One Number to Correlate Them All

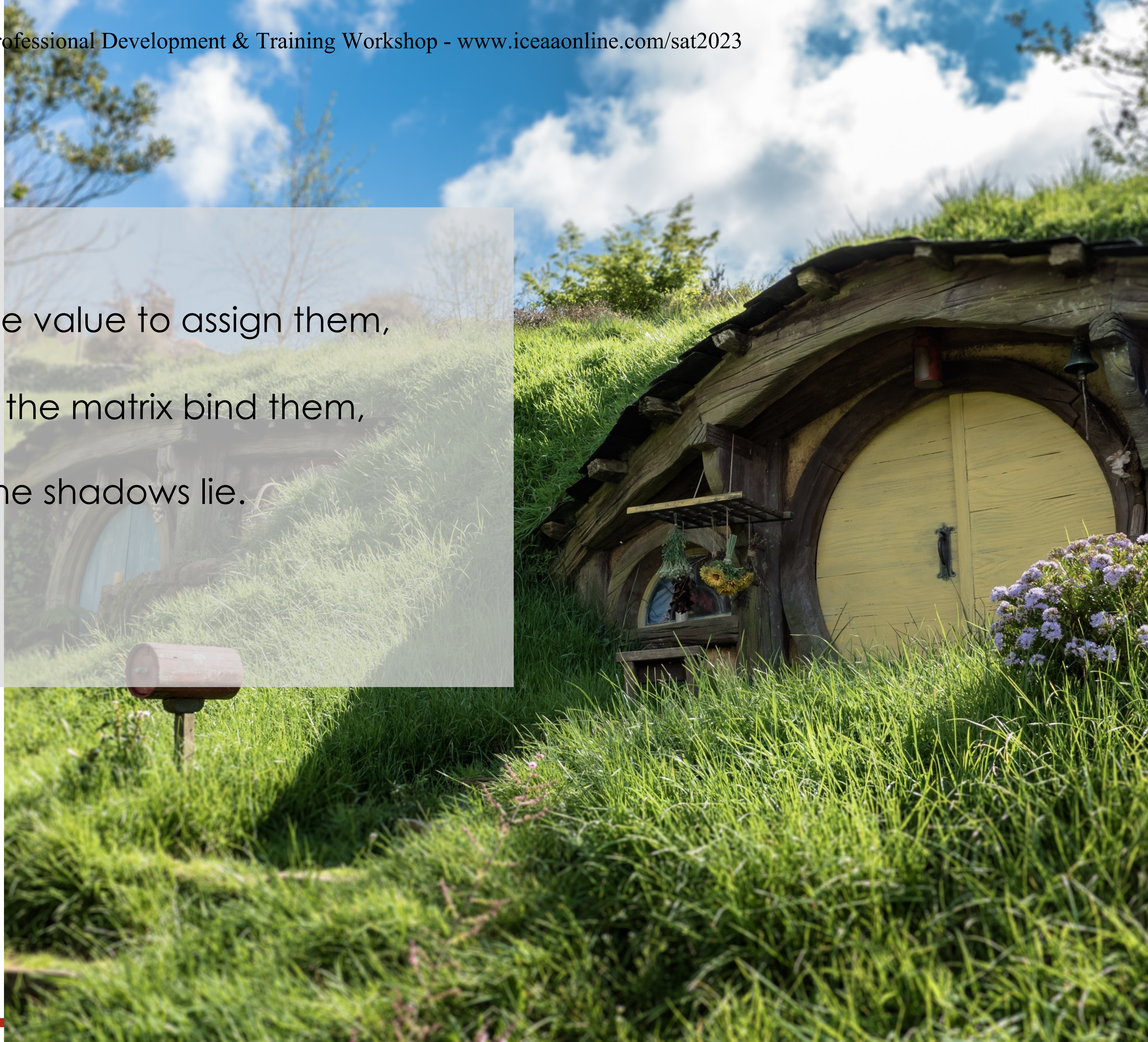
Christian Smart
Anh Harris
Karen McRitchie

One value to correlate them all, one value to assign them,

One value to bring them all, and in the matrix bind them,

In the Land of Uncertainty, where the shadows lie.

