calculations are requested by the board, no deterministic schedule even with contingencies will be accepted anymore. A "most likely schedule" could be used as target (basic) schedule.
As an optimum P70 to P80 values should be given to the board, generally the opposite takes place (P15 values) to satisfy board ambitions. Overall estimations are too optimistic (≈ P30). In real and in very rare cases P35 values were achieved.
The following counter strategy was suggested: high P values should be well explained to board, there has to be a value labelling: "unrisked", "P70 risked" etc. Additionally all statistics should be presented to the board, but an emphasis on the mean value is a reasonable approach, because this is the best approximation.
Always have in mind: transparent, profound estimations are commonly the most likely ones. Secondly: if you want to be very safe, keep a contingency of not less than ten per cent.

In general: what is the subjective effort/benefit-ratio for PSA? (Differentiate between people working in branch offices and people who give support from head office!)

For branch office people a budgeting without PSA is nearly impossible, because otherwise you cannot get a realistic and achievable plan. Consecutively this schedule can be sent with

all conscience. For the supporting head office this means a better ability to plan and a better visibility of the status quo.
Common sense is that you have, if well done, a win-win situation, which justifies quite an effort.

<u>Do you think a guideline as blueprint for conducting a PSA would be useful?</u>

You can find an overall agreement: yes, of course. But it must be transparent, thought through and usable!

4 Applied Probabilistic Schedule Analysis

Thus far plenty of theory concerning probabilistic schedule analysis was investigated and presented in this work. But do theory framework and its key approaches pass a reality check and furthermore can it be implemented in daily workflows?
A practical attempt to verify literature was already done with the company survey. Now results from this survey and literature`s key issues will be examined "in real" in the next chapter with MS Project, some test schedules and a special probabilistic analysis software called @Risk.

4.1 Material and Methods

4.1.1 MS Project

MS Project is a program designed and sold by Microsoft to manage projects in developing and analysing project plans, Gantt charts, budgets or resource allocations.

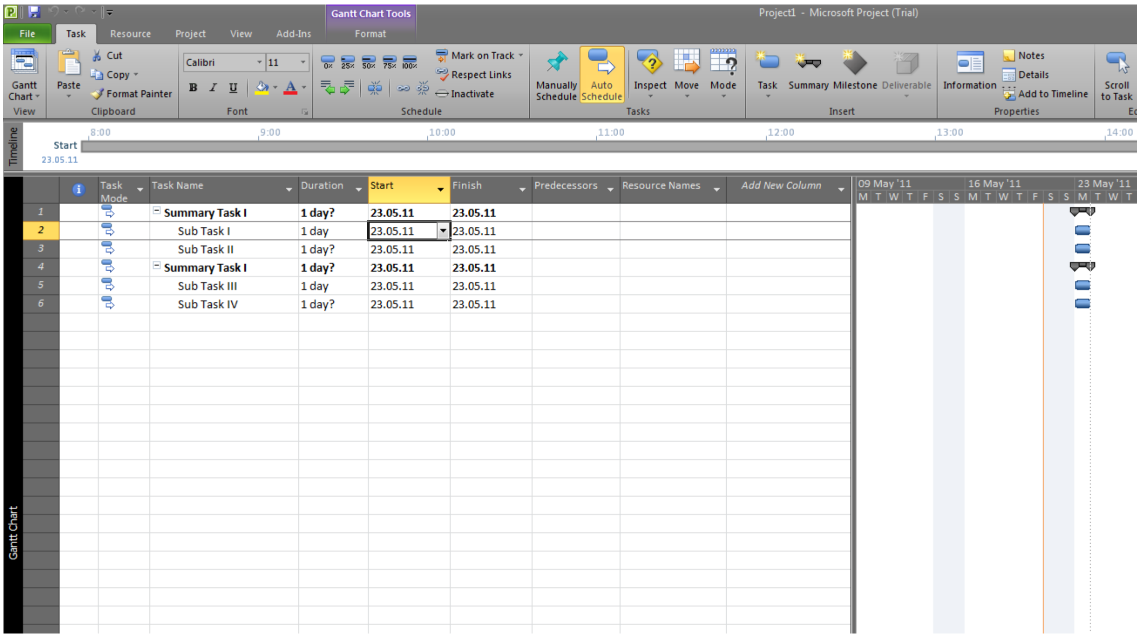


Figure 14: MS Project 2011, task view.

For more information on the product please visit www.microsoft.com. An excellent introduction for managing projects with MS Project 2010 is given by the book "Using Microsoft Project 2010" by Sonja Atchison[87].

4.1.2 Palisade`s @Risk for Project[88]

@Risk is an add-in for MS Project. It offers the possibility to turn a deterministic schedule into a probabilistic schedule by adding probability distributions to the single point estimates and running a Monte Carlo simulation.

87 Cp. Atchison (2011)
88 Cp. Wallace (2010)

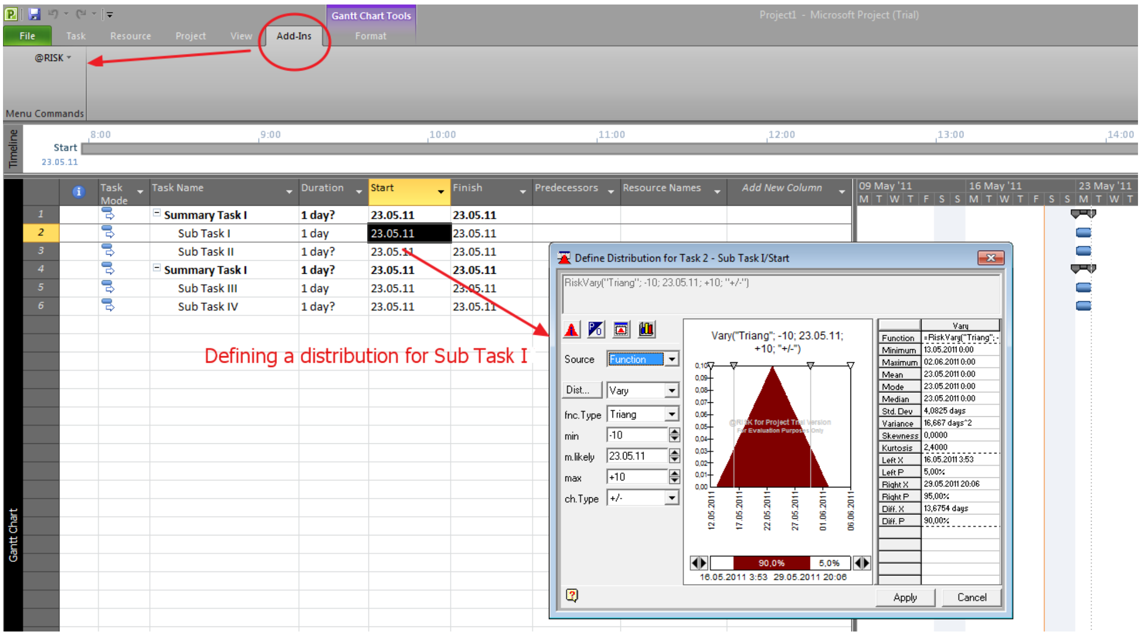


Figure 15: MS Project (Project I) with activated @Risk, the "Add distribution" window is opened.

For further information on @Risk two reference works are recommend: "Schedule risk and contingency using @Risk and probabilistic analysis"[89] by Ian Wallace (this chapter is mostly based on this paper) and "Guide to Using @Risk"[90] provided by www.palisade.com.

Basic Techniques to Use @Risk

Basic schedule with distributed task durations:

This is the basic way using @Risk. Uncertain task durations do not base on single values, but on distributed values (cp. Figure 15, Project I). Consecutively the Monte Carlo algorithm grabs samples within each duration distribution based on the density probability function defined by shape and area. Each time a sample is taken, it is returned to the schedule, so MS Project can recalculate the completion date. All in all, when we simulate 1000 different schedules, we will obtain 1000 potential completion dates along with a distribution of where most of them lie. The wider this distribution is, the more uncertain the outlook is.

IF/THEN conditions:

Supplemental this approach allows to include a risk register (a list of possibly risks linked with probability of occurrence and impact severity) into the schedule. The risks are entered as tasks with zero duration like milestones. These risk milestones are then linked to the tasks affected using Finish to Start (FS) dependency links.
Firstly the probabilities of occurrence for each risk are entered as variable in the model definition window using a Binomial distribution (cp. Figure 16, risk does not occur: value=0, risk occurs: value=1).

89 Cp. Wallace (2010)
90 Cp. N.N., Guide to Using @RISK, (2010)

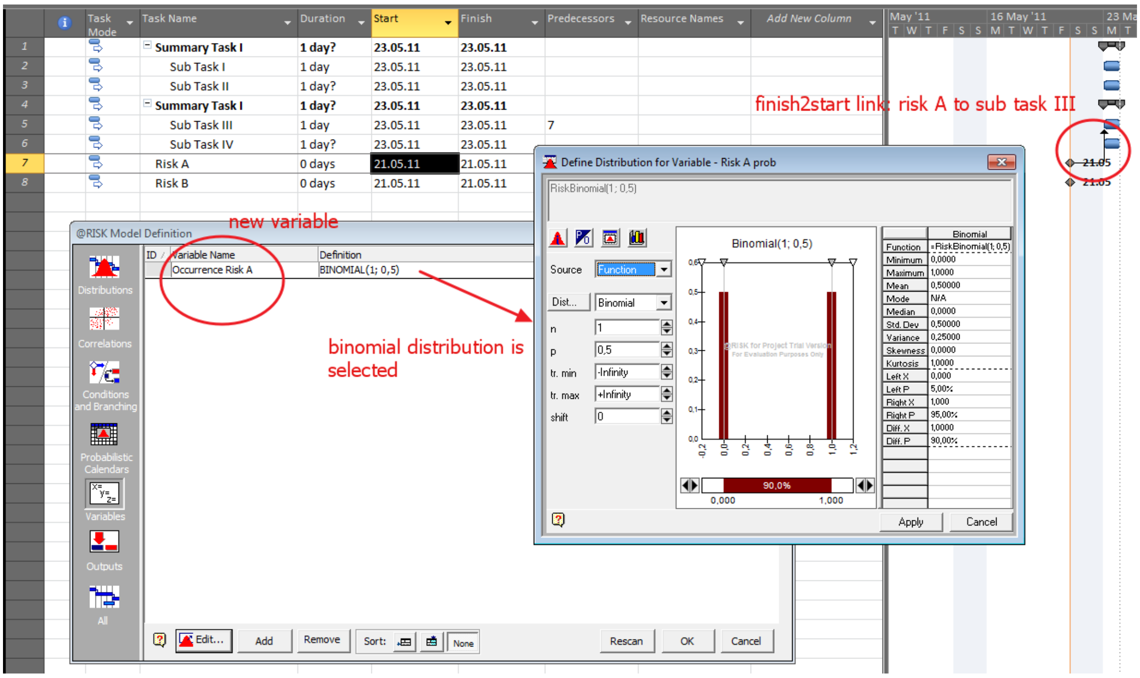


Figure 16: Defining a variable for risk A.

The variable will now turn either to 1 or to 0 during the Monte Carlo simulation. In the next step an IF/THEN statement for risk A is set up to link the occurrence to a probability distribution for the impact (cp. Figure 17).

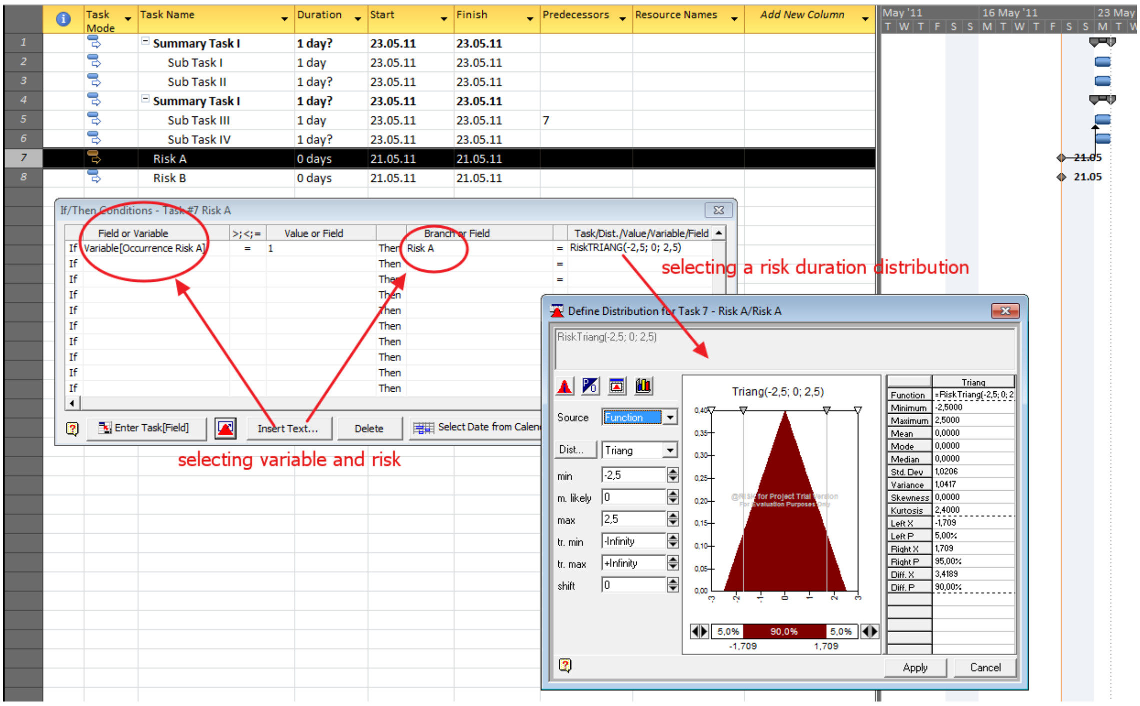


Figure 17: Choosing an IF/THEN condition.

If variable Occurrence Risk A=1, the probability distribution for the duration of the task "Risk A" is triangular distributed with the chosen parameters. Risk A will now take effect on sub task III in a certain amount in the associated Monte Carlo run.
IF/THEN is useful when you have a list of defined risks that can hit your schedule.

Probabilistic Branches:

This method can branch the schedule probabilistically to various task bundles, each of them with individual durations and probabilities.

For example the probabilistic branch leads either to the normal schedule sequence (90 %) or to an additional task sequence held in the risk task section (10 %) as showed in Figure 18:

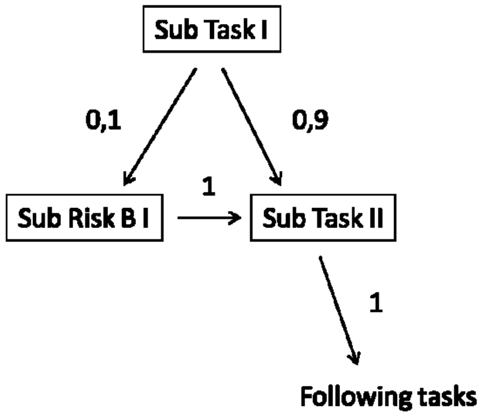


Figure 18: Probabilistic branching concept.

If the risk section is hit, task Sub Risk B I will occur with a PERT distribution (min=1 day, ml=4 days, max=10 days) and is relinked in the schedule (Sub Risk B I -> Sub Task II) again. Otherwise the original schedule is simulated.

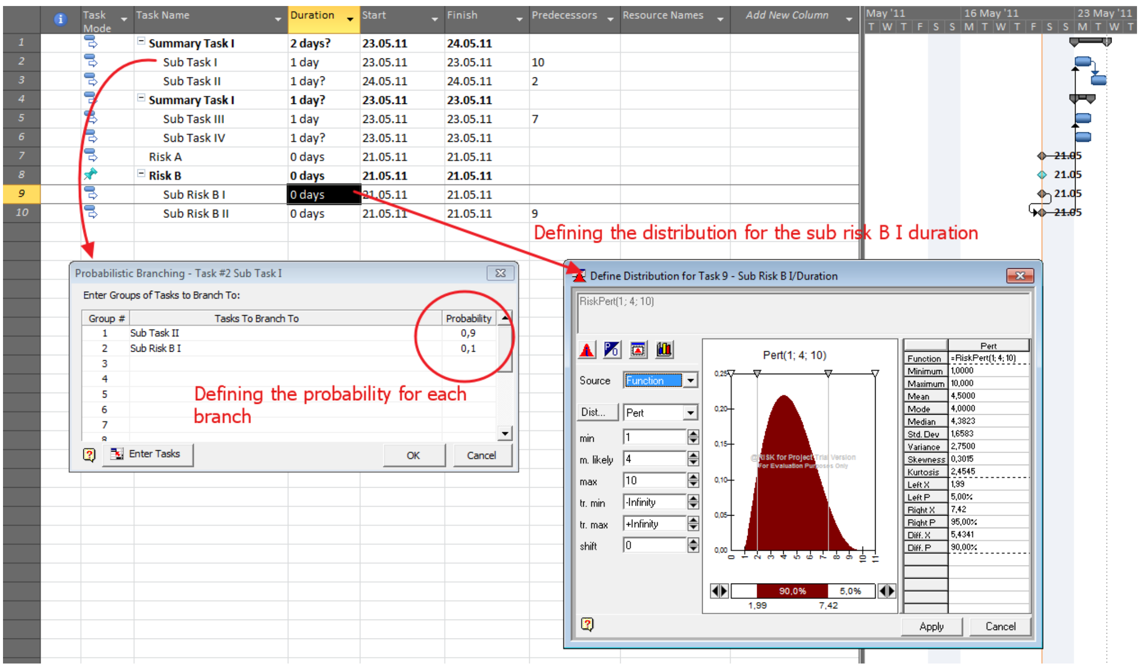


Figure 19: Probabilistic branching with @Risk.

It is important to link the risk task sequence back to the original plan so that it is correctly added to the schedule.

Probabilistic branches are useful to model alternatives or can give a comparison between two different scenarios.

Simple discrete distributions:

Discrete distributions are a simpler way than IF/THEN conditions to bring in risk events. Normally this is useful for accidents or other not well known risks, where discrete values define the possibilities and durations for the impact. If you not know when an accident can take place during the project, it is necessary to relink the impact to the final milestone, so that it can affect the finish date if it occurs.

Risk C has been set up as shown in Figure 20:

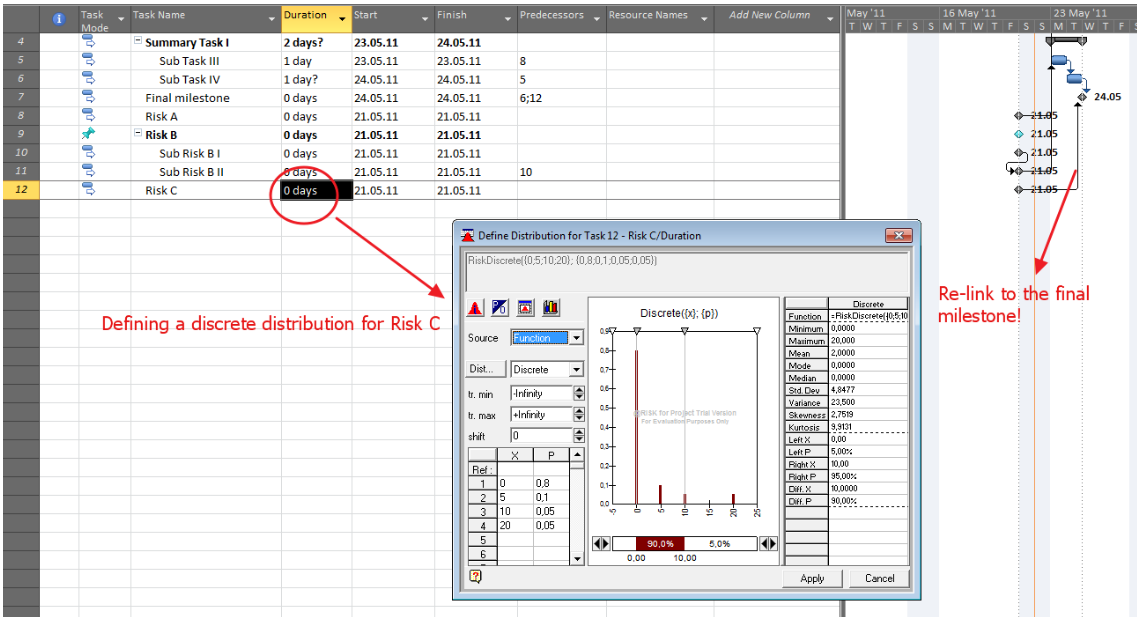


Figure 20: Risk C is discrete distributed and re-linked to the final milestone.

Now the probability for no impact is 90 %, the possibility for a 5 days delay is 10 %, for a 10 days delay 5 % and for 20 days delay 5 %.

Correlations:

Often durations or other parameters of different tasks are related, for example when duration of task A is long, duration of task B should also be long. Or when duration of task B is long, costs of task B are high at the same time. This is called correlation. More precisely, the distributions of the 2 tasks are still independent, but the sampling should be correlated. Correlations themselves are defined by a correlation coefficient. This coefficient can have values between -1 and 1.

A coefficient of -1 indicates that two parameters are negatively correlated. When the first parameter is sampled at the high end of the min-max range, the other should be sampled at the low end.

A coefficient of 1 indicates that two parameters are positively correlated. When the first parameter is sampled at the high end of the min-max range, the other should be also sampled at the high end.

A coefficient of 0 says that there is no correlation between the two parameters. The sampling of one input will be independent of the other sampling then.

As example you can see in Figure 21 that Sub Task I and Sub Task II are positively correlated (@Risk->Model->Model Definition->Distributions: Select tasks, right mouse click->Correlate Distributions).

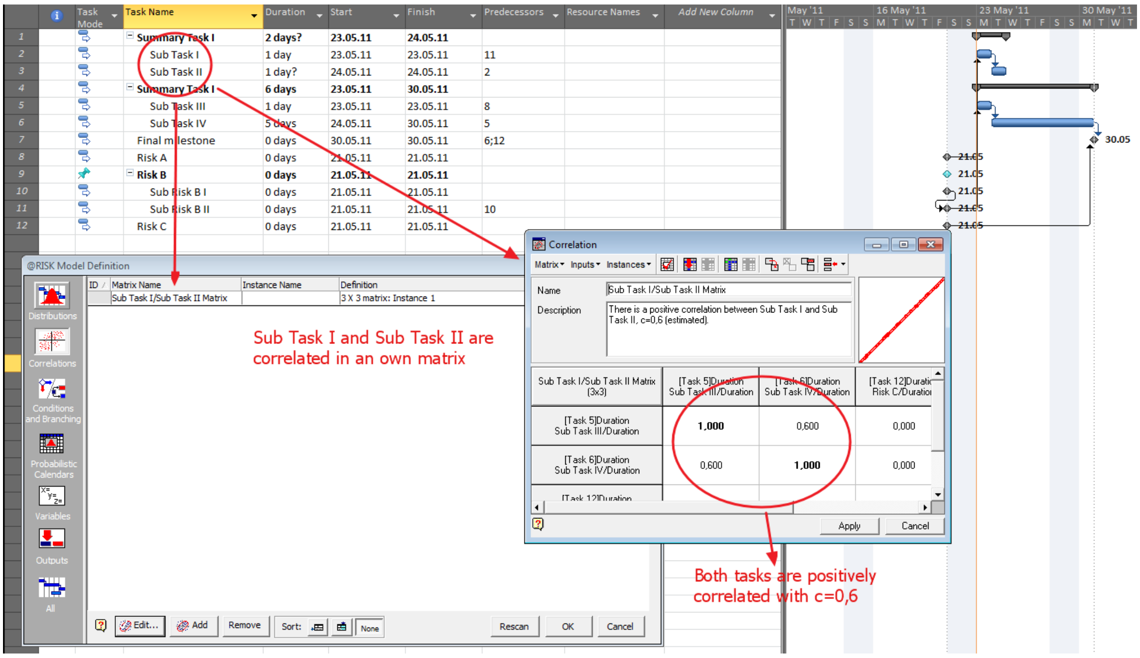


Figure 21: Correlation of Sub Task I and Sub Task II.

All in all, the impact of correlation tends to increase the overall uncertainty in a project. Therefore it is prudent to estimate correlation coefficients in parallel to task durations or other task parameters.

Define Parameters and Run a Simulation

The basic sequence of actions to run a simple simulation is as follows (cp. Figure 23):
1) Enter all input distributions.
@Risk->Model->Model Definition->Define Distribution
2) Add the output(s), like the final milestone.
@Risk->Model->Model Definition->Add Output
3) Enter the number of iterations (as example 1000) to take during the simulation.
@Risk->Simulation->Settings
4) Run the simulation.
@Risk->Simulation->Start

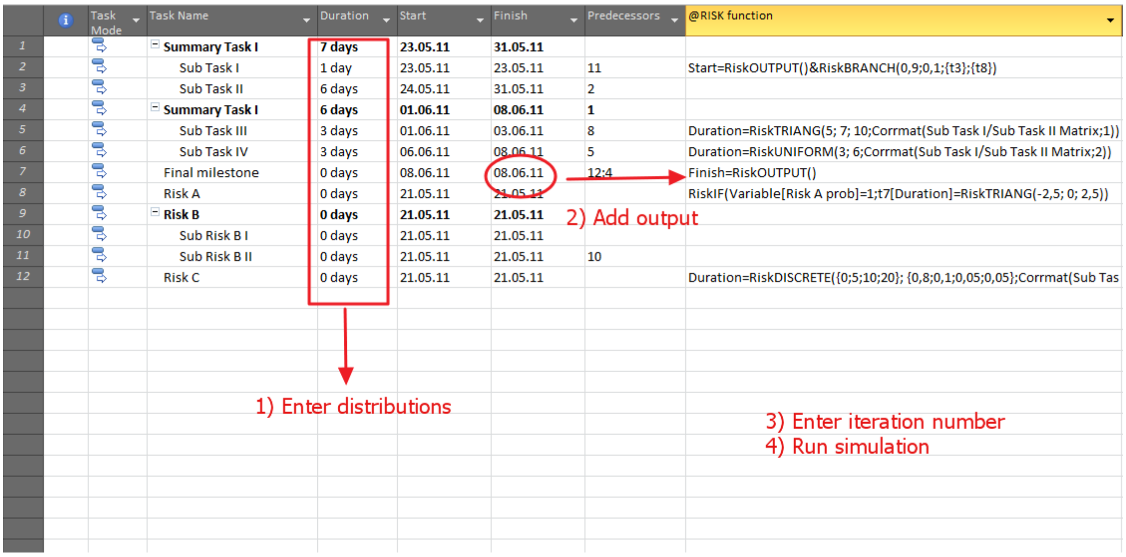


Figure 22: Basic sequence of simulation actions.