(Apologies to J.R.R. Tolkien)



What Will Be Covered

Importance of Correlation

Understand how correlation can impact an estimate across many WBS elements and what happens when it is ignored.

Specifying Correlation

In an estimate of many WBS elements, how best to specify correlation. The correlation matrix.

The Right Value

What value should be used for correlation?

Monte Carlo in SEER

Overview of how **the Monte Carlo sampling process** along with the concepts of correlation.

Correlation is a key consideration in cost and schedule risk analysis, as its exclusion causes significant underestimation of uncertainty. When assigning values in the absence of functional correlation, this can be accomplished by considering every WBS element. However, this can be time-consuming for a detailed estimate. In this presentation, we discuss an alternative method that uses a single value, which offers significant time savings, and discuss its implementation in the SEER® model suite.

Definitions

- Consider two random variables, *X* and *Y*.
- The mean of X, E(X), is denoted by μ_x, and similarly, the mean of Y, E(Y), is denoted by μ_y
- The variance of X, Var(X), is denoted by σ_X^2, and similarly, the variance of Y, Var(Y), is denoted by σ_Y^2
- The variance of X and Y are equal to:

$$Var(X) = Cov(X, X) = E(X^2) - [E(X)]^2$$
$$Var(Y) = Cov(Y, Y) = E(Y^2) - [E(Y)]^2$$

- Correlation, denoted by the Greek letter *r* ("rho"), is defined by

$$\rho_{XY} = Corr(X,Y) = \frac{cov(X,Y)}{\sqrt{Var(X)Var(Y)}} = \frac{E(XY) - E(X)E(Y)}{\sqrt{Var(X)Var(Y)}} = \frac{E(XY) - \mu_X \mu_Y}{\sigma_X \sigma_Y}$$

Total System Mean and Variance

- For *n* WBS elements, the mean and the variance of the total cost are defined by:

$$E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} \mu_i$$

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=}^{n} \sigma_i^2 + 2\sum_{j=2}^{n}\sum_{i=1}^{j-1} \rho_{ij}\sigma_i\sigma_j$$

Specifying Correlation – One vs Many

- Here is a notional correlation matrix for a set of "n" WBS elements
- You need to specify correlation for each unique pair
- And each element has a 1.0 correlation with itself
- So some basic math indicates you need to specify this many values:

$$\frac{n^2}{2} - \frac{n}{2} = \frac{n(n-1)}{2}$$

	1	2	3	…	…	n-2	n-1	n
1	1	20%	40%	40%	23%	30%	20%	10%
2		1	30%	50%	5%	50%	40%	10%
3			1	80%	10%	25%	10%	60%
⋮				1	33%	90%	60%	10%
⋮					1	2%	15%	75%
n-2						1	40%	42%
n-1							1	12%
n								1

Correlation Challenge

Not practical as n gets large, or even moderately large

Who has the time and knowledge?

Specifying a correlation matrix for even a moderate number of WBS elements, it's a lot of data entry and will likely have to involve different SMEs. It could easily turn into a herding cats situation.