Interpretation of Results

Figure 23 obtained from @Risk shows a histogram of the simulation results for Project I according to the date of the final milestone. As you can see, 90% (i.e. 900 of 1000) of the simulation results ended with a finish date between 10.06.2011 and 01.07.2011. The mean outcome is 21.06.2011. There is only a less than 5% chance to meet the baseline date (deterministic date) of 8.6.2011 due to involved risks and task distributions.

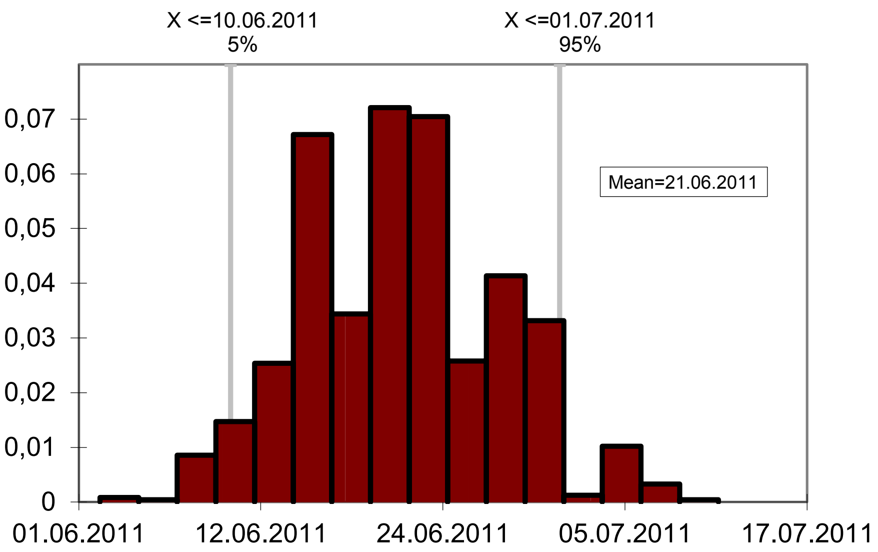


Figure 23: Probability density function (histogram): date of final milestone (Project I).

Next important information source is the cumulative curve (cp. Figure 24). Here, an 80th percentile of 1000 simulations were less than or equal to 27.6.2011, 19 days later than the original baseline date.

So to imply a contingency buffer, it would be a possibility not to quote below this 80th percentile officially and set the internal target at the 50th percentile.

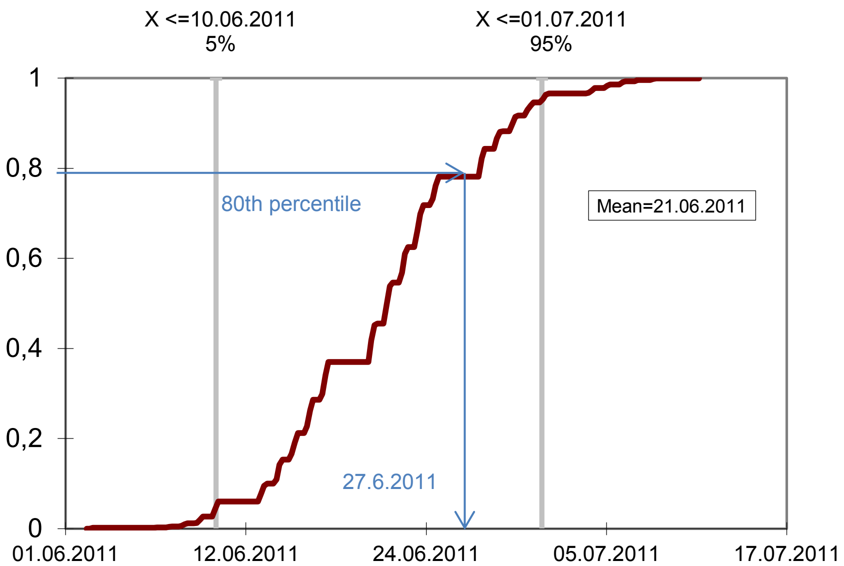


Figure 24: Cumulative probability function: date of final milestone (Project I).

If this contingency is unacceptable, one has to mitigate risks or re-schedule the plan. Here another diagram is very helpful, the tornado graph (→ *tornado graph*), a sensitivity analysis (cp. Figure 25).
In this graph you can see a positive correlation of four tasks to the final milestone date. As you can see, Risk A/duration impact according to Sub Task III and the duration impact of Sub Task II are important. So a schedule improvement contains more precise duration estimations, maybe a better probability parameter setting to these two tasks and a risk mitigation strategy for Risk A.

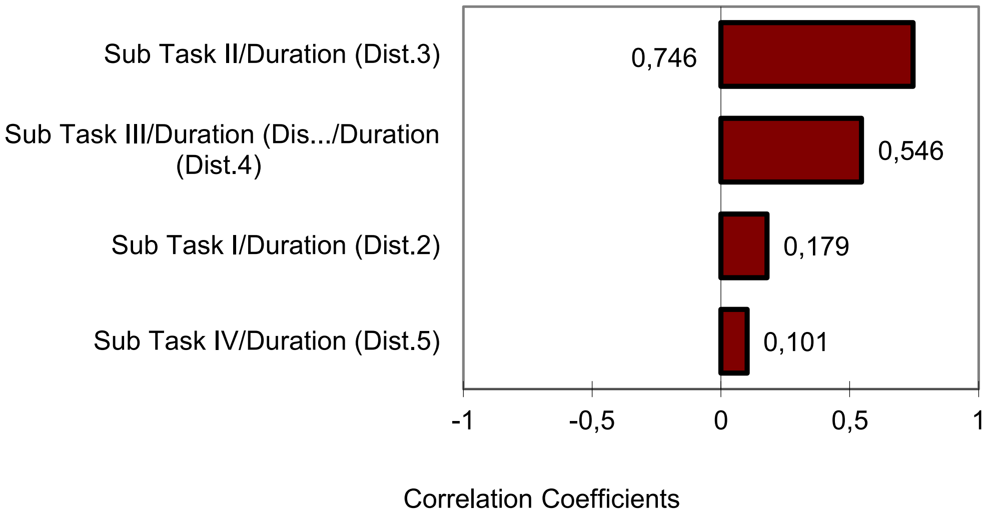


Figure 25: Tornado graph: correlation impact of tasks on final milestone date.

4.1.3 Model schedules

Three model schedules (OMV_A, OMV_B, OMV_C) that cover original OMV exploration & production projects were examined. These schedules are coded due to confidentiality and cannot be displayed in full detail.
The schedules comprise a lot of typical exploration & production characteristics and thus will provide a sufficient basis for comprehensive simulation results.
For testing CLT, correlations and constraints (→ *constraint*) dummy schedules without original OMV background were used due to easier handling.

4.2 Experimental Design

4.2.1 Simulation Settings[91]

Sampling

Sampling (→ *sample*) is used in @Risk simulations to generate random values (→ *random* value) from probability distribution functions. These sets of values are then used to evaluate your project schedule. So sampling is the basis for the hundreds or thousands of "what-if" scenarios of possible schedules. Choosing a sampling method affects both the quality of results, and the length of time necessary to simulate the schedule.
Sampling is done repetitively with one sample drawn every iteration from each input probability distribution. With enough iteration, the sampled values become distributed in a manner which approximates the original input probability distribution. The statistics of the

[91] Cp. N.N., Guide to Using @RISK (2010)

sampled distribution (mean, standard deviation and higher moments) approximate therefore the true input statistics. The graph of the sampled distribution looks like a graph of the true input distribution.

Another important factor to examine when evaluating sampling techniques is the number of iterations required to accurately recreate an input distribution through sampling because accurate results for output distributions depend on a complete sampling of input distributions. If one sampling method requires more iterations, and longer simulation runtimes than another to approximate input distributions, it is the less "efficient" method.

There are several techniques for drawing random samples. @Risk uses two of them called Monte Carlo Sampling and Latin Hypercube Sampling. They differ in the number of iterations required until sampled values approximate input distributions. Monte Carlo is the less exact one, it often requires a large number of samples to approximate an input distribution, especially if the input distribution is highly skewed or has some outcomes of low probability. Latin Hypercube Sampling, a newer sampling technique, forces the samples drawn to correspond more closely with the input distribution, and thus converges faster on the true statistics of the input distribution.

In the following section these methods are described more detailed.

Cumulative Distribution

It is helpful, when reviewing different sampling methods, to understand the concept of a cumulative distribution (→ *cumulative density distribution*) first.

Any probability distribution may be expressed in cumulative form. A cumulative curve is typically scaled from 0 to 1 on the Y-axis, with Y-axis values representing the cumulative probability up to the corresponding X-axis values. The 0 cumulative value is the minimum value (0% of the values will fall below this point) and the 1.0 cumulative value is the maximum value (100% of the values will fall below this point). The 0.5 cumulative value is the point of 50% cumulative probability. Fifty per cent of the values in the distribution fall below and 50% are above (see Figure 26).

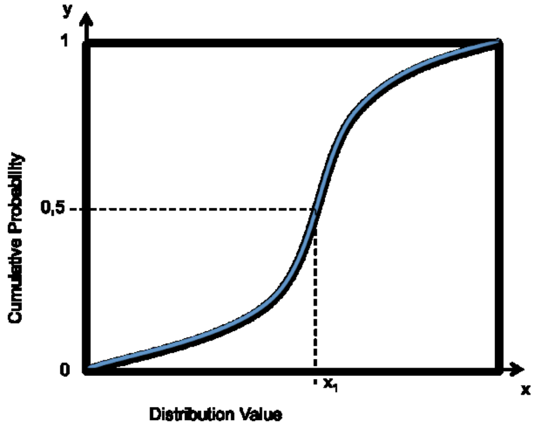


Figure 26: Example of a cumulative distribution.

Simultaneously the 0 to 1.0 scale of the curve is the range of the possible random numbers generated during sampling. In a typical Monte Carlo Sampling sequence, the computer will generate a random number between 0 and 1 with any number in the range equally likely to occur. This number is then used to select a value from the cumulative curve. In the figure above, the value sampled for the distribution shown would be x_1 if a random number of 0.5 was generated during sampling. Since the shape of the cumulative curve is based on the shape of the input probability distribution, it is more probable that more likely outcomes defined by the input distribution will be sampled. The more likely outcomes are in the range where the cumulative curve is the "steepest".

Monte Carlo Sampling

Monte Carlo sampling refers to the traditional technique for using numbers to sample from a probability distribution. It is applied to a wide variety of complex problems involving random behaviour. Monte Carlo sampling techniques are entirely random, so any given sample may fall anywhere within the range of the input distribution. Samples, of course, are more likely to be drawn in areas of the distribution which have higher probabilities of occurrence. In the cumulative distribution shown earlier, each Monte Carlo sample uses a new random number between 0 and 1. With enough iteration, Monte Carlo sampling “recreates” the input distributions through sampling. A problem of clustering, however, arises when a small number of iterations are performed.

In the illustration shown below (Figure 27), each of the 4 samples drawn falls in the middle of the distribution. The values in the outer ranges of the distribution are not represented in the samples, and thus their impact on the results is not included in your simulation output. Clustering becomes especially pronounced when a distribution includes low probability outcomes, which could have a major impact on your results. So if probability is low enough, a small number of Monte Carlo iterations may not sample sufficient quantities of these outcomes to accurately represent their probability. This problem has led to the development of stratified (→ *stratification*) sampling techniques such as the Latin Hypercube sampling.

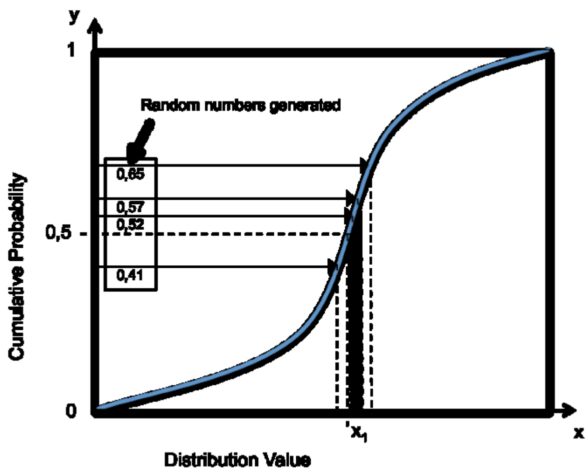


Figure 27: Monte Carlo Sampling.

Latin Hypercube Sampling

Latin Hypercube sampling is a recent development in sampling technology, designed to accurately recreate the input distribution through sampling in fewer iterations compared to the Monte Carlo method. The key to Latin Hypercube sampling is stratification of the input probability distributions. Stratification divides the cumulative curve into equal intervals on the cumulative probability scale (0 to 1). A sample is then randomly taken from each interval or “stratification” of the input distribution. Sampling is forced to represent values in each interval and thus recreate the input probability distribution.

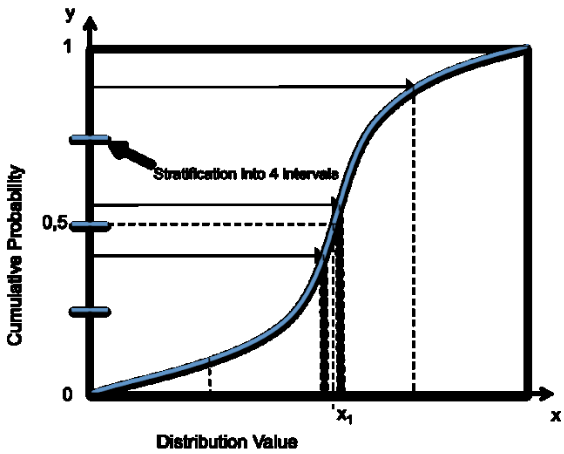


Figure 28: Latin Hypercube sampling.

In the illustration above (Figure 28), the cumulative curve has been divided into 5 intervals. During sampling, a sample is drawn from each interval (compare this to the 5 clustered samples drawn using the Monte Carlo method). With Latin Hypercube, the samples more accurately reflect the distribution of values in the input probability distribution. However, once a sample is taken from stratification, this stratification is not sampled from again.
As a more efficient sampling method, Latin Hypercube offers great benefits in terms of increased sampling efficiency and faster runtimes (due to less iteration).
In literature typically 1/3 as many Latin Hypercube iterations are required to get equal or better results as the equivalent amount of Monte Carlo iterations. Of course, the number of iterations required for good stable results depends on the nature of the model being analysed.

Testing Convergence

The concept of convergence (→ *convergence*) is used to test a sampling method. At the point of convergence, the output distributions are stable (additional iterations do not markedly change the shape or statistics of the sampled distribution). The sample mean versus the true mean is typically a measure of convergence. @Risk provides a solution for testing convergence with the so-called convergence monitor. Simply run a simulation and the built-in convergence monitoring capability in @Risk can estimate how many iterations it takes to stabilize the percentiles, mean and standard deviation.

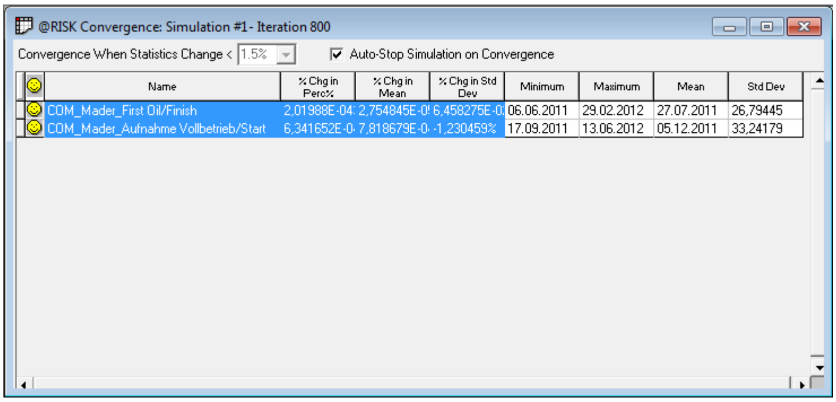


Figure 29: Convergence monitor for project OMV_A.[92]

92 Source: @RISK for Project

As you can see in Figure 29, the simulation result for project OMV_A is on convergence (convergence level < 1,5%) after 800 iterations. Therefore 1000 iterations should be a sufficient trade-off between effort and accuracy in this case.
Project OMV_B and OMV_C were examined the same way and in all cases 1000 iterations were effectual.

4.2.2 Design of Experiments (DoE)

Before starting an experiment it is necessary to create an experimental design to plan, fulfil and control experimental process.
In the first place the following parameters for schedule simulation were selected in an expert meeting:

- distribution shape,
- distribution central value and spread values,
- number of tasks (testing CLT),
- correlations,
- and constraints.

As testing basis three original OMV schedules with exploration & production background:

- OMV_A,
- OMV_B,
- OMV_C,

and three dummy schedules without OMV background:

- CLT (CLT testing),
- COR (correlation testing),
- and CON (constraints testing)

were checked.

Consecutively every parameter got its own DoE to examine the impact systematically. In the following section this process is described more deeply.

<u>Distribution Shape</u>

All test schedules had a various amount of different distribution shapes built in. When testing the impact of a shape change, it was necessary to find a test setting that produces comparable results.
Consecutively the general approach was to change an original distribution with a similar distribution because in praxis the problem is to find the right distribution shape amongst similar shapes. The following criteria defined a similar distribution:

- Similar shape
- Same boundary conditions
- Same mean value

For example a Lognormal function is replaced by a Weibull function due to these criteria. Additionally varying change percentage, you get a DoE as follows in Table 16.

Table 16: DoE “distribution shape” for OMV_A, OMV_B.

Schedule ID	**OMV_A** **distribution change (%)**	**OMV_B** **distribution change (%)**
OMV_X (Original)	0	0
OMV_X1	50	50
OMV_X2	100	100

OMV_C got a different DoE (cp. Table 17), because it uses Vary functions. The Vary function allows you to assign minimum and maximum values for a field using per cent or +/- changes of a central value like the mode. It also describes how a field's value will be distributed across a minimum-maximum range: you can choose a Triangular, Uniform, PERT or TRI1090 distribution (a Triangular distribution where the min value is at the 10th percentile of the distribution and the max value is at the 90th percentile of the distribution). For example: if the selected distribution type takes three arguments, the entered min value is the minimum argument of the distribution, the existing value for the task field is the most likely value, and the entered max value is the maximum argument of the distribution. During simulation @Risk converts the entered Vary function into a standard Triangular, Uniform or PERT distribution for sampling.[93] Consecutively it was possible to vary distributions in OMV_C quickly and to examine the impact of triangular (Triangular), curvy (PERT) and rectangular (Uniform) distributions.

Table 17: DoE "distribution shape" for OMV_C.

Schedule ID	**OMV_C distribution content (%)**
OMV_X (Original)	100, Triangular
OMV_X1	100, PERT
OMV_X2	100, Uniform

<u>Distribution Central Value and Spread Values</u>

Beside shape the central values like mean, mode or median and spread values like minimum and maximum or standard deviation are often estimated parameters for input distributions. Hence, the effect of modifying these values' percentage was examined with the following DoE (cp. Table 18):

Table 18: DoE "mean and spread".

Schedule ID	**Δ varied = constant**	**OMV_A value change (%)**			**OMV_B value change (%)**			**OMV_C value change (%)**		
OMV_X (Original)	= mean = spread	0			0			0		
OMV_X3+X%	Δ mean = spread	+10	+50	+100	+10	+50	+100	+10	+50	+100
OMV_X4+X%	= mean Δ spread	+10	+50	+100	+10	+50	+100	+10	+50	+100
OMV_X5+X%	Δ mean Δ spread	+10	+50	+100	+10	+50	+100	+10	+50	+100

<u>CLT: Number of Tasks</u>

The Central Limit Theorem (CLT) is an important parameter in literature and therefore it was examined in this work. As basis a dummy schedule was produced due to increase the number of tasks easily and to insert an adequate distribution mix.

This dummy schedule starts with 10 tasks and goes up to 1000 tasks. Simultaneously the project duration remains the same (1000 days), because otherwise the results are not comparable. As distribution mix 30 % PERT, 30 % Triangular and 30 % Uniform were implemented with a standard parameter set (@Risk default setting). 10 % of the tasks remained deterministic.

shows the according DoE:

Table 19 shows the according DoE:

93 Cp. @RISK for Project Help (2005)

Table 19: DoE “CLT”.

Schedule ID	# Tasks	Task durations (days)	Project duration (days)
CLT10	10	100	1000
CLT25	25	40	1000
CLT50	50	20	1000
CLT100	100	10	1000
CLT500	500	2	1000
CLT1000	1000	1	1000

Constraints

As already mentioned each task has a certain rule applied that helps the scheduling software to figure out when the task should start or finish. These rules are called constraints. Also there are three types of constraints: flexible, semi-flexible and inflexible ones. Flexible constraints are default setting, so semi-flexible and inflexible ones are typically user-driven and therefore their impact on schedule uncertainty was examined here.

Two semi-flexible constraints: Finish No Earlier Than (FNET) and Finish No Later Than (FNLT) and one inflexible constraint: Must Finish On (MFO) were implemented in the dummy schedule. This dummy schedule was already used in CLT simulations (cp. dummy schedule “CLT10”). The following DoE (Table 20) wraps up the experimental setting:

Table 20: DoE “Constraints”.

Schedule ID	Constraint type	Constraint content (%)
CON_0% (Control)	/	0
CON_MFO+20%	Must Finish On	20
CON_MFO+40%	Must Finish On	40
CON_FNLT+20%	Finish No Later Than	20
CON_FNLT+40%	Finish No Later Than	40
CON_FNET+20%	Finish No Earlier Than	20
CON_FNET+40%	Finish No Earlier Than	40

Correlations

Many times your uncertain inputs are related. For example one input has a low value, some one other should also has. This is called correlation. Examination of different correlation inputs were conducted with the DoE`s in

Table 22 to Table 24. 3 genuine OMV schedules were used to simulate real-world conditions.

Correlation coefficients were set randomly and chosen regarding a “life-like” mix (see Table 21).

Table 21: Correlation mix.

Correlated Tasks	Correlation Coefficient
5 %	-1
15 %	-0,8
15 %	-0,5
15 %	-0,3
12 %	0
15 %	0,3
15 %	0,5
15 %	0,8
5 %	1

Table 22: DoE "random correlation" for OMV_A.

Schedule ID	Correlated Task #	Correlation Coefficients
OMV_A	/	/
COR_OMV_A	All distributed inputs (21)	Mix

Table 23: DoE "random correlation" for OMV_B.

Schedule ID	Correlated Task #	Correlation Coefficients
OMV_B	/	/
COR_OMV_B	All distributed inputs (27)	Mix

Table 24: DoE "random correlation" for OMV_C.

Schedule ID	Correlated Task #	Correlation Coefficients
OMV_C	/	/
COR_OMV_C	All distributed inputs (35)	Mix

4.3 Results and Conclusions

This chapter deals with the outcome of all @Risk simulations with various schedules. These results displayed by outcome distributions are shown in Figure 30 to Figure 86 and the according distribution data is summarized in
Table 25 to
Table 73.
The correlation results are presented in different form (Figure 89 to 92), because output distribution shapes change very little and moreover line charts are more effective in showing correlation impact.
In addition interpretations of the outcome are made in each case.

4.3.1 Distribution Shape

Firstly the impact of input distribution shapes on outcome distribution shape and parameters was investigated. DoE "Distribution shape" served as basis for the simulations.

<u>Shape: OMV_A</u>

As you can see in Figure 30, Figure 31 and Figure 32, there is a slight shift of the mean value to later dates. Kurtosis (→ *kurtosis*) remains basically the same, but skewness (→ *skewness*) shifts from left skewed towards symmetrical. Furthermore the 90 % spread (→ *range*, in this case it is called "Diff. X") increases heavily, this means, the output distribution is more uncertain or "riskier".
There are a lot of Lognormal distributions in OMV_A that are replaced by Weibull distributions. This explains the bigger spread and less skewness, because Weibull is not as slim as Lognormal (and therefore uncertainties with high probability are more scattered) and more symmetrical. Hence, the mean moves due to bigger outlier impact too.
All mentioned impacts increase with switching from 50 % to 100 % distribution change.