Data Handling Challenges

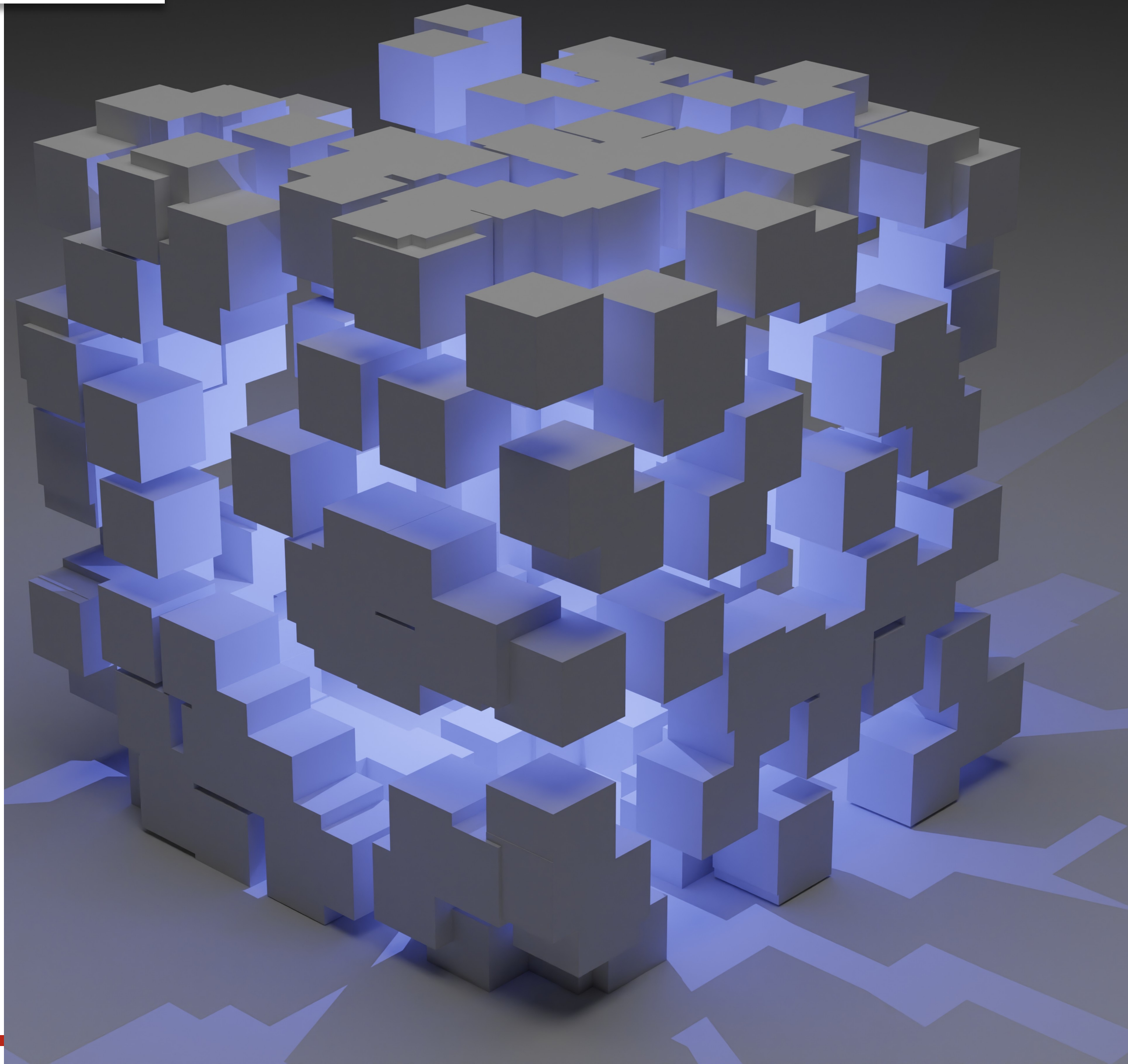
Even if you can get the data handling large matrices of data can be a challenge to document, clumsy to manipulate and can be prone to data entry errors.

Recommendation – Specify one value for all

A single value for all elements is recommended. The question is, what value should you use?

What Value to Use?

- Looking for a single value that results in the least amount of error in the variance
- A value that minimizes the amount by which the total standard deviation is misestimated due to the correlation assumption
- A value that would minimize the error when the assumed correlation differs from the actual underlying correlation



Total Variance with a Single Correlation

- **Suppose (for simplicity)**
  - **There are *n* WBS Elements:** $C_1,\ C_2,\ \ldots,\ C_n$
  - **Where:** $Var(C_i) = \sigma^2$ $Corr(C_i, C_j) = \rho < 1$
  - **Total Cost** $C = \sum_{k=1}^{n} C_i$

$$Var(C) = \sum_{k=1}^{n} Var(C_i) + 2\rho \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \sqrt{Var(C_i)\ Var(C_j)}$$

$$= n\sigma^2 + n(n-1)\rho\sigma^2$$

$$= n\sigma^2(1+(n-1)\rho)$$

Correlation	0	ρ	1
$Var(C)$	$n\sigma^2$	$n\sigma^2(1+(n-1)\rho)$	$n^2\sigma^2$

GALORATH

What if exercises

Equipped with the tools, we can look what happens in different scenarios

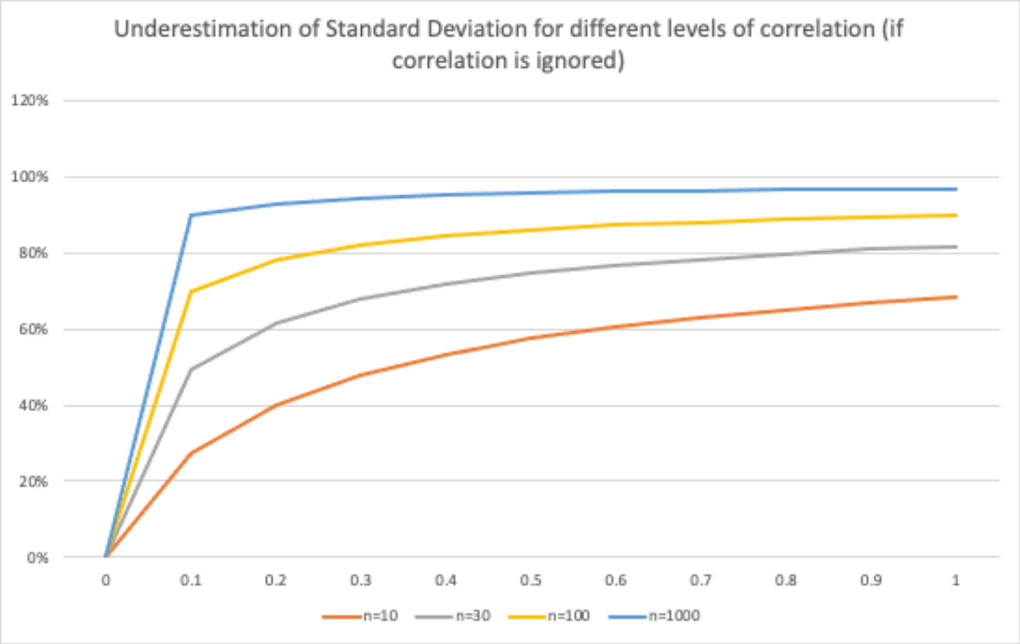
- **What happens when correlation is ignored and it is actually some non-zero value?**
- **What happens if you make the wrong assumption for correlation?**
- **Can we find a default correlation that minimizes error of the variance?**

Computing the Impact of Ignoring Correlation

Correlation	**0**	**ρ**	**1**
$Var(C)$	$n\sigma^2$	$n\sigma^2(1+(n-1)\rho)$	$n^2\sigma^2$