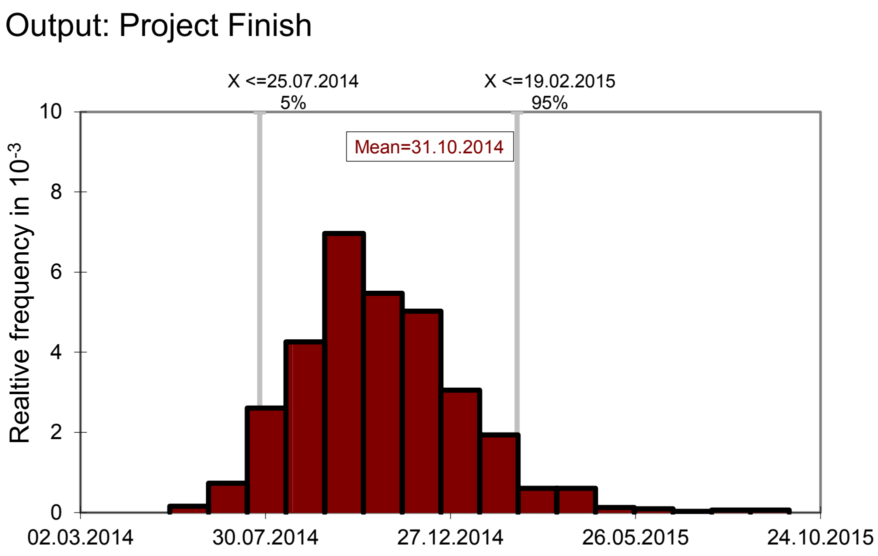


Figure 30: Output distribution for OMV_A.

Table 25: Distribution data OMV_A.

Minimum	13.05.2014	Skewness	0,76	Mode	07.07.2014
Mean	31.10.2014	Kurtosis	4,51	5th Perc.	25.07.2014
Maximum	28.09.2015			95th Perc.	19.02.2015
Std. Dev.	67,0			Diff. X	209 days

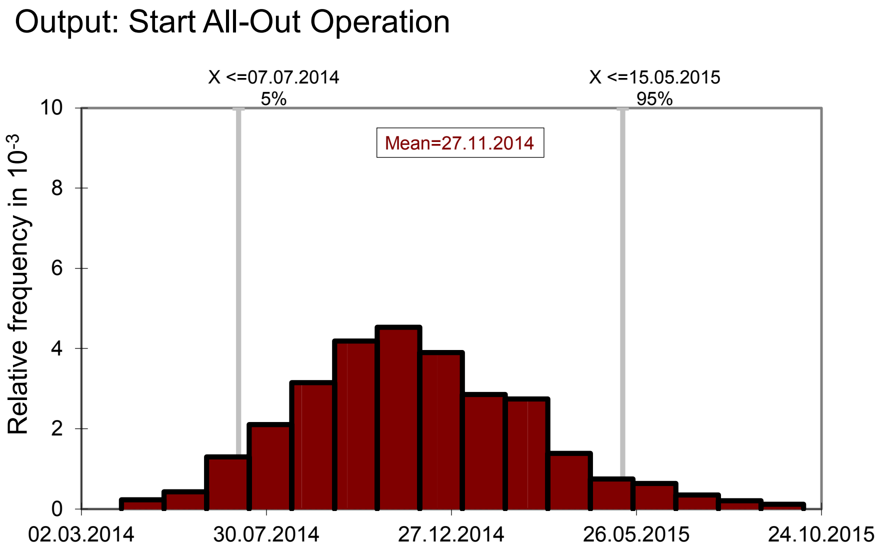


Figure 31: Output distribution for OMV_A1.

Table 26: Distribution data for OMV_A1.

Minimum	03.04.2014	Skewness	0,41	Mode	04.11.2014
Mean	27.11.2014	Kurtosis	3,14	5th Perc.	07.07.2014
Maximum	09.10.2015			95th Perc.	15.05.2015
Std. Dev.	94,1			Diff. X	312 days

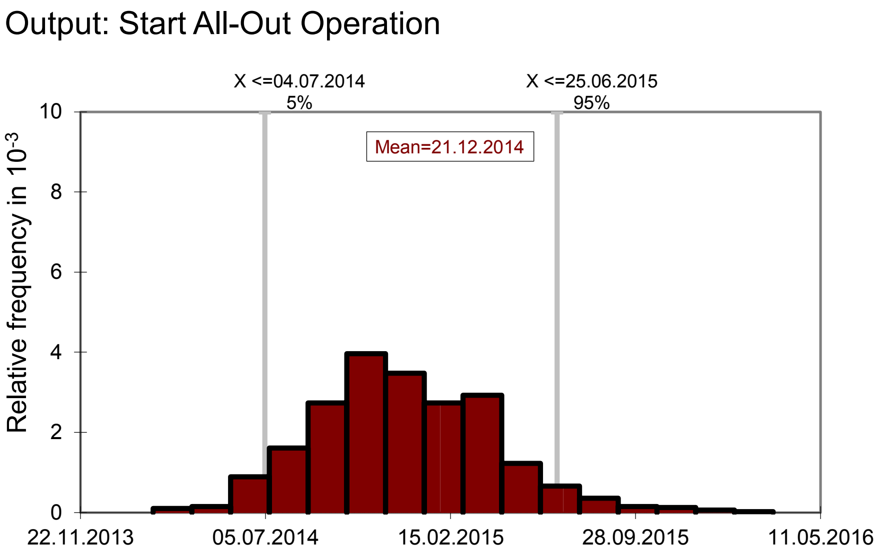


Figure 32: Output distribution for OMV_A2.

Table 27: Distribution data for OMV_A2.

Minimum	18.02.2014	Skewness	0,41	Mode	12.08.2014
Mean	21.12.2014	Kurtosis	3,47	5th Perc.	04.07.2014
Maximum	14.03.2016			95th Perc.	25.06.2015
Std. Dev.	109			Diff. X	356 days

Shape: OMV_B

Simulations with varied OMV_B affirmed the results obtained with OMV_A (see Figure 33, Figure 34 and Figure 35). Again a lot of Lognormal distributions are changed into Weibull and thus distribution dissolving takes place. Kurtosis is strongly reduced (9,6 to 2,8) and skewness tends to symmetrical values (1,3 to 0,3). Mean shifts only a little bit.
Again all mentioned impacts increase with switching from 50 % to 100 % distribution change.

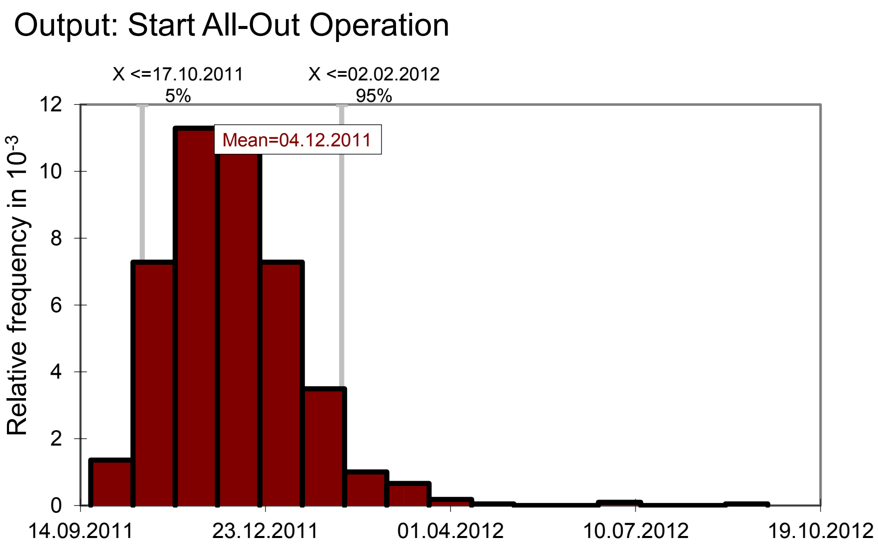


Figure 33: Output distribution for OMV_B.

Table 28: Distribution data for OMV_B.

Minimum	19.09.2011	Skewness	1,37	Mode	03.11.2011
Mean	04.12.2011	Kurtosis	9,06	5th Perc.	17.10.2011
Maximum	20.09.2012			95th Perc.	02.02.2012
Std. Dev.	35,9			Diff. X	108 days

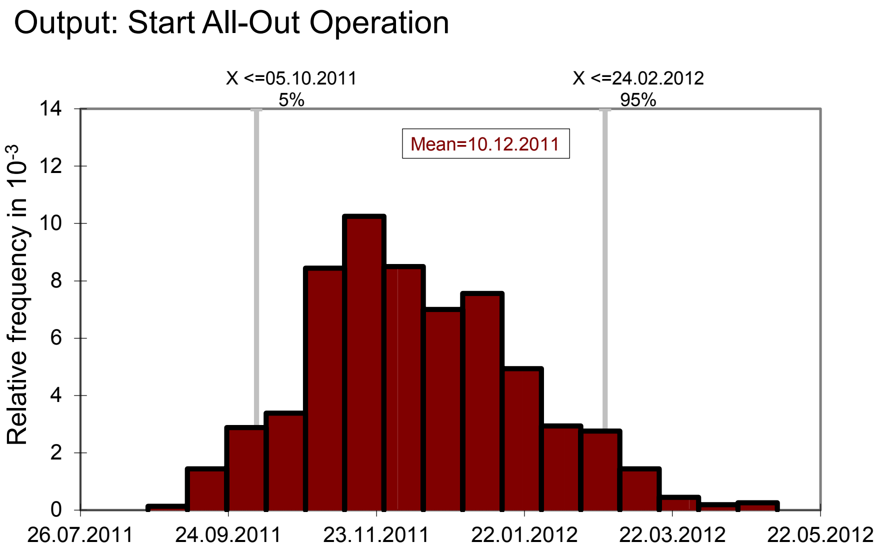


Figure 34: Output distribution for OMV_B1.

Table 29: Distribution data for OMV_B1.

Minimum	22.08.2011	Skewness	0,38	Mode	04.11.2011
Mean	10.12.2011	Kurtosis	2,89	5th Perc.	05.10.2011
Maximum	04.05.2012			95th Perc.	24.02.2012
Std. Dev.	42,8			Diff. X	142 days

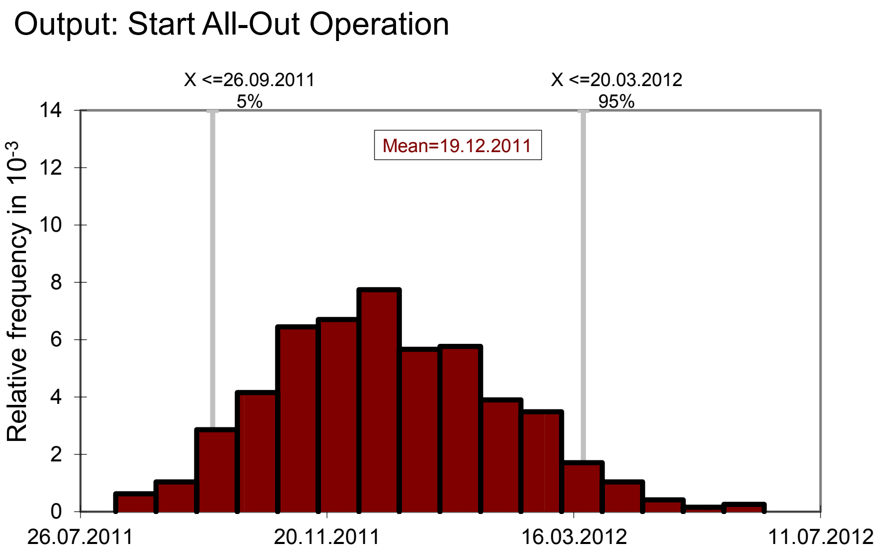


Figure 35: Output distribution for OMV_B2.

Table 30: Distribution data for OMV_B2.

Minimum	11.08.2011	Skewness	0,30	Mode	23.09.2011

Mean	19.12.2011	Kurtosis	2,83	5th Perc.	26.09.2011
Maximum	14.06.2012			95th Perc.	20.03.2012
Std. Dev.	53,9			Diff. X	176 days

Shape: OMV_C

The original C Schedule has only Vary distributions implied, so you can change distributions very easily to test distribution impact. So OMV_C is completely linked with Triangular distributions. These distributions were switched to PERT (OMV_C1) and Uniform (OMV_C2).

Triangular to PERT:

Most significant impact is the decrease of the 90% spread (see Figure 36 and Figure 37) after switching to PERT. However, with PERT values between the most likely and extremes are more likely to occur and the extremes are not as emphasized compared with Triangular. In practice, this means that we "trust" the most likely value. Thus, spread is smaller and the mean shifts to earlier date. Skewness and kurtosis are basically the same.

Triangular to Uniform:

In Figure 36 and Figure 38 means are nearly the same, but spread decreases one more time. That is not logically at first, because Uniform distributions are referred to as a "no knowledge" distribution. You have a base value but no clue, if the probability decreases moving away from that central value. All values have the same probability, even outliers. Therefore spread should increase and consecutively be bigger than OMV_C`s spread. One possible reason could be skewed Triangular distributions in OMV_C that over-emphasize values in the skew direction.

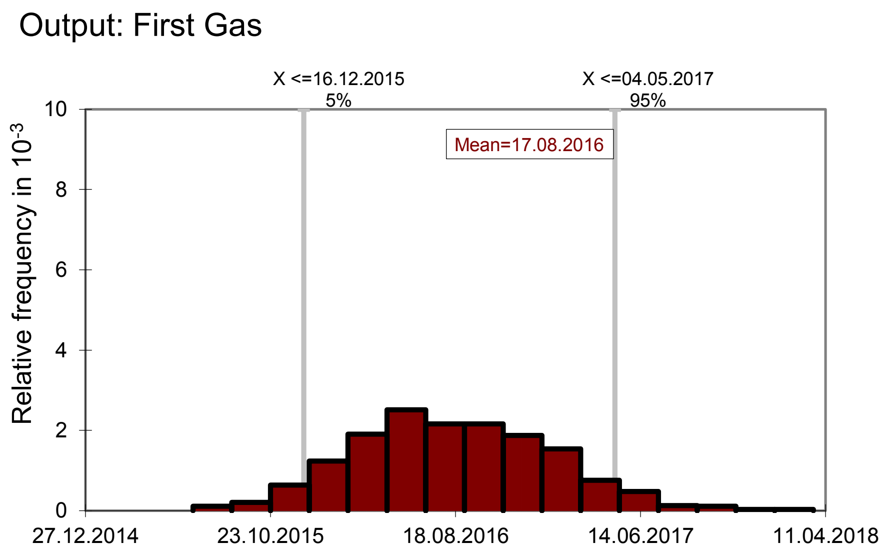


Figure 36: Output distribution for OMV_C.

Table 31: Distribution data for OMV_C.

Minimum	19.06.2015	Skewness	0,22	Mode	29.01.2016
Mean	17.08.2016	Kurtosis	2,99	5th Perc.	16.12.2015
Maximum	22.03.2018			95th Perc.	04.05.2017
Std. Dev.	159			Diff. X	505 days

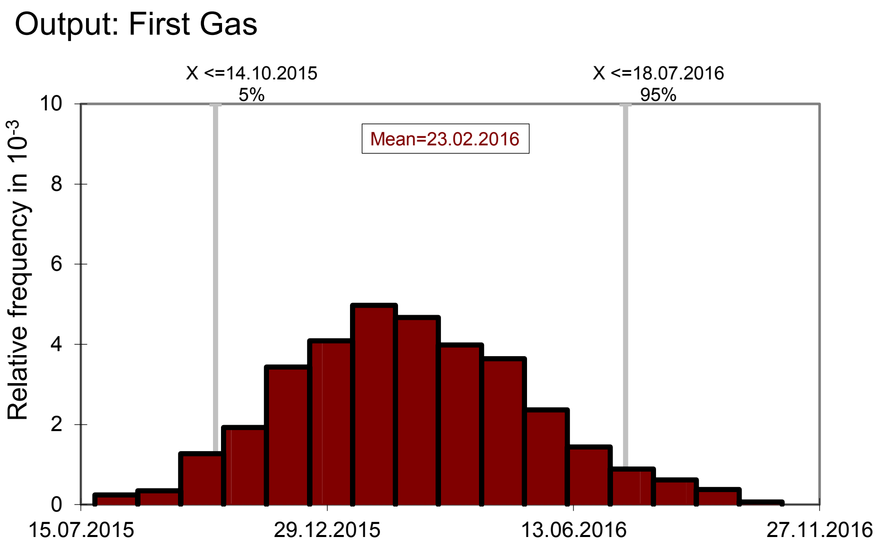


Figure 37: Output distribution for OMV_C1.

Table 32: Distribution data for OMV_C1.

Minimum	24.07.2015	Skewness	0,21	Mode	26.02.2016
Mean	23.02.2016	Kurtosis	2,87	5th Perc.	14.10.2015
Maximum	01.11.2016			95th Perc.	18.07.2016
Std. Dev.	81,6			Diff. X	278 days

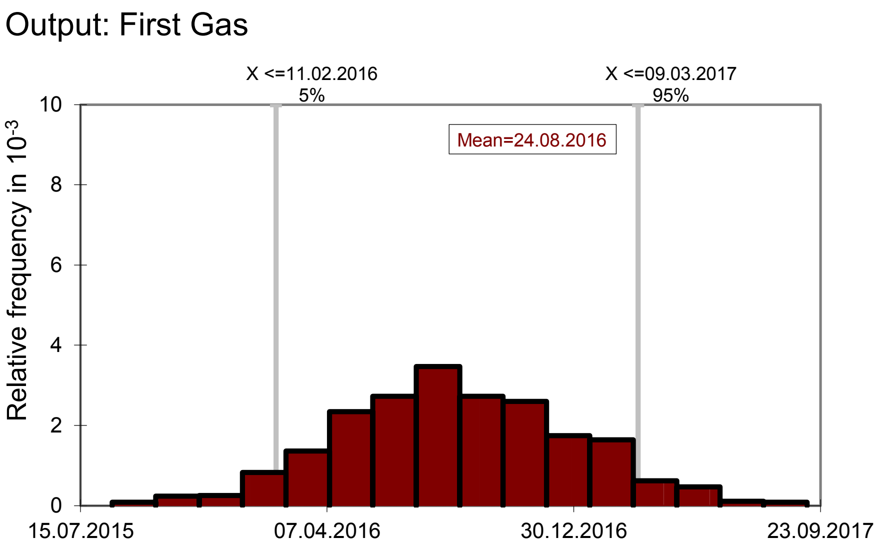


Figure 38: Output distribution for OMV_C2.

Table 33: Distribution data for OMV_C2.

Minimum	18.08.2015	Skewness	5,37E-02	Mode	16.03.2016
Mean	24.08.2016	Kurtosis	2,83	5th Perc.	11.02.2016
Maximum	08.09.2017			95th Perc.	09.03.2017
Std. Dev.	123			Diff. X	392 days

4.3.2 Mean and Spread

Next step was to examine the outcome effect of mean and spread variations when defining input distributions. DoE "Mean and spread" provided the simulation basis.

<u>Mean and spread: OMV_A</u>

Figure 39 to Figure 47 show the impact of varying mean and spread separately in first and then together. A summarize of the results is given by Figure 48 and Figure 49. If the inputs mean increases, the outputs mean increases too. Moreover the spread remains basically the same. If the inputs spread increases, the outputs mean increases too. Here, the mean does not differ much. If both mean and spread increase simultaneously, the outputs mean and spread increase consecutively.
All in all input mean and spread and output mean and spread are correlating positively and it seems that mean and spread are independent (note: without a correlation test there is no proof for the latter statement).
Moreover it seems like there is less impact on outcome shape when spread is increased separately (cp. OMV_A4+X%).

Output: Project Finish

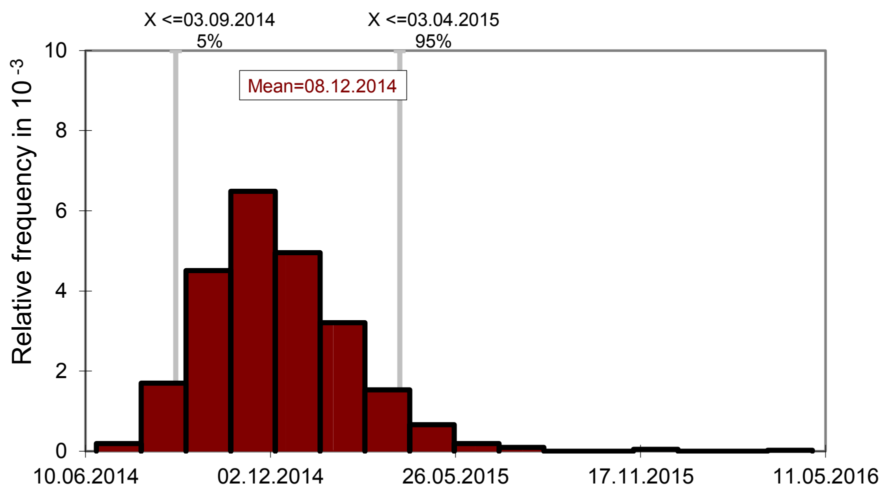


Figure 39: Output distribution for OMV_A3+10%.

Table 34: Distribution data for OMV_A3+10%.

Minimum	20.06.2014	Skewness	1,14	Mode	24.11.2014
Mean	08.12.2014	Kurtosis	7,25	5th Perc.	03.09.2014
Maximum	28.04.2016			95th Perc.	03.04.2015
Std Dev	68,5			Diff. X	212 days

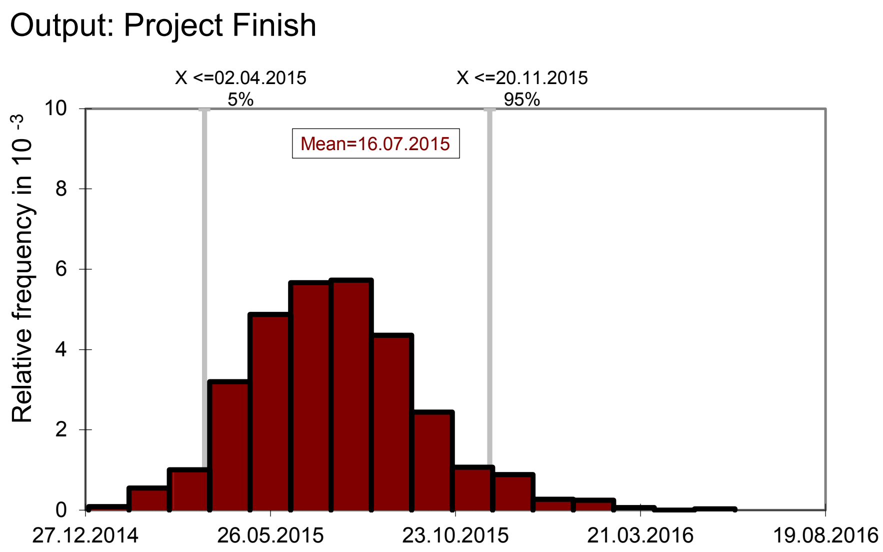


Figure 40: Output distribution for OMV_A3+50%.

Table 35: Distribution data for OMV_A3+50%.

Minimum	29.12.2014	Skewness	0,48	Mode	15.04.2015
Mean	16.07.2015	Kurtosis	3,70	5th Perc.	02.04.2015
Maximum	06.06.2016			95th Perc.	20.11.2015
Std Dev	69,8			Diff. X	232 days

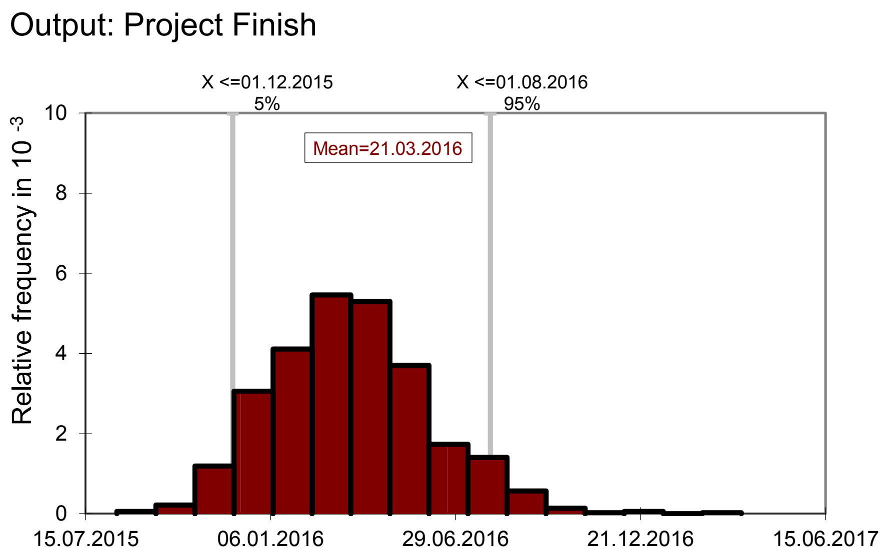


Figure 41: Output distribution for OMV_A3+100%

Table 36: Distribution data for OMV_A3+100%.

Minimum	13.08.2015	Skewness	0,44	Mode	27.10.2015
Mean	21.03.2016	Kurtosis	3,64	5th Perc.	01.12.2015
Maximum	27.03.2017			95th Perc.	01.08.2016
Std Dev	74,4			Diff. X	244 days

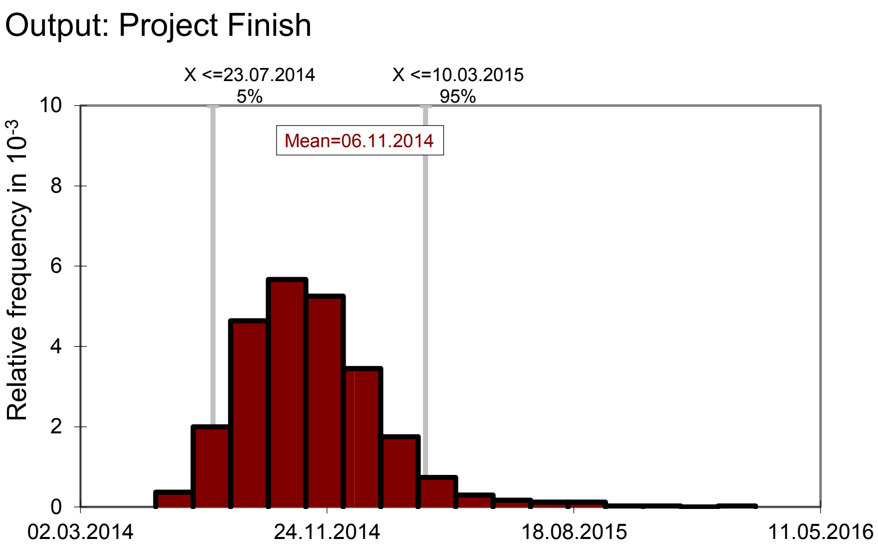


Figure 42: Output distribution for OMV_A4+10%.