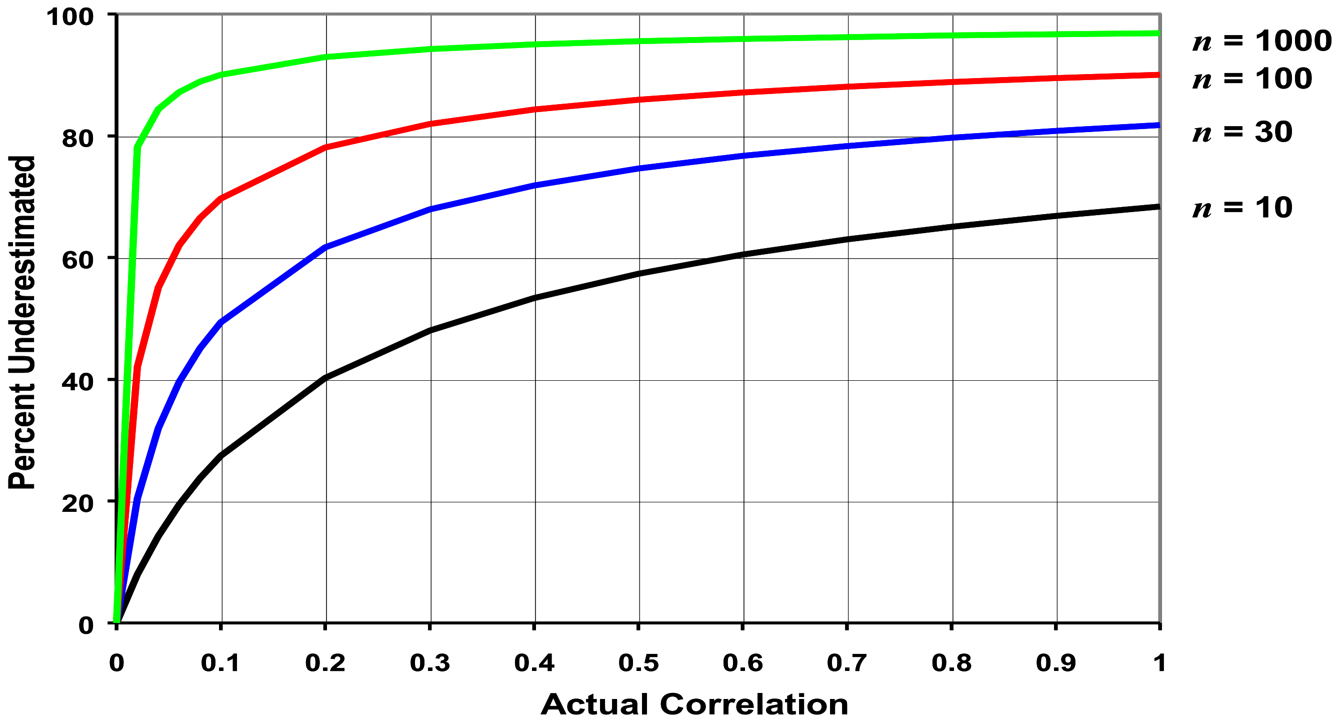
	Multiplier of Variance for "n" elements for different levels of correlation.. (all items have the same variance = σ^2).										
Correlation	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
n											
10	10	19	28	37	46	55	64	73	82	91	100
30	30	117	204	291	378	465	552	639	726	813	900
100	100	1090	2080	3070	4060	5050	6040	7030	8020	[illegible]	[illegible]
1000	1000	100900	200800	300700	400600	500500	600400	700300	800200	[illegible]	[illegible]

Correlation	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
		Underestimation of the standard deviation							
n	ABS(SQRT(EstimateVar)/SQRT(ActVar@0)-1)								
10	0%	27%	40%	48%	53%	57%	60%	63%	65%
30	0%	49%	62%	68%	72%	75%	77%	78%	80%
100	0%	70%	78%	82%	84%	86%	87%	88%	89%
1000	0%	90%	93%	94%	95%	96%	96%	96%	96%