Table 37: Distribution data for OMV_A4+10%.

Minimum	22.05.2014	Skewness	1,15	Mode	27.08.2014
Mean	06.11.2014	Kurtosis	6,27	5th Perc.	23.07.2014
Maximum	01.03.2016			95th Perc.	10.03.2015
Std Dev	76,6			Diff. X	230 days

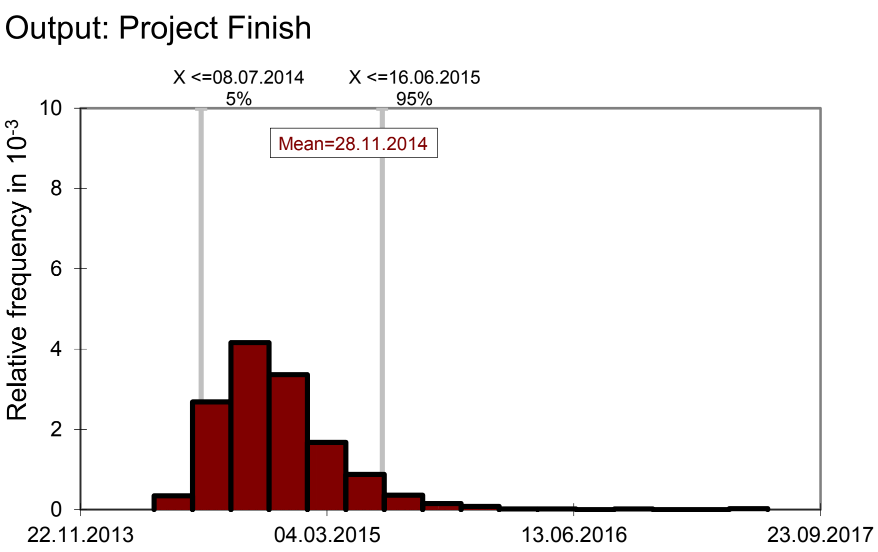


Figure 43: Output distribution for OMV_A4+50%.

Table 38: Distribution data for OMV_A4+50%.

Minimum	10.04.2014	Skewness	1,93	Mode	14.07.2014
Mean	28.11.2014	Kurtosis	12,2	5th Perc.	08.07.2014
Maximum	15.06.2017			95th Perc.	16.06.2015
Std Dev	115			Diff. X	343 days

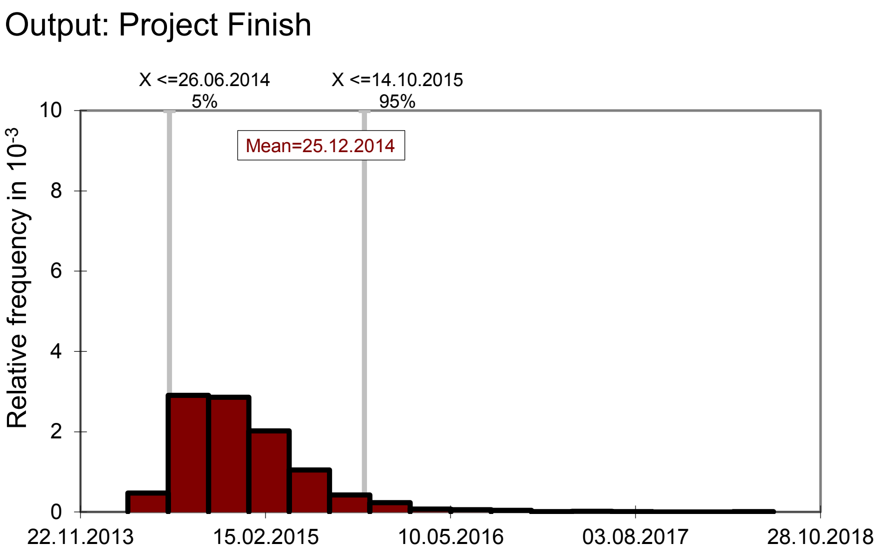


Figure 44: Output distribution for OMV_A4+100%.

Table 39: Distribution data for OMV_A4+100%.

Minimum	17.03.2014	Skewness	2,00	Mode	24.07.2014
Mean	25.12.2014	Kurtosis	10,7	5th Perc.	26.06.2014
Maximum	05.07.2018			95th Perc.	14.10.2015
Std Dev	162			Diff. X	475 days

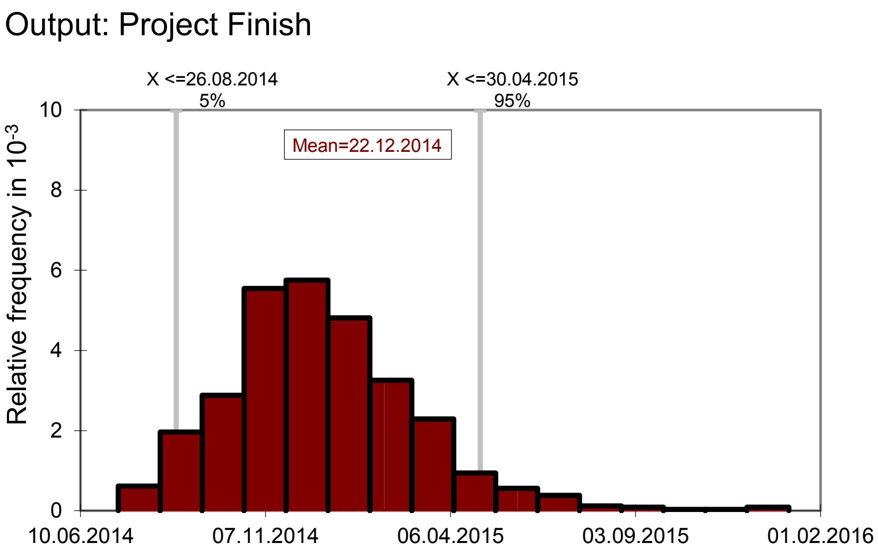


Figure 45: Output distribution for OMV_A5+10%.

Table 40: Distribution data for OMV_A5+10%.

Minimum	10.07.2014	Skewness	0,81	Mode	28.11.2014
Mean	22.12.2014	Kurtosis	4,79	5th Perc.	26.08.2014
Maximum	06.01.2016			95th Perc.	30.04.2015
Std Dev	76,6			Diff. X	247 days

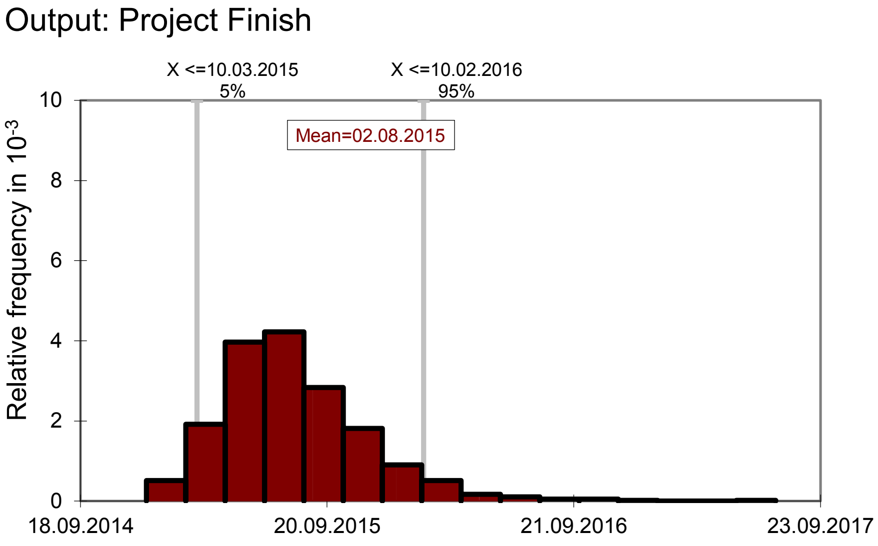


Figure 46: Output distribution for OMV_A5+50%.

Table 41: Distribution data for OMV_A5+50%.

Minimum	25.12.2014	Skewness	1,21	Mode	28.04.2015
Mean	02.08.2015	Kurtosis	6,39	5th Perc.	10.03.2015
Maximum	18.07.2017			95th Perc.	10.02.2016
Std Dev	108			Diff. X	337 days

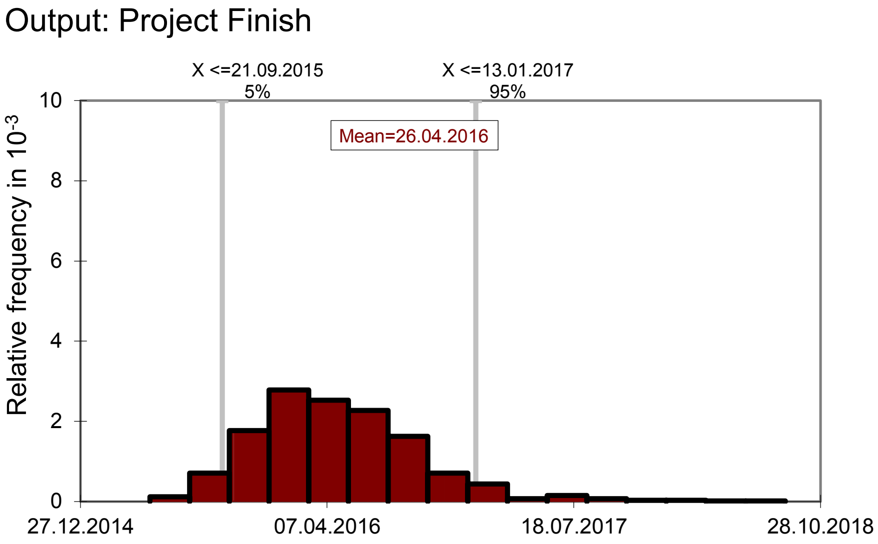


Figure 47: Output distribution for OMV_A5+100%.

Table 42: Distribution data for OMV_A5+100%.

Minimum	07.05.2015	Skewness	0,95	Mode	16.09.2015
Mean	26.04.2016	Kurtosis	5,10	5th Perc.	21.09.2015
Maximum	22.08.2018			95th Perc.	13.01.2017
Std Dev	154			Diff. X	480 days

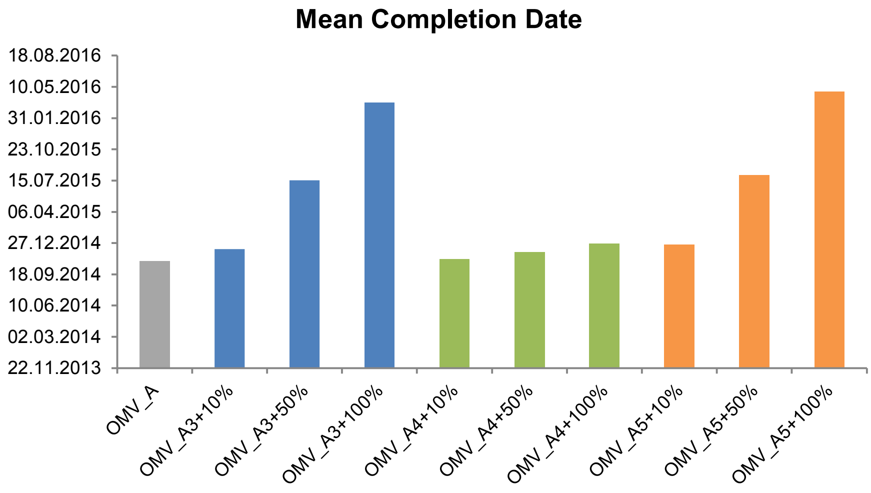


Figure 48: Mean distribution of OMV_A variations.

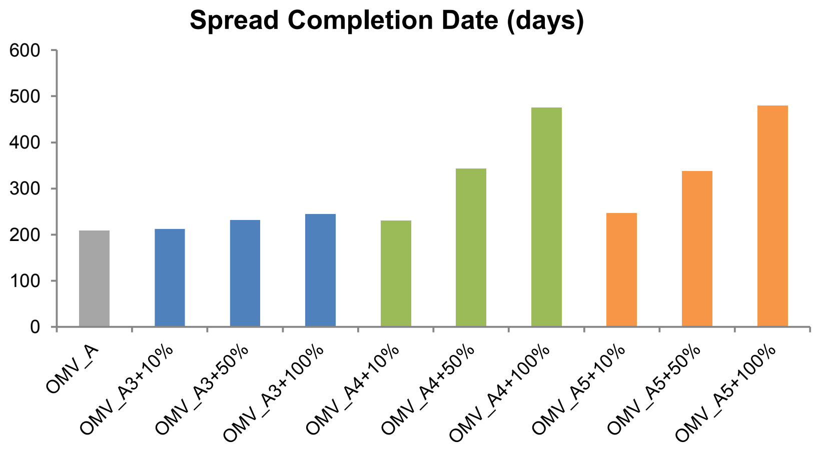


Figure 49: Spread distribution of OMV_A variations.

Mean and spread: OMV_B

Simulations with OMV_B and its derivatives proofed the results and conclusion from simulating OMV_A derivatives. Outcome means and spreads correlate positively with input means and spreads (cp. Figure 59 and Figure 60). A specific fact occurred with OMV_A5+50%. As you can see in Figure 59, this test schedule with 50 % input mean and spread escalation led to almost the same outcome mean as 100 % input mean and spread escalation did. Therefore it is possible that a synchronized mean and spread increase started some synergy effects. Small effects on output shape exist, because distributions are more symmetrical (lower positive skewness) and are less slim (lower positive kurtosis).

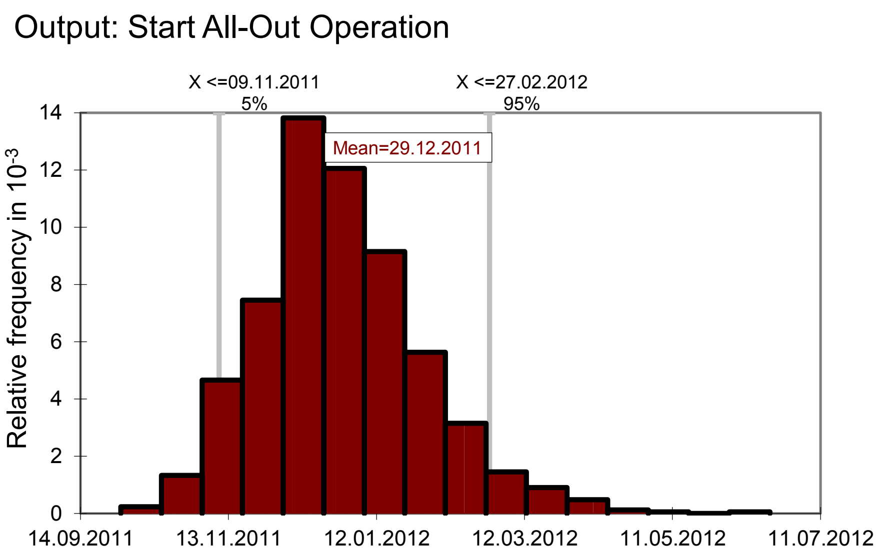


Figure 50: Output distribution for OMV_B3+10%.

Table 43: Distribution data for OMV_B3+10%.

Minimum	30.09.2011	Skewness	0,65	Mode	02.12.2011
Mean	29.12.2011	Kurtosis	4,05	5th Perc.	09.11.2011
Maximum	20.06.2012			95th Perc.	27.02.2012
Std Dev	33,6			Diff. X	110 days

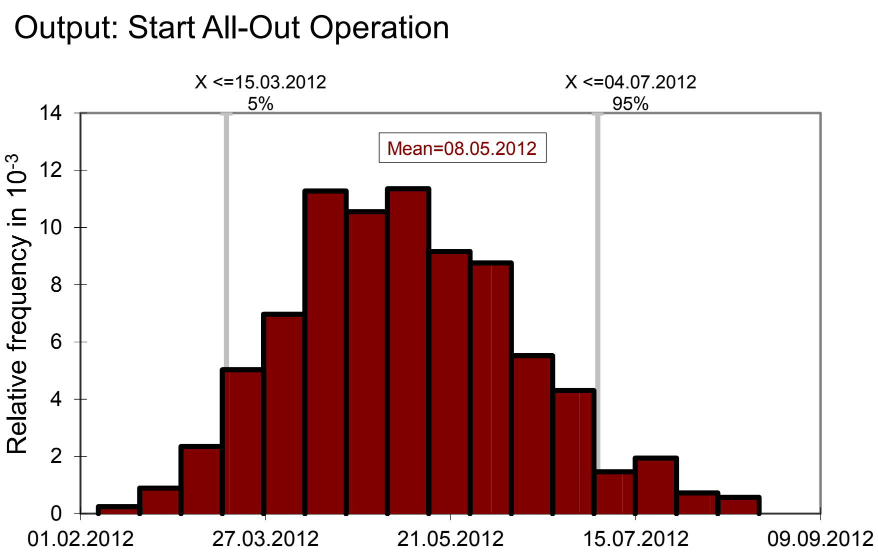


Figure 51: Output distribution for OMV_B3+50%.

Table 44: Distribution data for OMV_B3+50%.

Minimum	06.02.2012	Skewness	0,30	Mode	11.04.2012
Mean	08.05.2012	Kurtosis	2,85	5th Perc.	15.03.2012
Maximum	21.08.2012			95th Perc.	04.07.2012
Std Dev	34,6			Diff. X	111 days

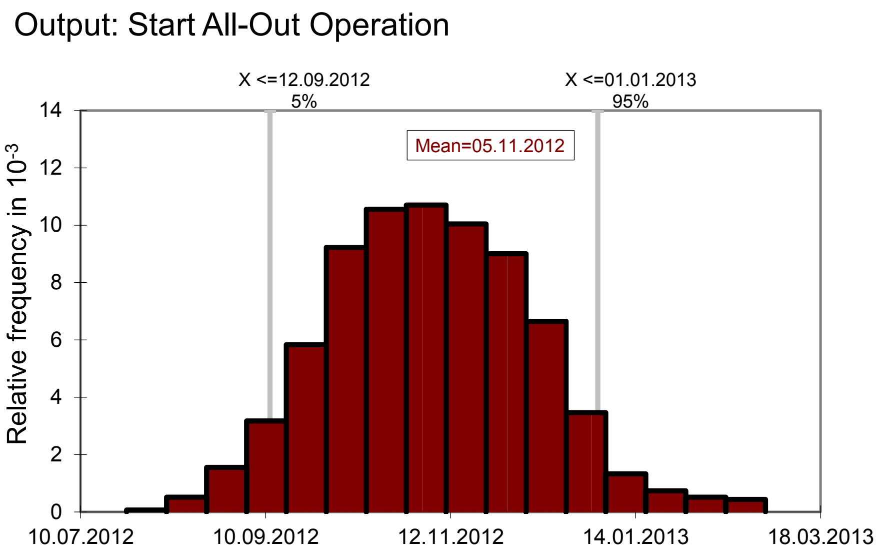


Figure 52: Output distribution for OMV_B3+100%.

Table 45: Distribution data for OMV_B3+100%.

Minimum	25.07.2012	Skewness	0,19	Mode	04.10.2012
Mean	05.11.2012	Kurtosis	2,95	5th Perc.	12.09.2012
Maximum	27.02.2013			95th Perc.	01.01.2013
Std Dev	35,0			Diff. X	111 days

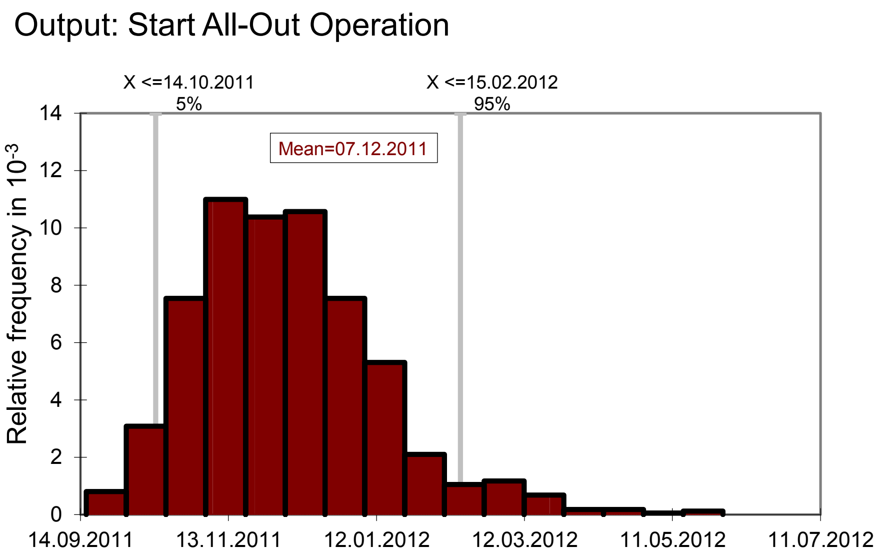


Figure 53: Output distribution for OMV_B4+10%.

Table 46: Distribution data for OMV_B4+10%.

Minimum	16.09.2011	Skewness	0,86	Mode	07.11.2011
Mean	07.12.2011	Kurtosis	4,43	5th Perc.	14.10.2011
Maximum	01.06.2012			95th Perc.	15.02.2012
Std Dev	37,3			Diff. X	124 days

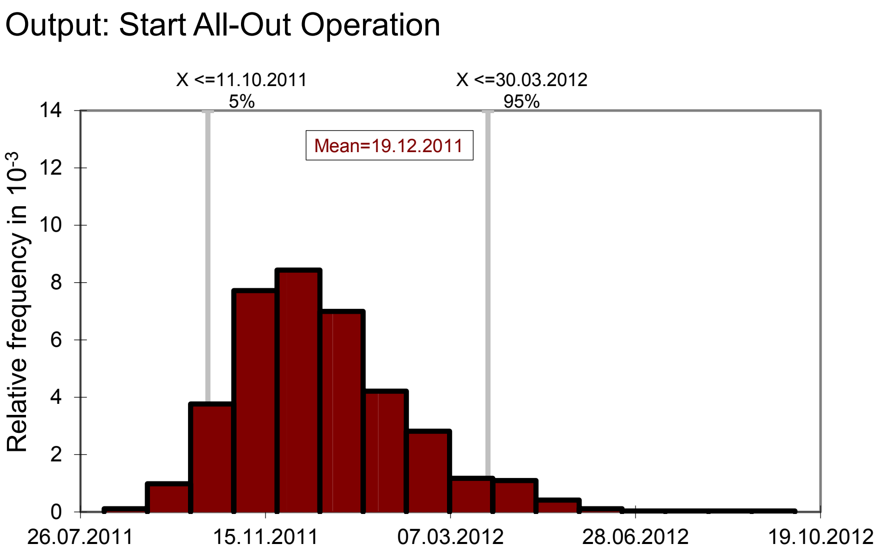


Figure 54: Output distribution for OMV_B4+50%.

Table 47: Distribution data for OMV_B4+50%.

Minimum	09.08.2011	Skewness	0,98	Mode	22.12.2011
Mean	19.12.2011	Kurtosis	4,83	5th Perc.	11.10.2011
Maximum	03.10.2012			95th Perc.	30.03.2012
Std Dev	52,1			Diff. X	171 days

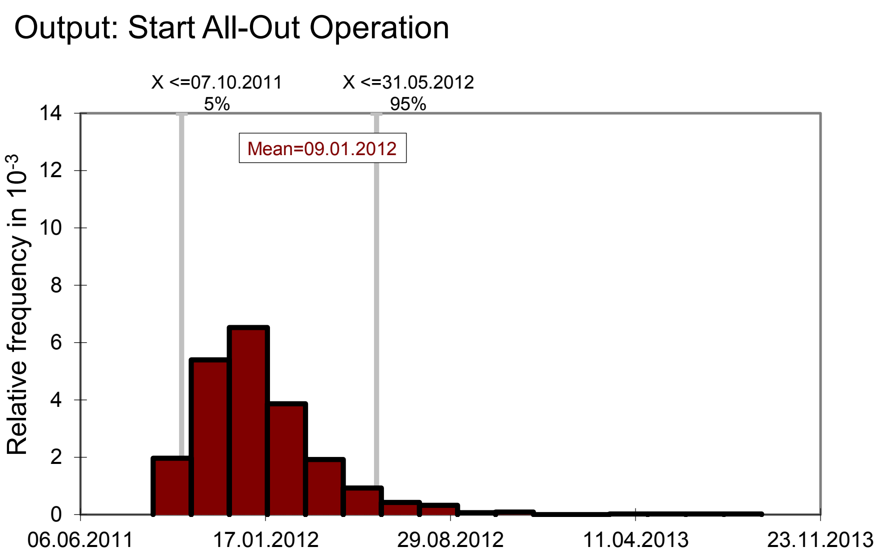


Figure 55: Output distribution for OMV_B4+100%.

Table 48: Distribution data for OMV_B4+100%.