The importance of correlation

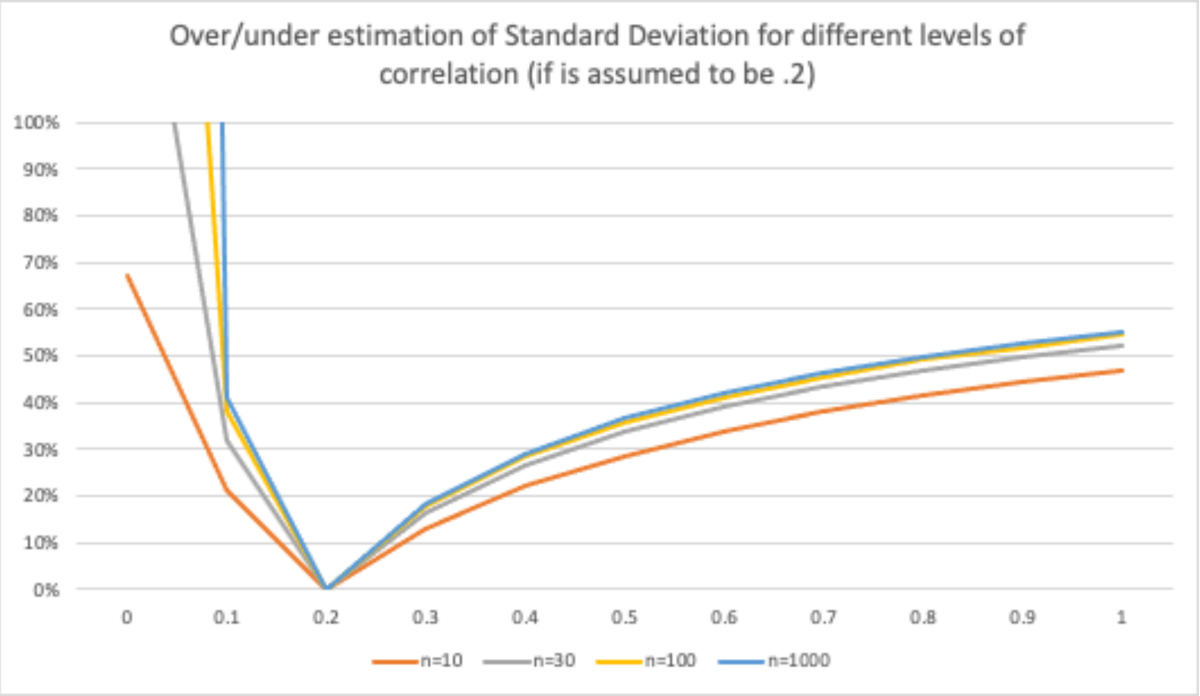
- Not considering correlation will lead to an understatement of the overall **standard deviation**
- N = number of WBS elements
- X-axis is the actual correlation
- Plot is the underestimation of standard deviation