Minimum	02.09.2011	Skewness	2,02	Mode	09.11.2011
Mean	09.01.2012	Kurtosis	11,6	5th Perc.	07.10.2011
Maximum	12.09.2013			95th Perc.	31.05.2012
Std Dev	79,4			Diff. X	237 days

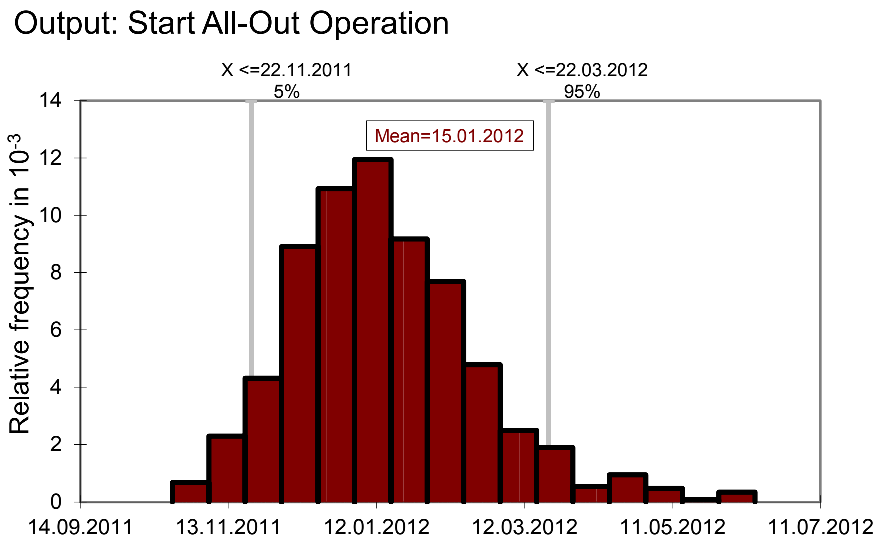


Figure 56: Output distribution for OMV_B5+10%.

Table 49: Distribution data for OMV_B5+10%.

Minimum	21.10.2011	Skewness	0,74	Mode	21.12.2011
Mean	15.01.2012	Kurtosis	3,98	5th Perc.	22.11.2011
Maximum	14.06.2012			95th Perc.	22.03.2012
Std Dev	37,8			Diff. X	121 days

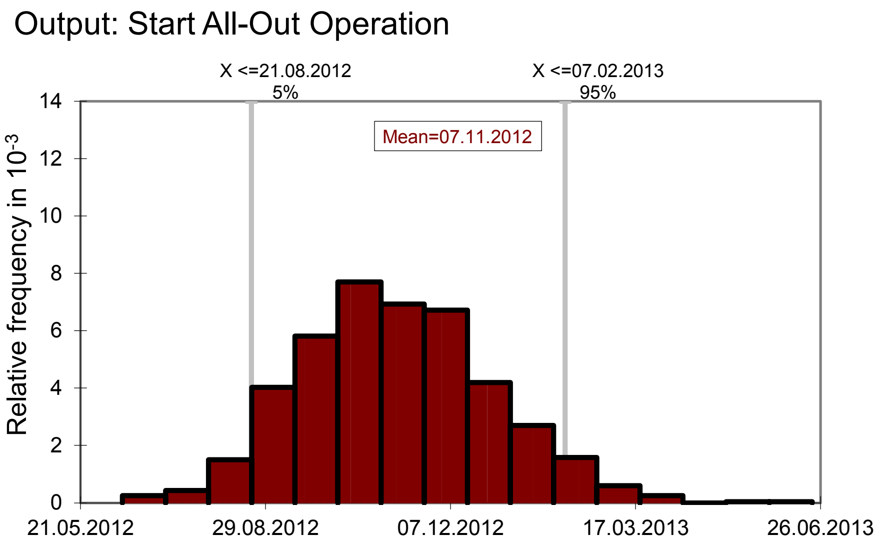


Figure 57: Output distribution for OMV_B5+50%.

Table 50: Distribution data for OMV_B5+50%.

Minimum	12.06.2012	Skewness	0,29	Mode	23.10.2012
Mean	07.11.2012	Kurtosis	3,17	5th Perc.	21.08.2012
Maximum	21.06.2013			95th Perc.	07.02.2013
Std Dev	51,9			Diff. X	170 days

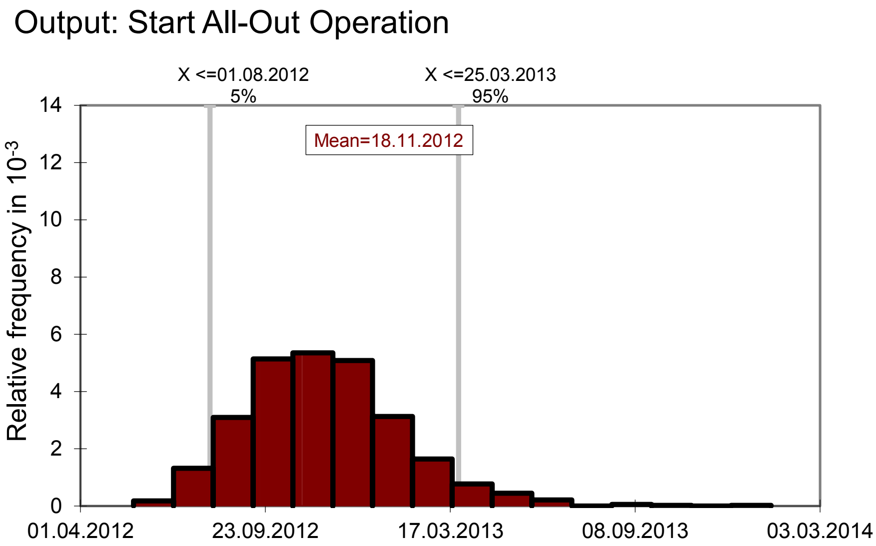


Figure 58: Output distribution for OMV_B5+100%.

Table 51: Distribution data for OMV_B5+100%.

Minimum	21.05.2012	Skewness	0,77	Mode	05.09.2012
Mean	18.11.2012	Kurtosis	4,74	5th Perc.	01.08.2012
Maximum	15.01.2014			95th Perc.	25.03.2013
Std Dev	73,2			Diff. X	236 days

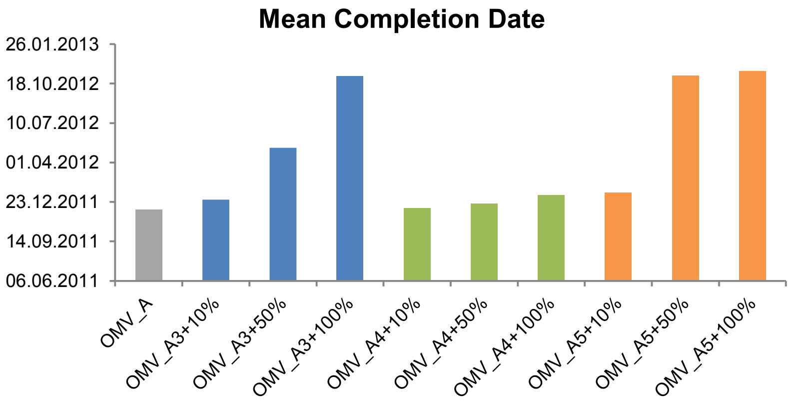


Figure 59: Mean distribution of OMV_B variations.

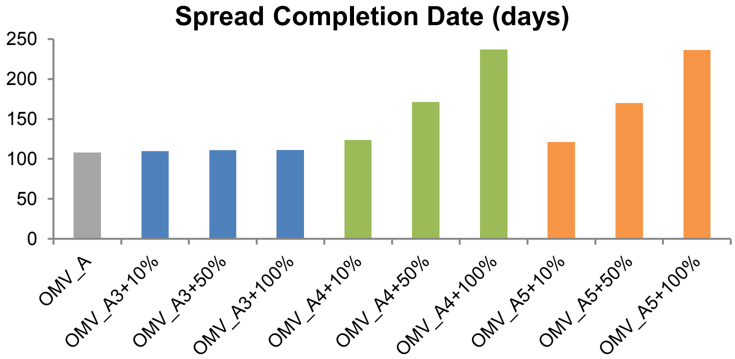


Figure 60: Spread distribution of OMV_B variations.

Mean and spread: OMV_C

Simulations with OMV_C and its derivatives give essentially the same results simulating OMV_A and OMV_B derivatives. Again outcome means and spreads correlate positively with input means and spreads (cp. Figure 70 and Figure 71). Furthermore mean and spread appear to be independent.
An influence on output shape (skewness, kurtosis) could not be detected.

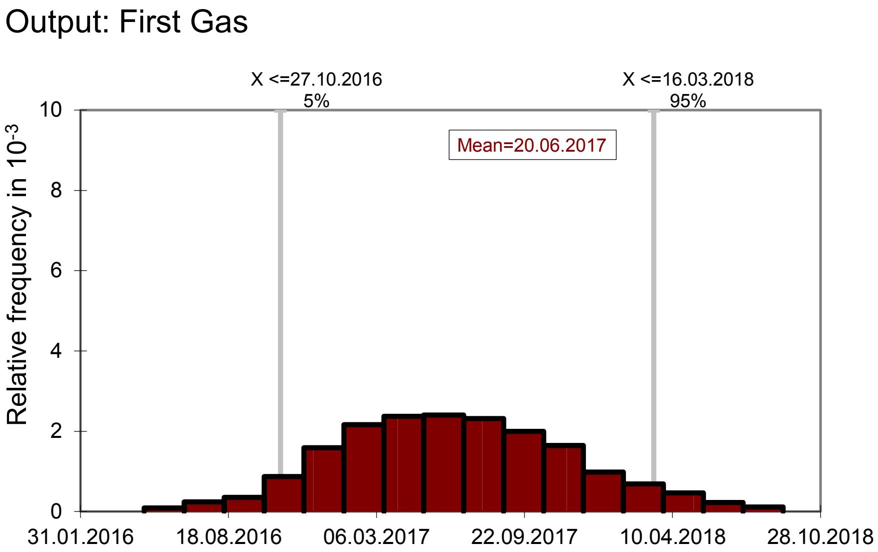


Figure 61: Output distribution for OMV_C3+10%.

Table 52: Distribution data for OMV_C3+10%.

Minimum	26.04.2016	Skewness	0,18	Mode	11.10.2016
Mean	20.06.2017	Kurtosis	2,75	5th Perc.	27.10.2016
Maximum	07.09.2018			95th Perc.	16.03.2018
Std Dev	154			Diff. X	505 days

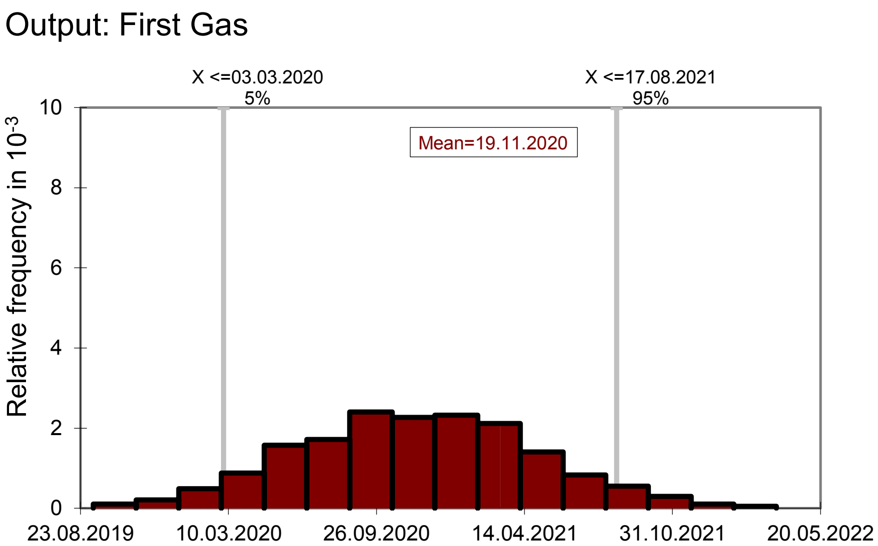


Figure 62: Output distribution for OMV_C3+50%.

Table 53: Distribution data for OMV_C3+50%.

Minimum	09.09.2019	Skewness	5,95E-02	Mode	06.04.2020
Mean	19.11.2020	Kurtosis	2,78	5th Perc.	03.03.2020
Maximum	21.03.2022			95th Perc.	17.08.2021
Std Dev	159			Diff. X	532 days

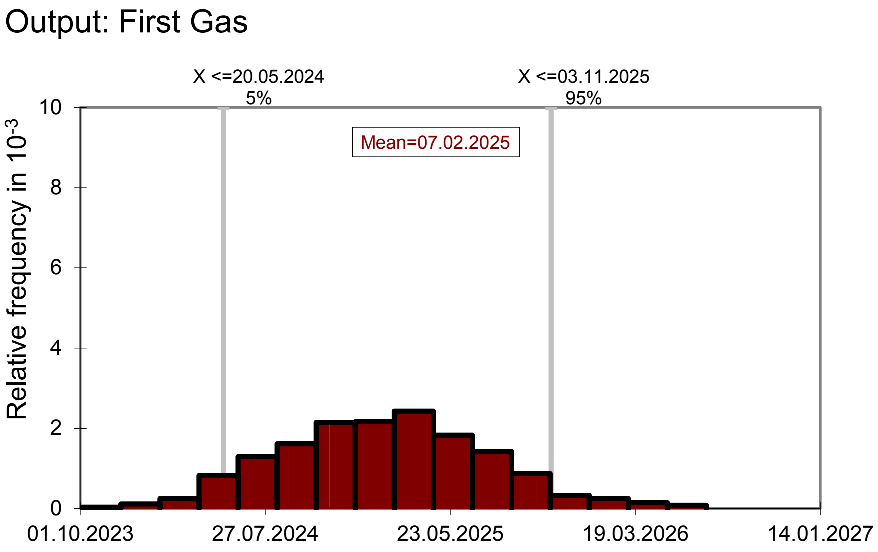


Figure 63: Output distribution for OMV_C3+100%.

Table 54: Distribution data for OMV_C3+100%.

Minimum	03.10.2023	Skewness	0,14	Mode	02.08.2024
Mean	07.02.2025	Kurtosis	2,88	5th Perc.	20.05.2024
Maximum	13.07.2026			95th Perc.	03.11.2025
Std Dev	166			Diff. X	532 days

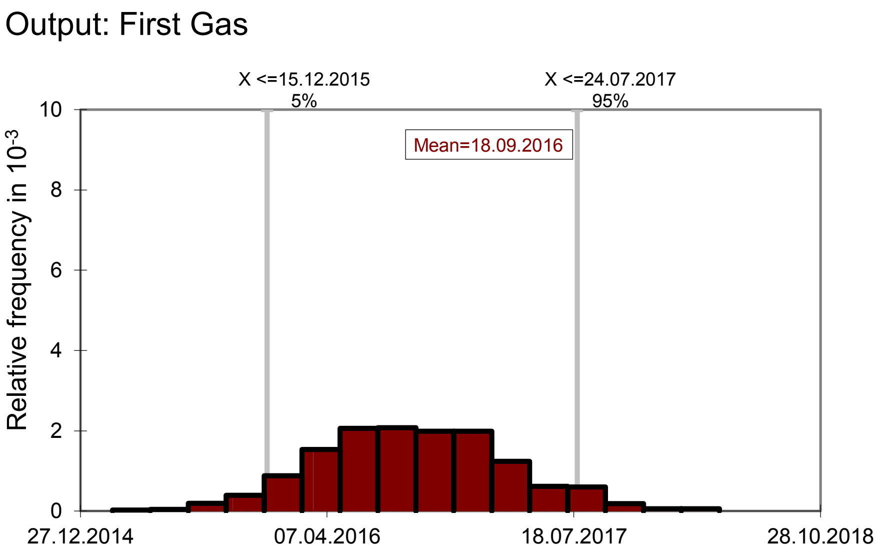


Figure 64: Output distribution for OMV_C4+10%.

Table 55: Distribution data for OMV_C4+10%.

Minimum	25.02.2015	Skewness	0,14	Mode	18.02.2016

Mean	18.09.2016	Kurtosis	2,98	5th Perc.	15.12.2015
Maximum	19.04.2018			95th Perc.	24.07.2017
Std Dev	177			Diff. X	587 days

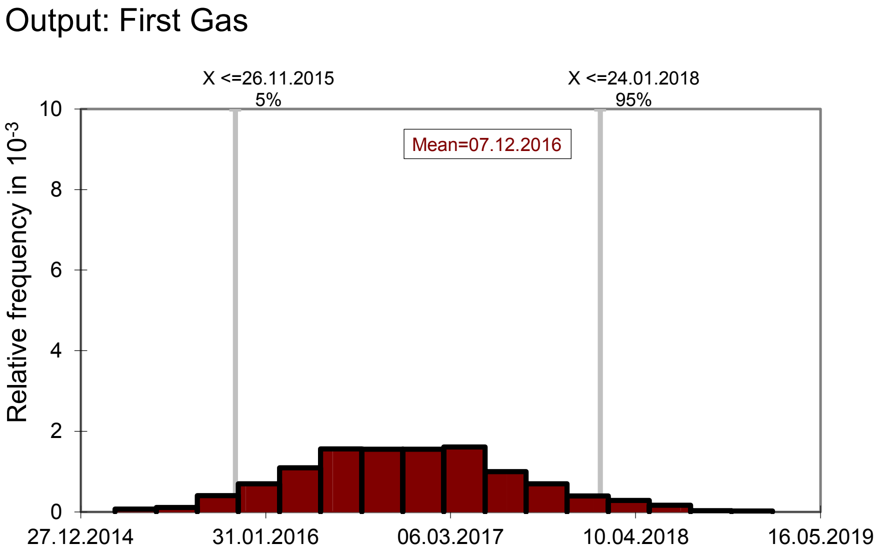


Figure 65: Output distribution for OMV_C4+50%.

Table 56: Distribution data for OMV_C4+50%.

Minimum	11.03.2015	Skewness	0,23	Mode	20.05.2016
Mean	07.12.2016	Kurtosis	2,91	5th Perc.	26.11.2015
Maximum	01.02.2019			95th Perc.	24.01.2018
Std Dev	235			Diff. X	790 days

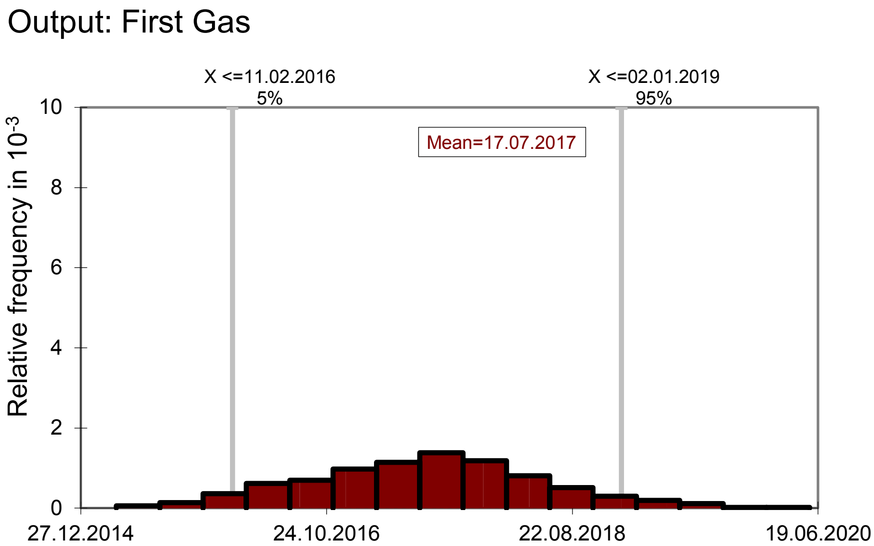


Figure 66: Output distribution for OMV_C4+100%.

Table 57: Distribution data for OMV_C4+100%.

Minimum	02.04.2015	Skewness	7,81E-02	Mode	05.01.2017
Mean	17.07.2017	Kurtosis	2,92	5th Perc.	11.02.2016
Maximum	27.05.2020			95th Perc.	02.01.2019
Std Dev	314			Diff. X	1056 days

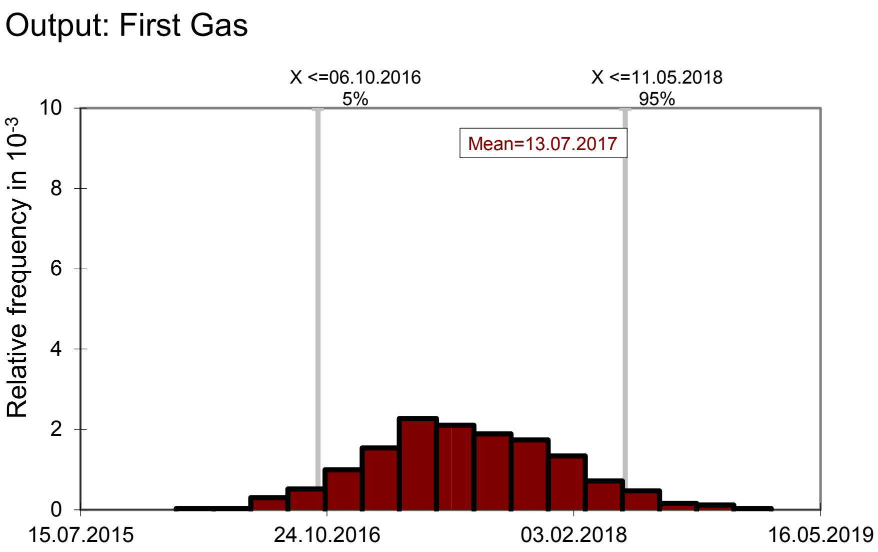


Figure 67: Output distribution for OMV_C5+10%.

Table 58: Distribution data for OMV_C5+10%.

Minimum	12.01.2016	Skewness	0,15	Mode	05.01.2017
Mean	13.07.2017	Kurtosis	2,91	5th Perc.	06.10.2016
Maximum	11.02.2019			95th Perc.	11.05.2018
Std Dev	177			Diff. X	582 days

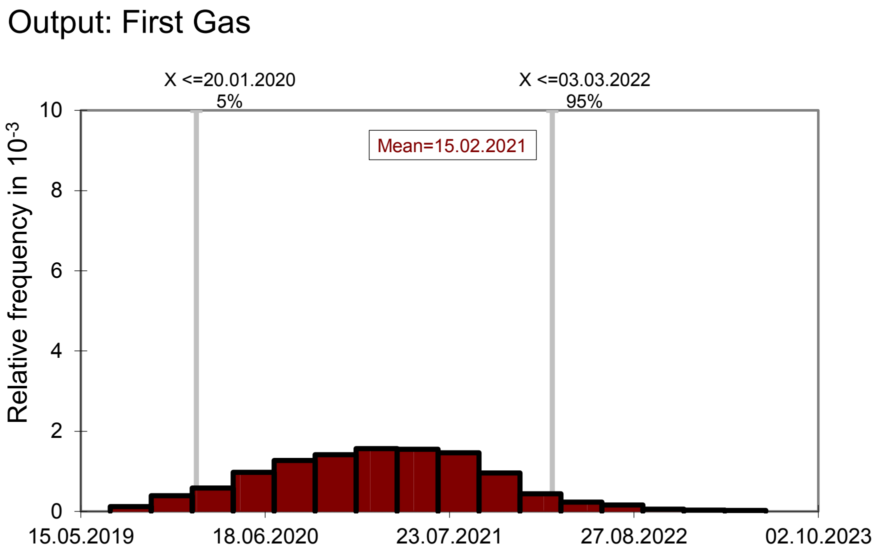


Figure 68: Output distribution of OMV_C5+50%.

Table 59: Distribution data for OMV_C5+50%.

Minimum	17.07.2019	Skewness	0,15	Mode	26.03.2020
Mean	15.02.2021	Kurtosis	2,94	5th Perc.	20.01.2020
Maximum	09.06.2023			95th Perc.	03.03.2022
Std Dev	237			Diff. X	773 days

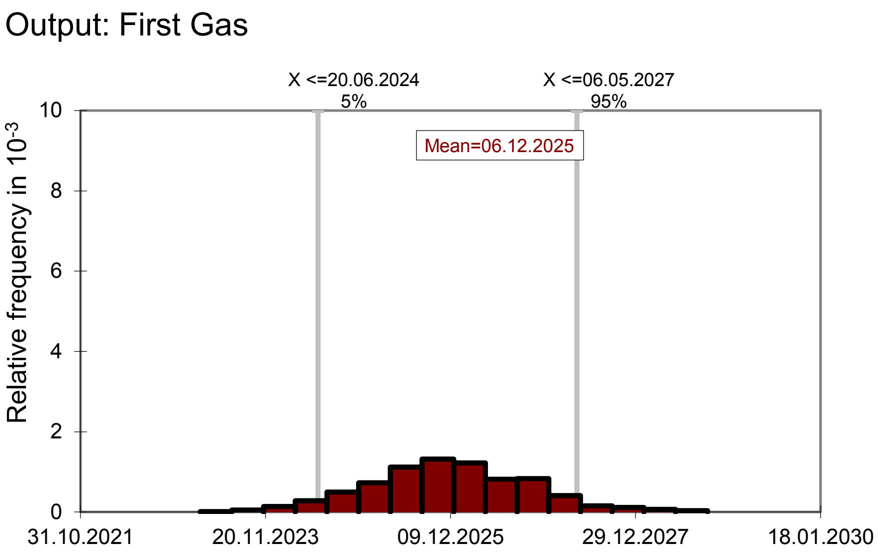


Figure 69: Output distribution for OMV_C5+100%.

Table 60: Distribution data for OMV_C5+100%.

Minimum	28.02.2023	Skewness	0,13	Mode	27.05.2024
Mean	06.12.2025	Kurtosis	3,06	5th Perc.	20.06.2024
Maximum	16.10.2028			95th Perc.	06.05.2027
Std Dev	321			Diff. X	1050 days

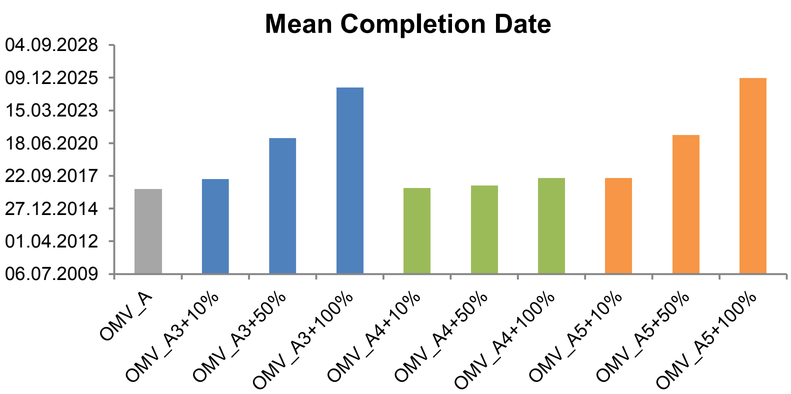


Figure 70: Mean distribution of OMV_C variations.

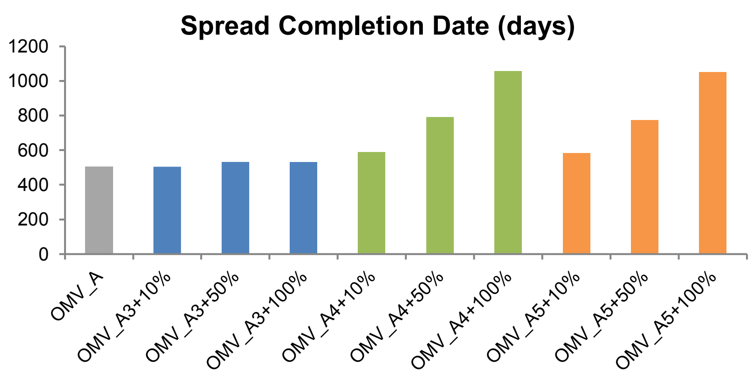


Figure 71: Spread distribution of OMV_C variations.

4.3.3 CLT

Third investigated parameter was the number of schedule tasks that might have influence on standard deviation and spread due to the Central Limit Theorem (literature says so, but engineers always want to try out). DoE “CLT” gave the framework for simulation procedure.

As demonstrated in Figure 72 to Figure 77 and even better presented in Figure 78 and Figure 79 CLT does have a significant effect on standard deviation and spread. Both decrease if the number of tasks goes up. That is not a realistic outcome because there are more and more uncertain tasks in the schedule. In contrast mean is settling in a two month range which is almost a constant value in comparison. So outcome distribution is fixed on the same place but becomes slimmer and slimmer which means less risky in other words. Thus, there is danger to underestimate risk and give in to unrealistic small spread.

Especially in case of lumping (summarize sub tasks to one task) it could be possible that due to “reverse CLT” spread and standard deviation increases if you lump heavily. For example lumping CLT25 into CLT10 will bring about 40 % more spread and about 50 % more standard deviation.

One interesting consequence of higher task number is that the output distribution becomes bi-modal or even tri-modal. This is simply because there are non-working days like weekends, where the possibility of project completion is zero. Thus if the spread is very small (typical for high task numbers), the X-axis resolution is high and you can see the non-working days as “zero possibility divider”.

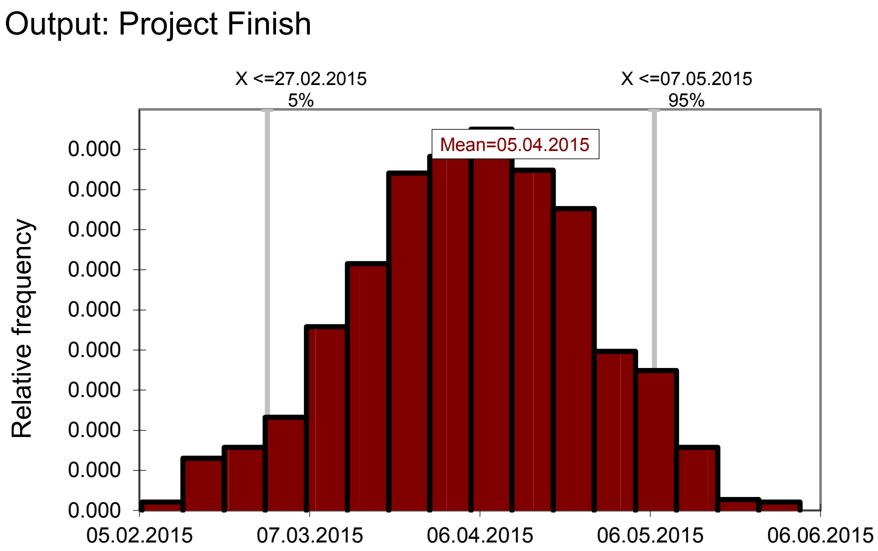


Figure 72: Output distribution for CLT10.

Table 61: Distribution data for CLT10.

Minimum	05.02.2015	Skewness	-0,13	Mode	25.03.2015
Mean	05.04.2015	Kurtosis	2,82	5th Perc.	27.02.2015
Maximum	02.06.2015			95th Perc.	07.05.2015
Std Dev	19,9			Diff. X	70 days

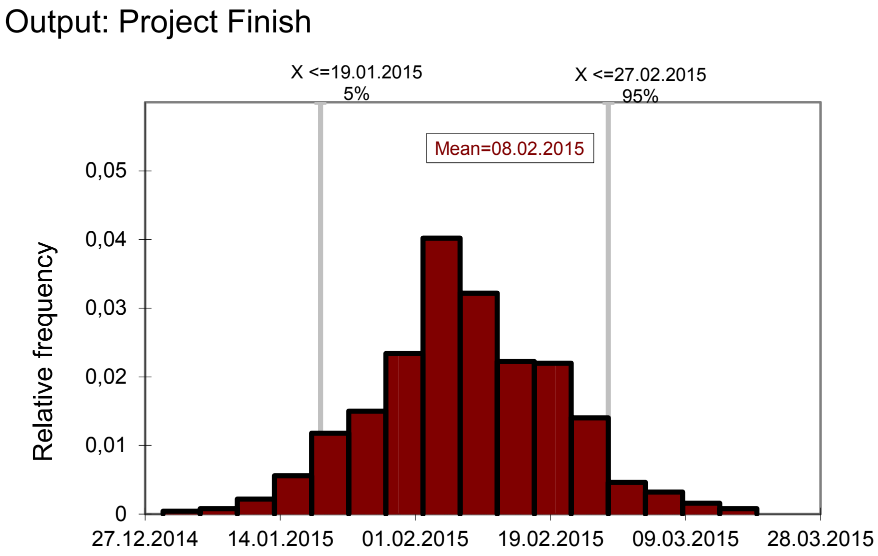


Figure 73: Output distribution for CLT25.

Table 62: Distribution data for CLT25.

Minimum	29.12.2014	Skewness	2,18E-02	Mode	02.02.2015
Mean	08.02.2015	Kurtosis	3,17	5th Perc.	19.01.2015
Maximum	19.03.2015			95th Perc.	27.02.2015
Std Dev	12,2			Diff. X	39 days

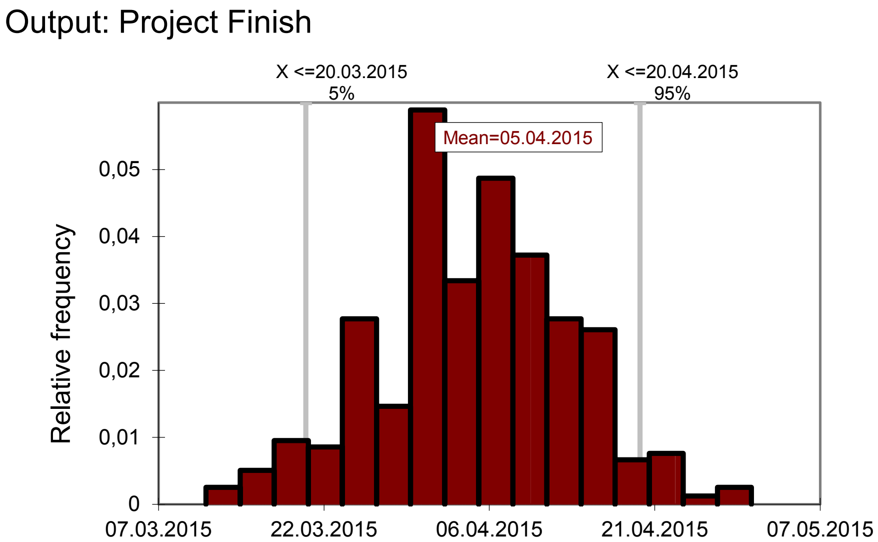


Figure 74: Output distribution for CLT50.

Table 63: Distribution data for CLT50.

Minimum	11.03.2015	Skewness	1,27E-03	Mode	26.03.2015
Mean	05.04.2015	Kurtosis	2,82	5th Perc.	20.03.2015
Maximum	30.04.2015			95th Perc.	20.04.2015
Std Dev	8,93			Diff. X	31 days

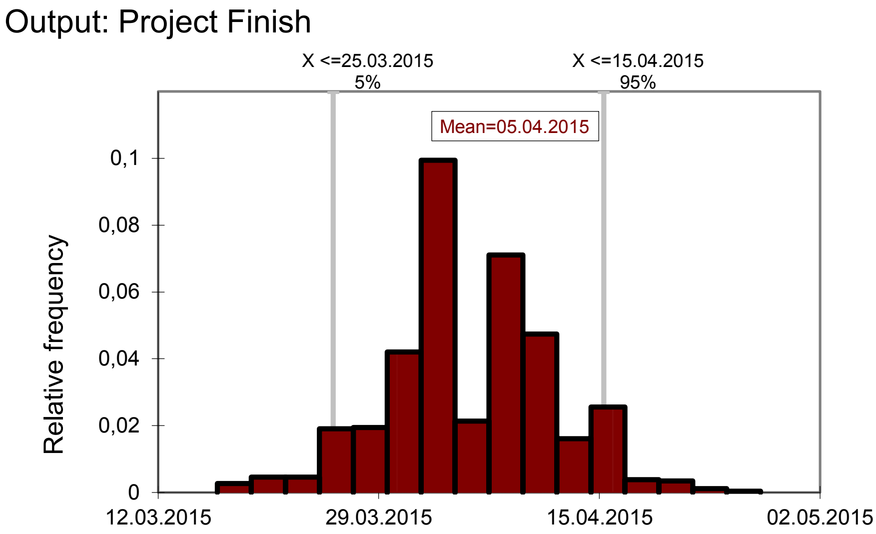


Figure 75: Output distribution for CLT100.

Table 64: Distribution data for CLT100.

Minimum	16.03.2015	Skewness	-5,39E-02	Mode	02.04.2015
Mean	05.04.2015	Kurtosis	3,14	5th Perc.	25.03.2015
Maximum	27.04.2015			95th Perc.	15.04.2015
Std Dev	6,34			Diff. X	21 days

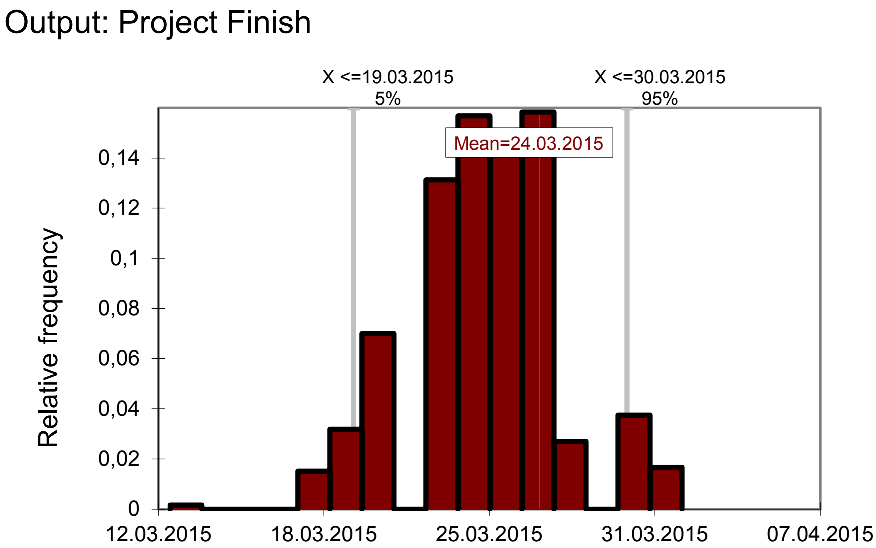


Figure 76: Output distribution for CLT500.

Table 65: Distribution data for CLT500.

Minimum	12.03.2015	Skewness	-0,18	Mode	23.03.2015
Mean	24.03.2015	Kurtosis	3,98	5th Perc.	19.03.2015
Maximum	01.04.2015			95th Perc.	30.03.2015
Std Dev	2,77			Diff. X	11 days

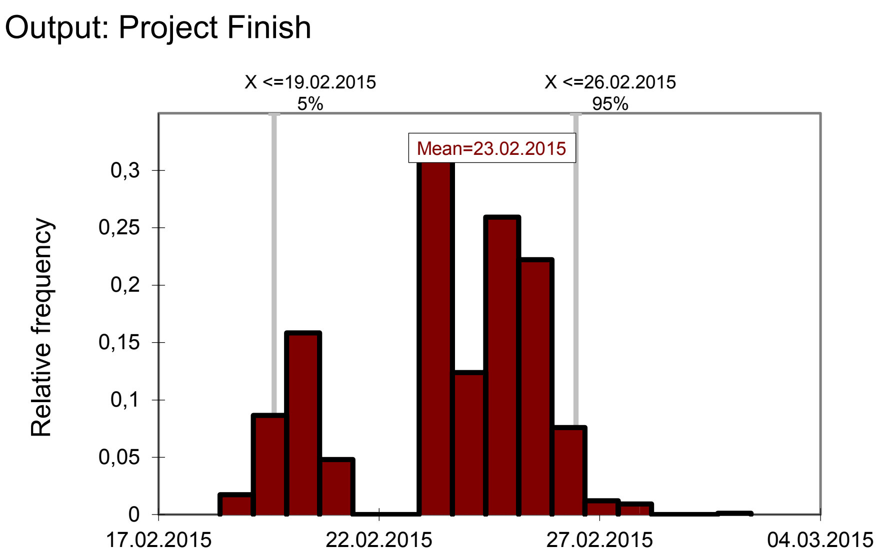


Figure 77: Output distribution for CLT1000.

Table 66: Distribution data for CLT1000.

Minimum	18.02.2015	Skewness	-0,63	Mode	23.02.2015
Mean	23.02.2015	Kurtosis	2,58	5th Perc.	19.02.2015
Maximum	02.03.2015			95th Perc.	26.02.2015
Std Dev	2,10			Diff. X	7 days

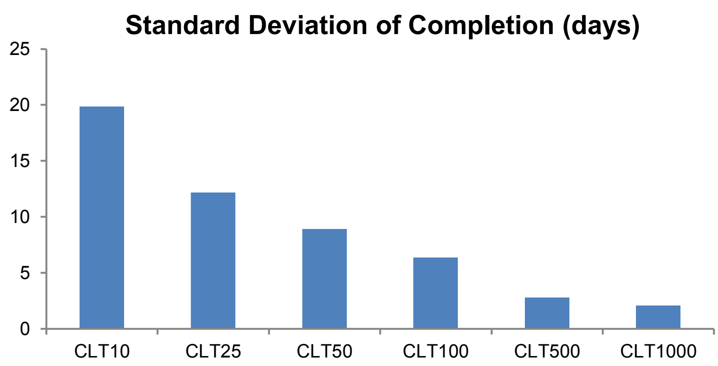


Figure 78: Standard deviation for different CLT dummy schedules.

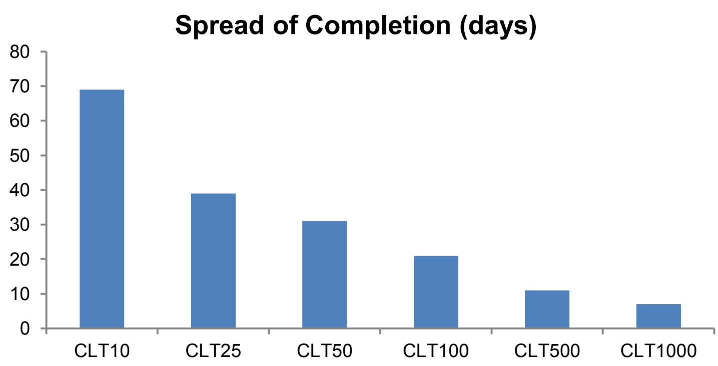


Figure 79: Spread for different CLT dummy schedules.

4.3.4 Constraints

Constraints have a great impact to schedule and completion dates. Testing this important parameter (often mentioned by project schedulers) is based on the DoE "Constraints". Figure 80 to Figure 86 (wrapped up in Figure 87 and Figure 88) indicate that constraints principally truncate input distributions and thus produce outcomes with less standard deviation and spread. The schedule is more inflexible and outliers are more rarely sequentially. MFO constraints are in this way more effective. FNLT and FNET have the same but less significant impact. With MFO`s mean and shape are not affected. With FNET constraints mean roams towards later dates, on the opposite FNLT constraints are bringing mean to earlier dates.

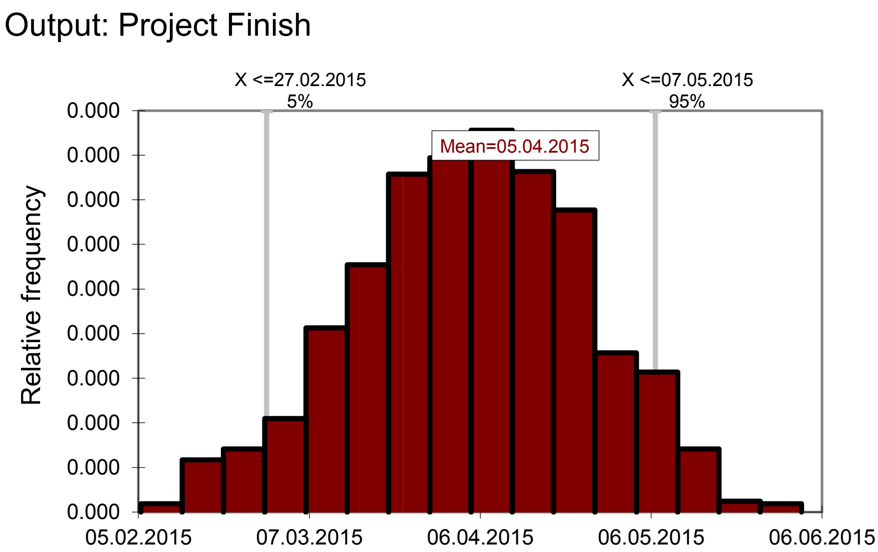


Figure 80: Output distributions CON_0%.
Table 67: Distribution data for CON_0%.

Minimum	05.02.2015	Skewness	-0,13	Mode	25.03.2015
Mean	05.04.2015	Kurtosis	2,82	5th Perc.	27.02.2015
Maximum	02.06.2015			95th Perc.	07.05.2015
Std Dev	19,9			Diff. X	69 days

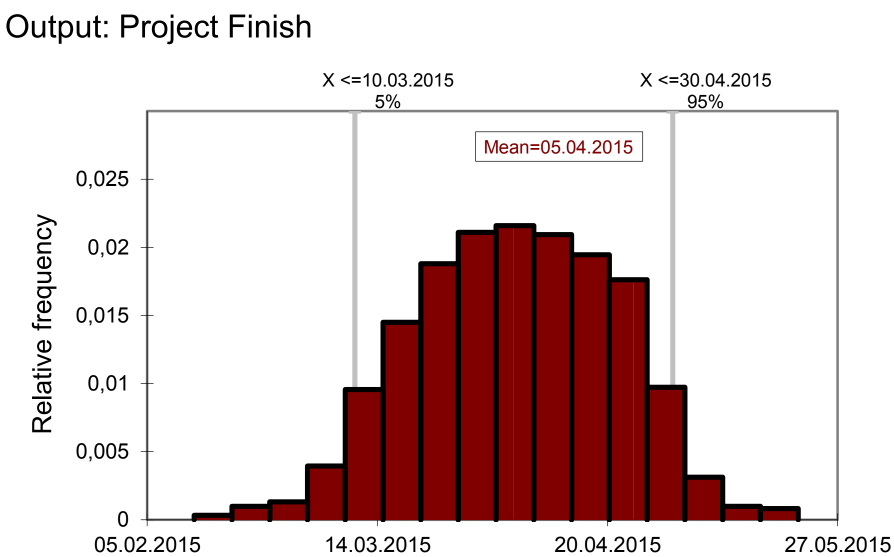


Figure 81: Output distribution for CON_MFO+20%.

Table 68: Distribution data for CON_MFO+20%.

Minimum	12.02.2015	Skewness	-0,11	Mode	03.04.2015
Mean	05.04.2015	Kurtosis	2,83	5th Perc.	10.03.2015
Maximum	20.05.2015			95th Perc.	30.04.2015
Std Dev	15,9			Diff. X	51 days

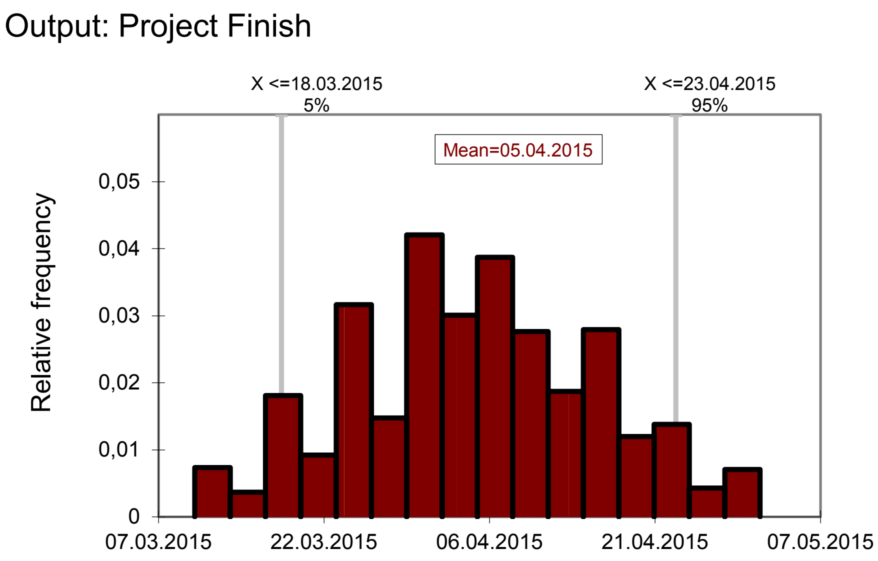


Figure 82: Output distribution for CON_MFO+40%.

Table 69: Distribution data for CON_MFO+40%.

Minimum	10.03.2015	Skewness	8,41E-02	Mode	30.03.2015
Mean	05.04.2015	Kurtosis	2,44	5th Perc.	18.03.2015
Maximum	01.05.2015			95th Perc.	23.04.2015
Std Dev	11,3			Diff. X	36 days

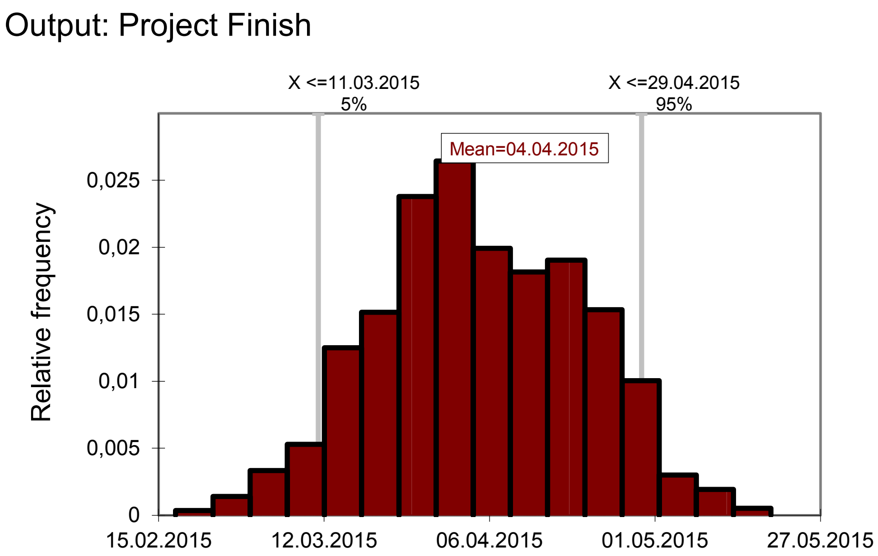


Figure 83: Output distribution for CON_FNLT+20%.

Table 70: Distribution data for CON_FNLT+20%.

Minimum	17.02.2015	Skewness	-1,49E-02	Mode	24.03.2015

Mean	04.04.2015	Kurtosis	2,57	5th Perc.	11.03.2015
Maximum	19.05.2015			95th Perc.	29.04.2015
Std Dev	15,6			Diff. X	49 days

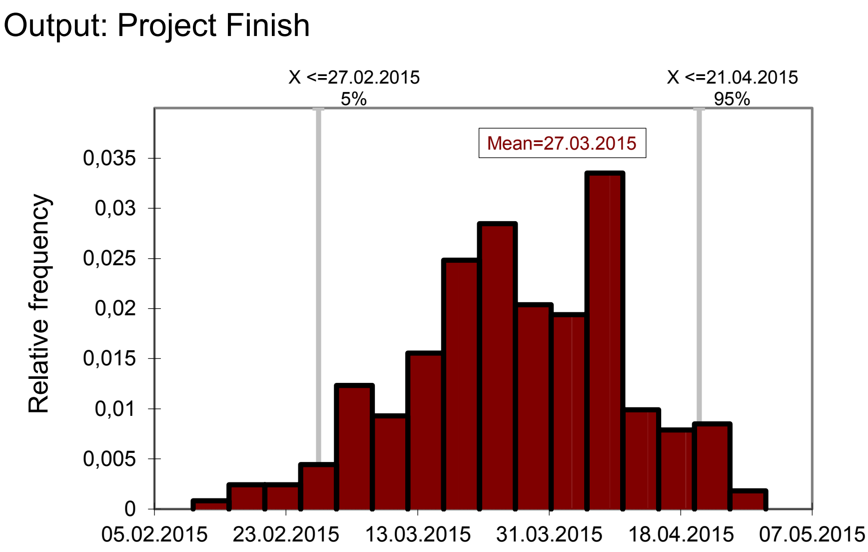


Figure 84: Output distribution for CON_FNLT+40%.

Table 71: Distribution data for CON_FNLT+40%.

Minimum	10.02.2015	Skewness	-0,23	Mode	04.03.2015
Mean	27.03.2015	Kurtosis	2,68	5th Perc.	27.02.2015
Maximum	30.04.2015			95th Perc.	21.04.2015
Std Dev	14,9			Diff. X	53 days

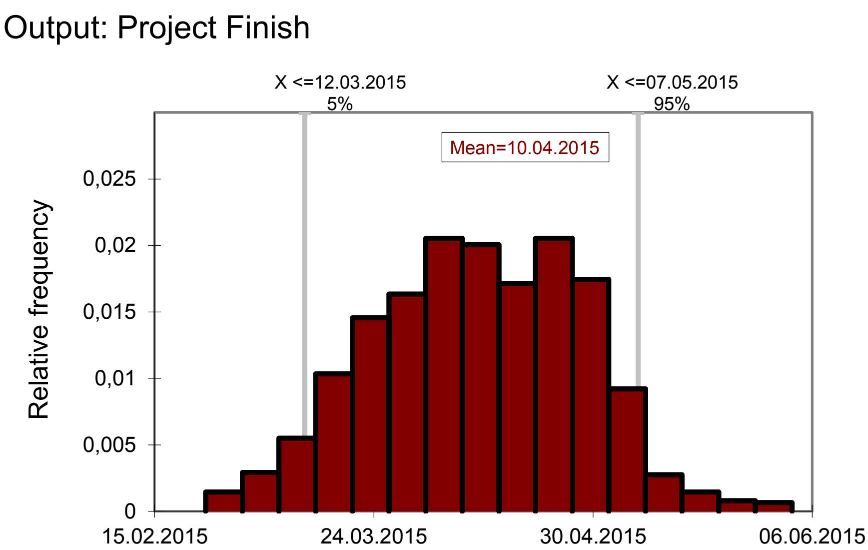


Figure 85: Output distribution for CON_FNET+20%.

Table 72: Distribution data for CON_FNET+20%.

Minimum	23.02.2015	Skewness	-1,14E-02	Mode	06.04.2015
Mean	10.04.2015	Kurtosis	2,72	5th Perc.	12.03.2015
Maximum	02.06.2015			95th Perc.	07.05.2015
Std Dev	17,3			Diff. X	56 days

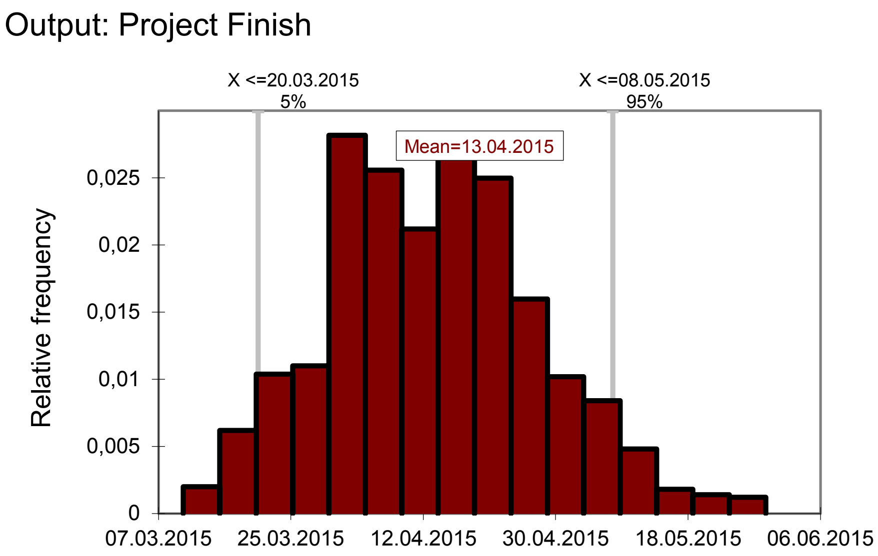


Figure 86: Output distribution for CON_FNET+40%.

Table 73: Distribution data for CON_FNET+40%.

Minimum	10.03.2015	Skewness	0,28	Mode	09.04.2015
Mean	13.04.2015	Kurtosis	2,85	5th Perc.	20.03.2015
Maximum	29.05.2015			95th Perc.	08.05.2015
Std Dev	14,5			Diff. X	49 days

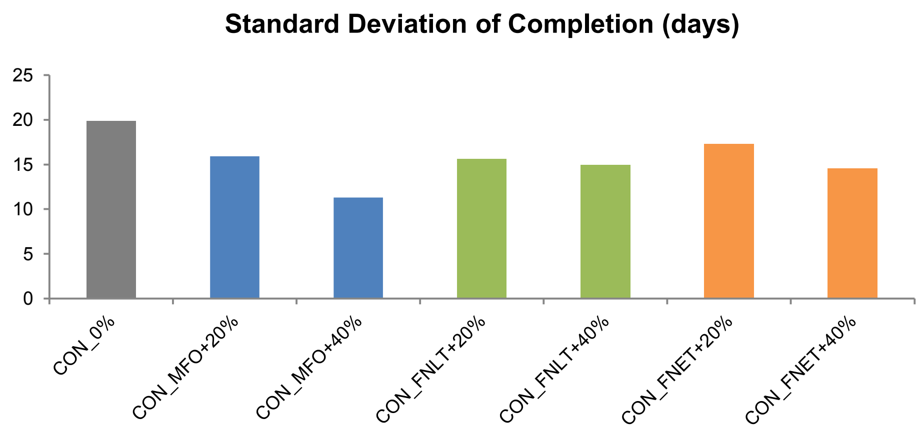


Figure 87: Standard deviation of different constraint dummy schedules.

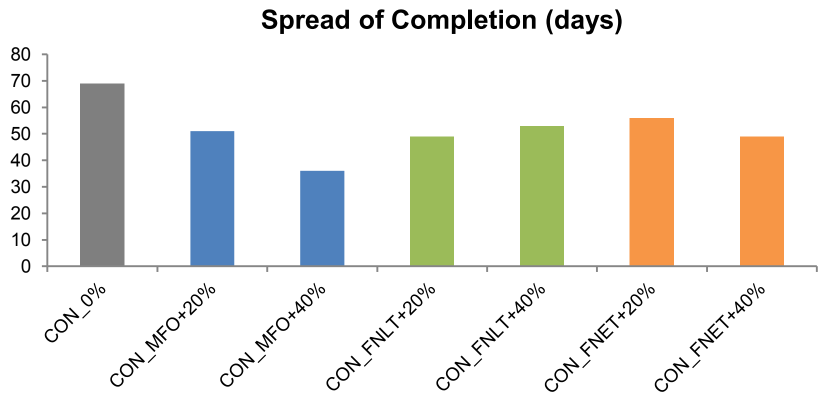


Figure 88: Spread of different constraint dummy schedules.

4.3.5 Correlations

Literature states that correlation makes your schedule more uncertain and produces more spread and different means in your output distribution. In other words, not taking correlation into account tends to be over-optimistic. Proving these statements simulation of 3 genuine OMV schedules containing a correlation coefficient mix (cp. Table 21) was conducted. DoE basis is given by Table 23 and Table 24.

Table 23: DoE “random correlation” for OMV_B.

Schedule ID	Correlated Task #	Correlation Coefficients
OMV_B	/	/
COR_OMV_B	All distributed inputs (27)	Mix

As you can see in Figure 89, Figure 90 and Figure 91, correlation seems to widen expected project completion data. Mean value and spread increases as correlation takes place. That is what literature says and it is comprehensible because linking between tasks is kind of an additive uncertainty dimension. In reality there is a high possibility that tasks are not independent from each other, so implementing correlation in your simulation model can give a more reliable output.

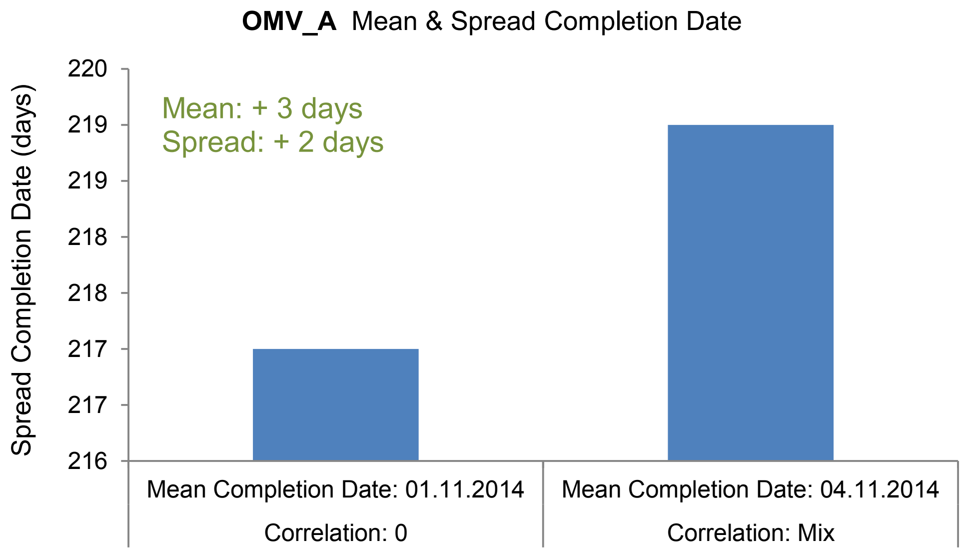


Figure 89: Correlation simulation for OMV_A.

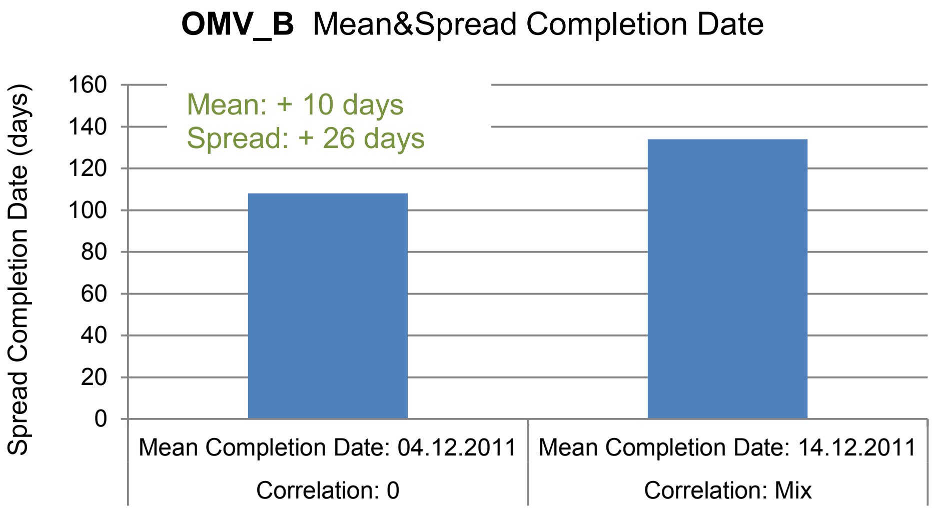


Figure 90: Correlation simulation for OMV_B.

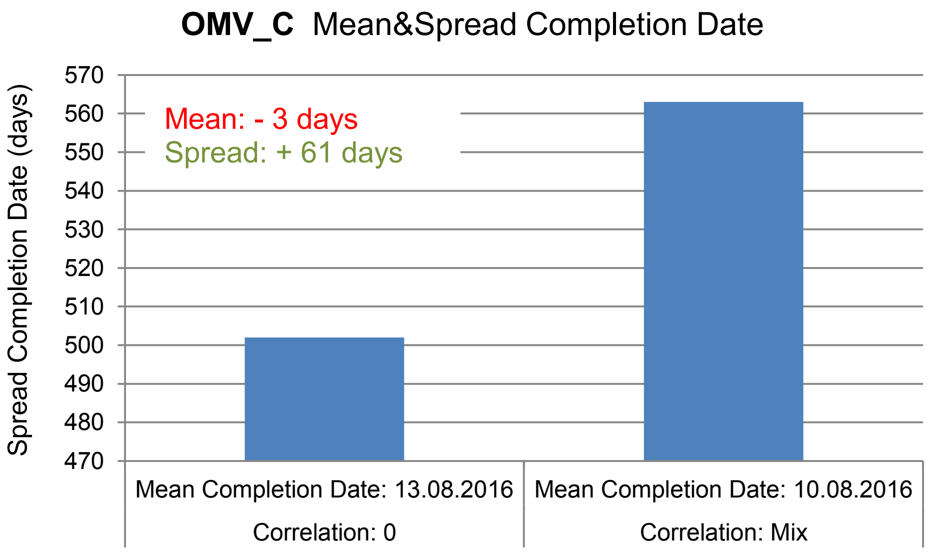


Figure 91: Correlation simulation for OMV_C.

4.4 PSA Key Factors Given by MC Simulations

Comparing key factors given by literature and applied PSA the following parameters are affirmed by MC simulations:

- Input distribution shape: Use distribution catalogue!
- Input mean and spread: Use precise/realistic input values!
- CLT: Minimize schedule task number and/or correlate them!
- Constraints: Do not use them at all and make schedule flexible!
- Correlations: Correlations are important regarding literature and simulations.
  - Determine correlation coefficients parallel to duration data estimation!
  - Additionally the relationship between correlation and CLT has to be investigated in future. It will be interesting, if literature notion (correlation decrease CLT effect) goes hand in hand with MC simulation.

5 Summary of Work

This chapter will wrap up best practices and lessons learnt of PSA on basis of oil and gas field development projects. The key points of this work: theoretical probabilistic schedule analysis, input data estimation, company survey and applied probabilistic schedule analysis melt together in Chapter 5.1: How to do a PSA?

5.1 How to do a PSA?

Suggested basis for conducting a PSA is the flowchart shown in Figure 92. Firstly an estimation workshop takes place where probabilistic input data is generated. Ideally this is combined with producing the deterministic baseline schedule, thus the same people can work on both parts. As result a probabilistic schedule is achieved that can enter now the MC simulation stage. Schedule risk drivers are (hopefully) detected afterwards and the schedule can be optimised. Now there are two options: either the reviewed schedule goes back to simulation (looping) or it is perfectly enough to execute.
One important process step is filling, checking and maintaining an MC input value database. This database builds the foundation for subsequent estimation workshop and will be an assessment reference in the following process steps. Figure 93 illustrates embedding of database in all important process steps.
Finally the true intent of PSA is to encompass the range of uncertainty giving some sort of confidence intervals in order to make better decisions and highlight important task durations and cost drivers.

5.1.1 PSA Flowchart

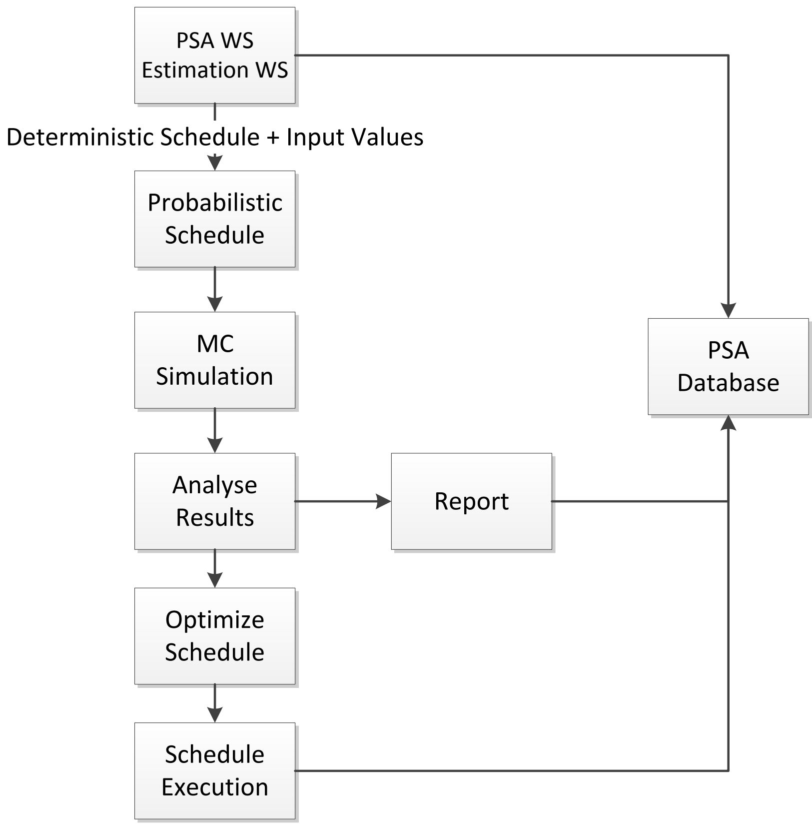


Figure 92: Suggested flowchart of a PSA.

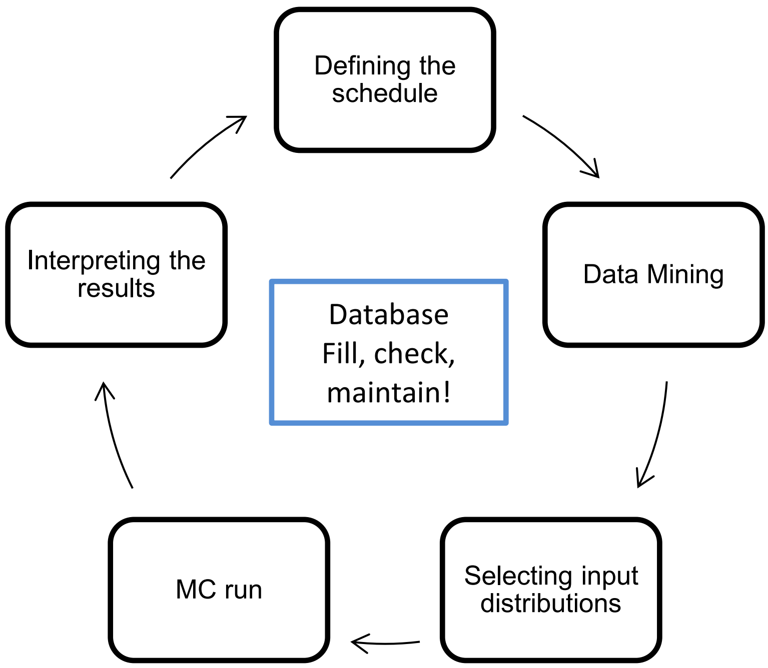


Figure 93: Cycle process and working with a database.

Figure 96 emphasizes that PSA is a looped process. As the project moves on, project plans (with all data) must be revisited and updated. Some risks and uncertainties will then be obsolete. Some new one will be introduced and so on. All in all PSA is a dynamic process which must be permanently documented and communicated.

The following points deal with the most important flowchart steps in detail:

5.1.2 Input Data Estimation by PSA/Estimation Workshops

Using the main statements of Chapter 2 it will be necessary to initialize some kind of workshop for data gathering (group estimation, bias controlling etc.). So the main goal is to find out how an appropriate estimation workshop could be conducted. A good one will provide the right input data for the Monte Carlo analysis such as distribution parameters, distribution forms and maybe even correlation coefficients (see Figure 94).

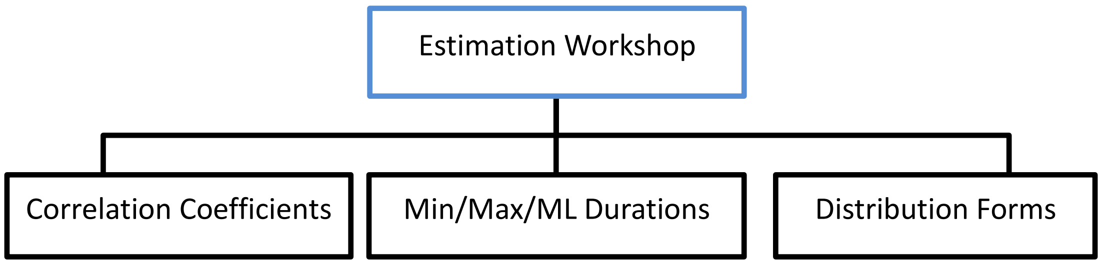


Figure 94: Output of an estimation workshop.

So we have to encounter the challenges of estimation biases and Black Swans and furthermore bring in the benefits of crowd wisdom and diversity. I propose the following way of doing it:

<u>Group Member Number</u>

You need 10 at minimum (the more the better).

Group Member Diversity

Take participants from different departments (like IT, Research, Sales etc.), but include the schedule-affected one!

TIP: in preparation of the estimation workshop you can staff a permanent personnel pool of interested and motivated people.

Interrogation Process

Interviews should always observe the rules of a wise group:

- Diversity (various departments)
- Independence (isolated individual estimation)
- Decentralization (no expert leading, full involvement of the group)
- Aggregation (workshop moderator has to aggregate opinions and loop when needed)

Subsequently there are two basic ways to conduct the interviews:

1) Delphi Method (workshop is decentralized in time and space)

The Delphi Method is a forecasting method relying on a panel of independent experts. These experts are answering a tailored questionary iteratively with feedback loops. After a certain number of iterations the collective values are calculated, for example the spread (Equation 5):

$$S = \mathrm{Max}(x_i) - \mathrm{Min}(x_i). \qquad 5$$

Feedback can evocate intensive examination and correction of the own estimations. Furthermore independence and decentralization of the participants are kept. A good selection of the interview partners should give you diversity. Finally a controlled aggregation of the information flow takes place.
The Delphi Method could be carried out in large circuits like e-mail loops or in small circuits like personal interviews (talk-estimate-talk).

2) Without Delphi Method

If the Delphi Method is not possible (too time consuming etc.) and there is a classical closed workshop estimation taking place, it is very important (besides all rules of a wise crowd) to isolate the group members during the personal estimation process. This avoids herding effects.

Estimation Process

Although extensive literature on subjective estimation[94,95,96] suggests different methods, I worked out an own approach based on Hawkins[97].

Three-Point Estimation

Normally you are asking for a three-point approximation of the distribution. For symmetrical distributions like Normal, min, max and a central value is mostly enough to characterize the distribution.

94 Cp. Spetzler (1975)
95 Cp. Howard (1989)
96 Cp. Goodwin (1999)
97 Cp. Hawkins (2002)

Skewed distributions need a little more elaborate approach. Mathematicians[98,99,100,101,102] provided weightings best suited for three-point estimation, based on a variety of points:[103]

- PERT: Extreme fractiles and the Mode (0,16; 0,66; 016)
- Moder & Rodgers: P5, the Mode, P95 (0,185; 0,63; 0,185)
- Swanson-Megill: P10, the Median, P90 (0,3; 0,4; 0,3)
- Extended Pearson & Tukey: P5, the Median; P95 (0,185; 0,63; 0,185)

The most accurate three-point estimation of a skewed distribution is hereby the extended Pearson & Tukey. However, it is more difficult to estimate low probability values, such as P5 or P95. If you take this into account, the better method is Swanson-Megill, which uses fractiles closer to the center like P10 etc. Pioneering in risk analysis regarding the oil industry Megill is the main influence behind the common use of the "80 % confidence interval" (→*confidence interval*) in today`s oil companies.[104]

Biases Confrontation

Confront participants with possible subjective biases to make them aware of it!

Clearly Define the Asked Variable

Remove ambiguity about the estimated value by discussing assumptions and units of measure![105]

Fix Extreme Outcomes

Ask for a list of events that can lead to extreme duration outcomes. This step encourages imagination of low probabilities!

Influence Diagram

Draw a simple influence diagram, so everybody can capture in short the key factors for the measured durations!

„Outside in" Method

This method can help you to get a correct distribution, because a better range leads to a better mean:

- Think of a range of uncertainty (spread), in which the true value could be. A realistic range is often more important than the distribution form!
- Then assign a relatively small probability to the range encompassing the truth, say 40 % (this is your confidence interval)!
- Now choose the right distribution form. Normal, Lognormal (for something beginning at 0 and reaching to infinity etc.) and Triangular are very often sufficient! In general, when the absolute difference between P10-P50 and P50-P90 is not equal, assume a skewed distribution. Otherwise it should be a symmetrical one.106
- Think in whole distributions, also reality-check the 60 %-point, the 90 %-point etc.!

98 Cp. Brown (1974)
99 Cp. Davidson (1976)
100 Cp. Megill (1977)
101 Cp. Pearson (1965)
102 Cp. Keefer (1983)
103 Cp. Hawkins (2002)
104 Cp. Hawkins (2002)
105 Cp. Hawkins (2002)
106 Cp. Hawkins (2002)

- Make some feedback-loops and do step 1-4 again, then you get more ranges with more distributions, playing one against each other!
- If you have in your opinion the right range and form, then set the mean value!

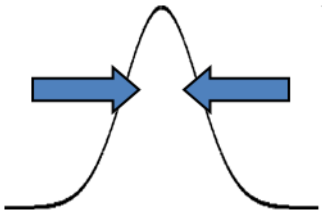


Figure 95: "Outside in" means iteratively narrowing the range from out sided start values to get a correct located mean value.

Reality Check

Performing a reality check can bring large deviations from historical or analogous time data to light. Furthermore check distributions, P5, P10, P50, P90, P95 etc. for logic and consistency!

5.1.3 Deterministic Schedule

The base schedule (normally a deterministic schedule) is the backbone of every simulation. A non-logical, wrongly linked and inflexible Gantt chart kills every simulation. Another problem associated with Monte Carlo simulations is that, if a project slips, project managers usually perform certain actions. So the base schedule is changing without simulation control. As already mentioned in Chapter 1.4.6, a regular schedule and simulation revision is needed.

TIP: Generally it can be said that a new schedule always needs a new simulation.

5.1.4 MC Simulation

Monte Carlo simulation is arguably better than other methods due to the following points:

- Monte Carlo can use any distribution form.
- Monte Carlo calculates the actual critical path within every simulation run.
- Monte Carlo gives you many "case studies" within a very short time.
- Monte Carlo offers additional information like critical indices (probability that a task lies on the critical path) or sensitivity charts based on a correlation analysis.

Moreover there are some key factors in simulation:

Precise and Accurate Input Values

Stochastic modelling is worthless when fed with incorrect data or in other words "garbage in – garbage out". Also a regular update of input data and associated distributions using performance measurement data is very useful (database!).

TIP: use data estimation workshops, Delphi Method etc. introduced in Chapter 2!

Distribution Shape

All tested OMV project schedules have a task number below 200 and moreover not every task does have probabilistic parameter.
CLT does not seem to have a big effect here. Instead the output distributions generated by @Risk simulations show specific characteristics of the input distributions.

Thus every task should get its appropriate distribution based on the distribution catalogue (cp. Chapter 1.3.4) and a profound task environment analysis.

TIP: use distribution catalogue introduced in Chapter 1.3.3.

Central Value and Spread Value

Varying input means and spread values positively correlates with output means and spreads. In addition it seems that for example a 10% increase on the input side produces a 10% increase on the output concerning the examined OMV schedules. Further investigations (in fact a higher number of schedules in the DoE) should be made to verify significance of correlation.

TIP: use data estimation workshops, Delphi Method etc. introduced in Chapter 2!

Central Limit Theorem (CLT)

The number of input distributions does have an effect on the spread and the standard deviation of the outcome. The higher this number is the minor spread and deviation are. Surprisingly this takes effect even in the common task number range (10 to 100) and it should be observed when lumping as a task number reducer is done (note: lumping makes the spread bigger, so lumping seems to bring you on the safe side).

- 3 main ways to reduce the impact of central limits:
  - Restrict the number of input variables (lumping)
  - Combat too narrow input ranges: importance of realistic Min/Max.
  - Use correlation to introduce dependencies between inputs.
    - Estimation group should give correlation coefficient from 0.0 (no correlation) to 1.0 (inputs perfectly correlated).

TIP: minimize schedule tasks and/or correlate them!

Constraints

Inflexible constraints like Must Finish On (MFO) do have an impact on probabilistic analysis. It makes output spreads and standard deviation smaller. Semi flexible ones like Finish No Later Than (FNLT) do have a similar but minor impact. For optimal schedule flexibility, even Microsoft recommends allowing MS Project to use flexible constraints to calculate the start and finish dates for tasks based on the entered durations and task dependencies (cp. http://office.microsoft.com). Only if you have unavoidable constraints, such as an event date that cannot be moved, it should be considered setting a constraint for a task manually. Only with no or flexible constraints a PSA can give you a picture of all possible project durations outcomes and therefore a realistic risk estimation basis.

TIP: Do not use them at all and make schedule flexible!

Correlations

Correlations increase the risk of unexpected and/or extreme completion dates. Therefore they must be estimated together with the base values like durations in the estimation workshop.

Sense Check

Always test, interrogate and review the model you using. Breaking the model into sequential phases could be advantageous. Furthermore the output can be broken into P10, P50 and P90 outcomes. Too small or too big differences between those values could be a warning sign. Remember that ranges are frequently more poorly estimated than central values. So P10 and P90 should be more affected by input changing than P50.[107]

5.1.5 Model Sensitivity Analysis

When a MC model is built, it should be run and re-run many times to check sensitivities. Tornado charts and scenario overlays are essential.
In first think about a base plan (1) with no explicit risk event, only uncertainties that still exist if everything goes right. With this rough plan we can find stages or phases to focus resources for biggest benefit. Next (2), we execute the plan with risks that are manageable, but not 100 % mitigable. This risk event sensitivity will rank those events that most affect our schedule. The following model runs (3) should focus on implemented mitigation strategies. That will measure mitigation effectiveness and highlight the greatest pay-offs. In the end repeat (2) and (3) using the base model, but with risk events that are beyond our influence (like dictatorship downfalls etc.). Doing so you can determine if the hierarchy of uncertainties or risks changes. Finally from the above sensitivities and percentiles you can draw a contingency and allocation plan.[108]

TIP: use tornado charts to detect and optimize schedule key drivers!

5.1.6 Database

As already mentioned a well maintained database is the foundation of well-done PSA. All estimations and further the real outcomes should go to a project archive, so project managers on future projects can learn from it.

[107] Cp. Akins (2005), p. 6 seq.
[108] Cp. Peterson (2005)

6 References

Allbery, R. (2008): Review: The Black Swan. URL: http://www.eyrie.org/~eagle/reviews/books/1-4000-6351-2.html. (Access: December 13, 2011).

Akins, W. A. et altera (2005): Enhancing Drilling Risk & Performance Management Through the Use of Probabilistic Time & Cost Estimating.

Brown, R. et altera (1974): Decision Analysis for the Manager. New York: Holt, Reinhart & Winston.

Book, Stephen A. (2002): Schedule Risk Analysis: Why It is Important and How to Do It, Presentation at the Ground System Architectures Workshop in El Segundo, CA.

Capen, E. (1976): The Difficulty of Assessing Uncertainty. Journal of Petroleum Technology, p. 843-850.

Capen, E. et altera (1995): A Problem With Experts. SPE Hydrocarbon Economics and Evaluation Symposium, 26-28 March 1995, Dallas, Texas.

Chase, A. P. (1991): A Survey of Research in the Application of Tolerance Analysis to the Design of Mechanical Assemblies. URL: http://adcats.et.byu.edu/Publication/91-1/DesRes_w_figs.html. (Access: May 21, 2011).

Cronquist, C. (1991): Reserves and Probabilities - Synergism or Anachronism? JPT.

Davidson, L.; Cooper, D. (1976): A Simple Way of Developing a Probability Distribution of Present Value. Journal of Petroleum Technology, p. 1069-1078.

Evers, J. F. (1973): How to Use Monte Carlo Simulation in Profitability Analysis. SPE Rocky Mountain Regional Meeting, 15-16 May 1973, Casper, Wyoming.

Goodwin, P.; Wright, G. (1999): Decision Analysis for Management Judgment. Wiley.

Hawkins, J. T.; Cunningham, P. C. (2002): Improving Stochastic Evaluations Using Objective Data Analysis and Expert Interviewing Techniques. SPE Annual Technical Conference and Exhibition 2002.

Hawkins, J. et altera (2002): Improving Stochastic Evaluations Using Objective Data Analysis and Expert Interviewing Techniques. SPE Annual Technical Conference and Exhibition, 29 September-2 October 2002, San Antonio, Texas.

Hertz, D. (1964): Risk Analysis in Capital Investments. Harvard Business Review.

Hogarth, R. (1987): Judgment and Choice. Wiley-Interscience, pp. 290.

Holtz, M. (1993): Estimating Oil Reserve Variability by Combining Geologic and Engineering Parameters. SPE Hydrocarbon Economics and Evaluation Symposium, 29-30 March 1993, Dallas, Texas.

Howard, R.; Matheson, J. (1989): The Principles and Applications of Decision Analysis. SDG.

Hulett, D. (2009): Practical Schedule Risk Analysis. England: Gower Publishing Company.

Keefer, D.; Bodily, S. (1983): Three-Points Approximations for Continuous Random Variables. Management Science, Vol. 22 (No. 3), p. 595-609.

Kahneman, D. (2011): Thinking - Fast and Slow, Macmillan Us (25. Oktober 2011), ISBN-10: 0374275637, ISBN-13: 978-0374275631

Mauboussin, M. J. (2007): The Importance of Diverse Thinking. Legg Mason Capital Management.

McCabe, B. (2003): Monte Carlo Simulation for Schedule Risks. Proceedings of the 2003 Winter Simulation Conference, p. 1561-1565.

McCray, A. (1977): Petroleum Evaluations and Economic Decisions. Englewood Cliffs, New Jersey: Prentice-Hall Inc.

McIntosh, J. (2004): Probabilistic Methods in Well Construction. MSc Thesis.

N.N.: MedCalc: easy-to-use statistical software. URL: http://www.medcalc.org/manual/summary_statistics.php. (Access: June 3, 2011).

Megill, R. (1977): An Introduction to Risk Analysis. Tulsa: Petroleum Publishing Company.

Megill, R. (1985): Evaluating & Managing Risk - A Collection of Readings. S. Publishing,.

N.N.: MindTools: Critical Path Analysis and PERT Charts. URL: http://www.mindtools.com/critpath.html. (Access: January 23, 2011).

Mochal, T. (2002): Use Monte Carlo Analysis for Sophisticated Scheduling. URL: http://www.techrepublic.com/article/use-monte-carlo-analysis-for-sophisticated-scheduling/1027705. (Access: February 2, 2011).

Murtha, J. (1987): Infill Drilling in the Clinton: Monte Carlo Techniques Applied to the Material Balance Equation. SPE Eastern Regional Meeting, 21-23 October 1987, Pittsburgh, Pennsylvania.

Murtha, J. (1993): Decisions Involving Uncertainties-An @RISK Tutorial for the Petroleum Industry. Newfield, New York: Palisade Corp.

Murtha, J. (1993): Incorporating Historical Data in Monte Carlo Simulation. SPE Computer Applications, Volume 6, Number 2 , p. 11-17.

Murtha, J. (1997): Monte Carlo Simulation: Its Status and Future. Distinguished Author Series (SPE), p. 361-373.

Murtha, J. A. (2002): Sums and Products of Distributions: Rules of Thumb and Applications. SPE 77422.

N.N.: EPM Live Workengine. URL: http://www.workengine.com (Access: May 21, 2011).

N.N.: MS Office Support: About Constraints. URL: http://office.microsoft.com (Access: February 2, 2011).

N.N.: The Wisdom of Crowds. URL: http://en.wikipedia.org/wiki/The_Wisdom_of_Crowds (Access: November 21, 2011).

N.N.: Book review: "The difference: How the power of diversity creates better groups, frims, schools, and societies by Scott E. Page". URL: http://press.princeton.edu/titles/8353.html. (Access: March 11, 2011).

N.N.: Guide to Using @Risk. URL: http://www.palisade.com.(Access: August 20, 2010).

N.N.: www.riskAMP.com. URL: http://www.riskAMP.com.(Access: March 11, 2011).

Page, S. E.: Presentation: The Power of Diversity. URL: http://fora.tv/2007/03/14/Power_of_Diversity#fullprogram. (Access: March 6, 2011, my birthday ;).

Page, S. E. (2007): The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies. Princeton, New Jersey: Princeton University Press.

Page, S. E. (8. January 2008): In Professor's Model, Diversity = Productivity. (C. Dreifus, Interviewer) New York Times.

Palisade Corporation (2005): @RISK for Project Help. @RISK Version 4.1.4-Professional Edition.

Pearson, E.; Tukey, J. (1965): Approximate Means and Standard Deviations Based on Distances Between Percentage Points of Frequency Curves. Biometrika, Vol. 52 (No. 3-4), p. 533-546.

Peterson, S. K. et altera (2005): Risk and Uncertainty Management-Best practices and Misapplications for Cost and Schedule Estimates. SPE 97269.

Purvis, D. C. (2003): Judgement in Probabilistic Analysis. (S. o. Engineers, Hrsg.) SPE Hydrocarbon Economics and Evaluation Symposium 2003.

Rose, P. (1987): Dealing with Risk and Uncertainty in Exploration: How Can We Improve? AAPG Bulletin, Vol. 71 (No. 1), p. 1-16.

Saleri, N. (1996): Reengineering Simulation: A Bottom-Line Approach to Managing Complexity and Complexification. SPE Annual Technical Conference and Exhibition, 6-9 October 1996, Denver, Colorado.

Spetzler, C.; Holstein, S. S. (1975): Probability Encoding in Decision Analysis. Management Science, Vol. 22(3), p. 340-358.

Staber, S. (2008): Decision Biases. OMV AG.

Sheskin, D. (2004): Handbook of parametric and nonparametric statistical procedures. Boca Raton: Chapman&Hall, 3rd edition.

N.N.: Microsoft Office Project Support. URL: http://office.microsoft.com/en-us/project-help. (Access: December 11, 2011).

N.N.: Microsoft Office Support: Use a PERT analysis to estimate task durations. URL: http://office.microsoft.com/en-us/project-help/use-a-pert-analysis-to-estimate-task-durations-HA010113082.aspx. (Access: March 11, 2011).

Surowiecki, J. (2005): The Wisdom of Crowds. New York: Anchor Books.

Taleb, N. (2007): The Black Swan. United States: Random House.

Tversky, A.; Kahneman, D. (1974): Judgment Under Uncertainty: Heuristics and Bias. Science(Vol. 185), p. 1124-1131.

Tversky, A.; Kahneman, D. (1979): Intuitive Prediction: Biases and corrective procedures. TIMS Studies in the Management Science, p. 313-327.

Tschoegl, J. A. (2007): Review of Philipp E. Tetlock, Expert political judgement: How good ist it? How can we know? International Journal of Forecasting (Volume 23), p. 339-342.

Vose, D. (2003): Risk Analysis. Chichester: John Wiley & Sons.

Wallace, I. (2010): Schedule risk and contingency using @RISK and probabilistic analysis. URL: www.palisade.com. (Access: April 23, 2011).

7 Glossary

7.1 Basic Statistic Terms

Continuous Distribution

A probability distribution where any value between the minimum and maximum is possible (see Figure 96 on the right).

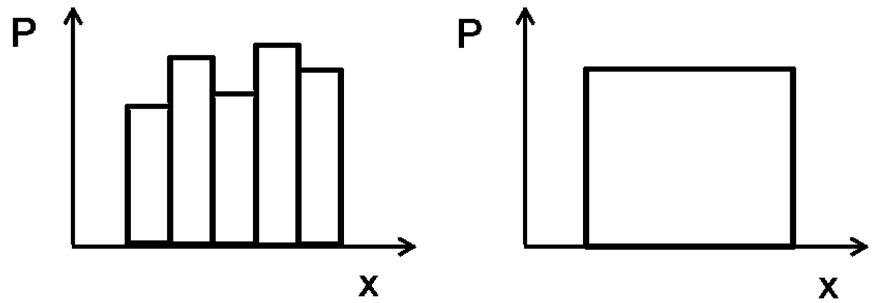


Figure 96: An example for a discrete (left) and a continuous (right) distribution.

Cumulative Density Distribution

A cumulative distribution (CDF) is the set of points, each of which equals the integral of a probability distribution, starting at the minimum value and ending at the associated value of the random variable.

Confidence Interval

Is used to indicate the reliability of an estimate. It is the probability that your estimated value lies between a defined upper and lower distribution value.

Convergence

Some functions and sequences approach a limit (value) under certain conditions.

Discrete Distribution

A probability distribution, where only a finite number of discrete values are possible between the minimum and maximum (see Figure 96 on the left).

Distribution Parameters

Define the characteristics of a distribution like location (central values like mean), spread (range, variance) and shape (skewness, kurtosis).

Frequency Distribution

Frequency distributions are constructed from data by arranging values into classes and representing the frequency of occurrence in any class by the height of bar. The frequency of occurrence corresponds to probability.

Kurtosis

Kurtosis is a measure for the degree of peakedness/flatness of the distribution (see Figure 97).

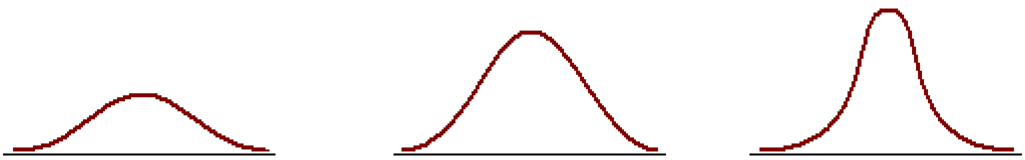


Figure 97: Kurtosis.[113]

Mean

Measure of central tendency of a distribution, the average value: sum all values and divide it by the number of values. Works well if the data on which it is based are more or less normally distributed. Skewed or multimodal distributions and the presence of extreme values distort the mean.
Note: you CAN ADD mean values!

Median

Measure of central tendency of a distribution, the midpoint of this distribution: the point above which and below which 50 % of the values lie. Works well if extreme values occur.

Mode

Measure of central tendency of a distribution, the most frequently occurring value. Works well if skewness or bi(multi)-modality occurs.

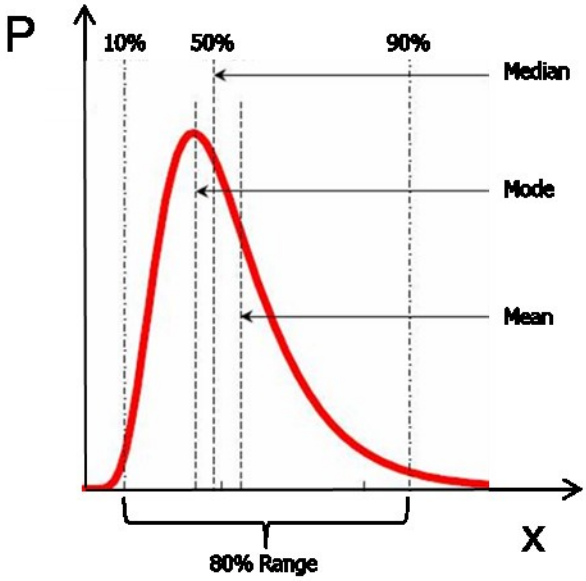


Figure 98: Various measures of central tendency of a positive skewed distribution. Note: if the distribution is normal, median, mode and mean are the same.

Percentiles (the "P levels"):[109]

Percentiles are values that divide your calculated project duration into 100 equal parts. The percentile rank is the proportion of values in a distribution that a specific value is lesser than or equal to. For example, if your presented project duration value is on a P95 (aka 95th

109 Cp. Akins (2005)

Percentile) level, than all possible project duration in future will be equal to that level or lesser with a probability of 95 per cent.
Note: DO NOT ADD percentiles!

Population

The set of all possible outcomes of the process being studied.[110]

Probability Density Distribution

A probability density function (PDF) is the proper statistical term for a frequency distribution constructed from an infinitely large set of values where the class size is infinitesimally small (see Figure 99).

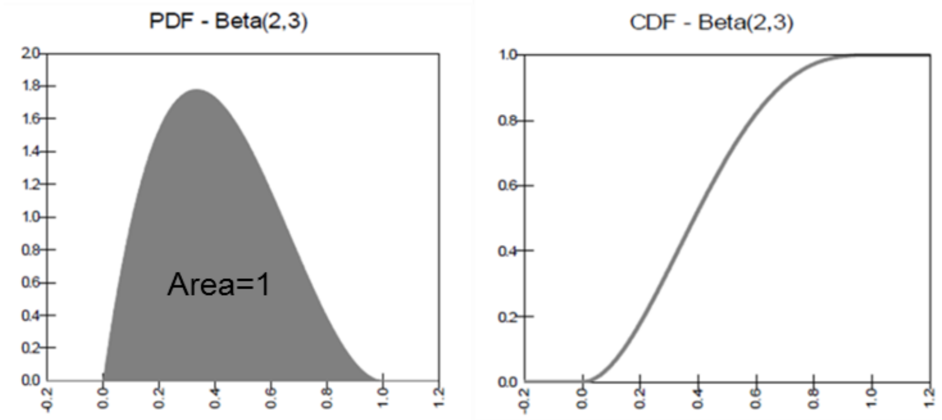


Figure 99: PDF and CDF of a Beta distribution.[111]

Random Variable (Stochastic Variable, Chance Variable)

A real number taken from a set of real numbers which have a specific probability distribution function. The random number is used to select a random event.[112]

Range

The range is the absolute difference between the maximum and minimum values in a set of values. The range is the simplest measure of the dispersion or "risk" of a distribution (see Figure 101): broader range means that more values are possible and therefore as example completion dates of a project cannot be limited to a certain area.

Sample

A subset of a population. Statistics are calculated from samples so you can make inferences or extrapolations from the sample to the population.

Skewness

Skewness is a measure of the shape of a distribution. Skewness indicates the degree of asymmetry in a distribution. Skewed distributions have more values to one side of the peak or most likely value, so one tail is much longer than the other. A skewness of 0 indicates a symmetric distribution, while a negative skewness means the distribution is skewed to the left. Positive skewness indicates a skew to the right (see Figure 100).

110 Cp. Evers et al. (1973)
111 N.N., Guide to Using @RISK (2010)
112 Cp. Evers et al. (1973)

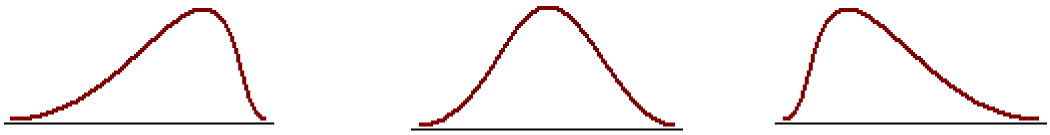


Figure 100: Skewness.[113]

Stratification

Stratification divides the cumulative curve into equal intervals on the cumulative probability scale.

Truncation

Truncation is the process by which a user chooses a minimum-maximum range for a random variable that differs from the range indicated by the distribution type of the variable. A truncated distribution has a smaller range than the non-truncated distribution, because the truncation minimum is bigger than the distribution minimum and/or the truncation maximum is smaller than the distribution maximum.

Variance

The variance is a measure of how widely dispersed values are and thus it is another indication for the "risk" content of the distribution. The variance gives disproportionate weight to "outliers", values that are far away from the mean. Technically variance is the square of the standard deviation (see Figure 101).
Note: DO NOT ADD variances/standard deviations!

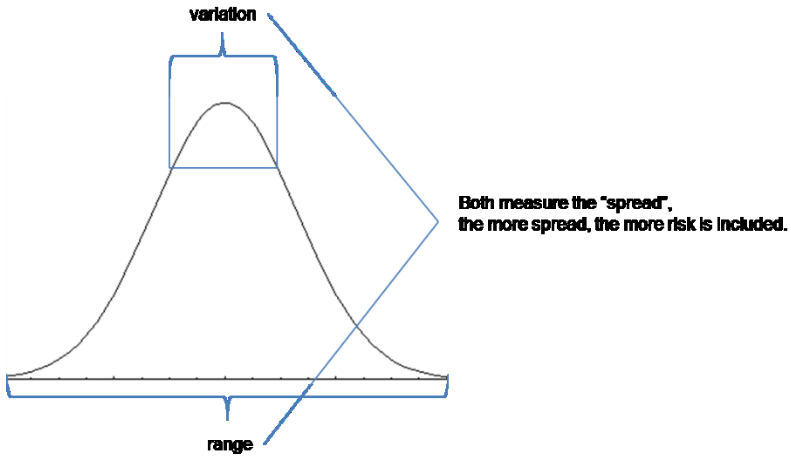


Figure 101: Variance and range.

[113] Source: MedCalc (2011)

7.2 Specific Project Scheduling Terms[114]

Constraint

A restriction on the start or finish date of a task. You can specify that a task must start on or finish no later than a particular date. Constraints can be flexible (not tied to a specific date) or inflexible (tied to a specific date).[115]

Critical Index

Probability that a task lies on the critical path.

Critical Path

“A series of critical tasks makes up a project's critical path.”[116]

Critical Tasks

Tasks that have to be completed on schedule for the project to finish on time. If such a critical task is delayed, the project completion date might also be delayed.[117]

Lag Time (positive Lag)

A delay between tasks that have a dependency. For example, if you need two-day space between the finish of the first task and the start of the second, you can establish a finish-to-start dependency and specify a two-day lag time.

Lead Time (negative Lag)

An overlap between tasks that have a dependency. For example, if second task can start when the first task is half-finished, you can specify a finish-to-start relation with a lead time of 50 % for the second task.

Lumping

In scheduling: Summarizing many sub-tasks to fewer top-tasks. Advantages could be simpler schedules, fewer tasks to estimate, less risk-driven schedules etc.
Other meaning: Simplifying experienced facts by personal prediction skills.

Risk

Loss or gain, i.e. a change in assets associated with some chance of occurrence.[118]

Slack (Float)

The amount of time that a task can slip before it has impact on another task (“free slack”) or the project's finish date (“total slack”).

Tornado Chart

A special type of Bar chart, where the data categories are listed vertically instead of the standard horizontal presentation. Categories are ordered so that the largest bar appears at the top of the chart, the second largest appears second from the top etc. The name comes from a tornado-like appearance.

114 Cp. Microsoft Support (2011)
115 Microsoft Support (2011)
116 Microsoft Support (2011)
117 Microsoft Support (2011)
118 Cp. Murtha (1997)