Computing the impact of using the wrong correlation

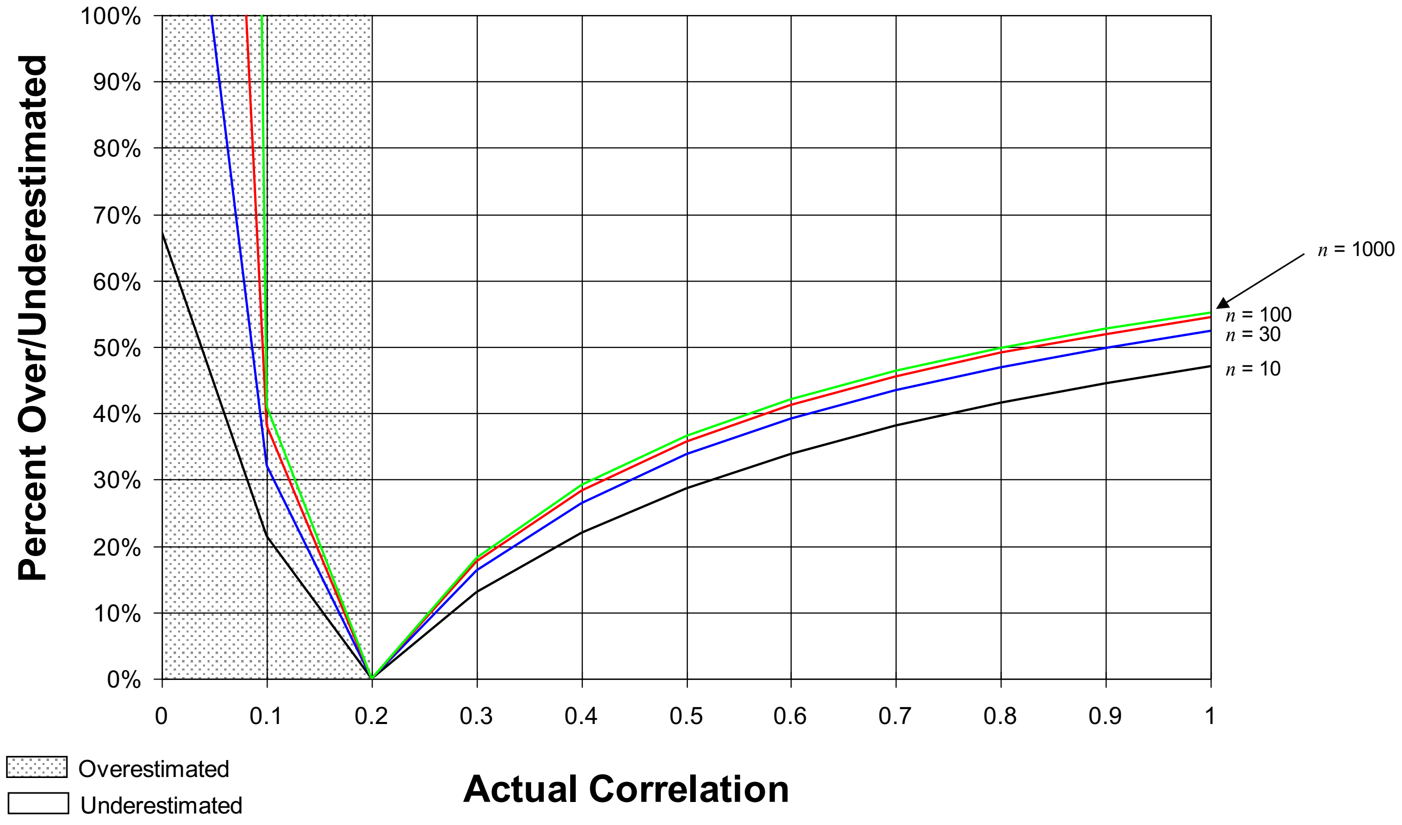
Correlation	0	ρ	1
$Var(C)$	$n\sigma^2$	$n\sigma^2(1+(n-1)\rho)$	$n^2\sigma^2$

	Multiplier of Variance for "n" elements for different levels of correlation.. (all items have the same variance = σ^2).									
Correlation	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
n											
10	10	19	28	37	46	55	64	73	[illegible]	[illegible]	[illegible]
30	30	117	204	291	378	465	552	639			
100	100	1090	2080	3070	4060	5050	6040	7030			
1000	1000	100900	200800	300700	400600	500500	600400	700300	[illegible]		
Correlation	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7			
	Over Estimate			Under Estimate							
n	ABS(SQRT(EstimateVar)/SQRT(ActVar@.2)-1)										
10	67%	21%	0%	13%	22%	29%	34%	38%			
30	161%	32%	0%	16%	27%	34%	39%	43%			
100	356%	38%	0%	18%	28%	36%	41%	46%			
1000	1317%	41%	0%	18%	29%	37%	42%	46%			



Another view – what if you did not assume sufficient correlation

- This chart assumes 20% correlation and shows the over/under estimate of standard deviation if the actual correlation is different
- Even with 20% correlation assumption, standard deviation can be underestimated



Look at minimizing error in variance

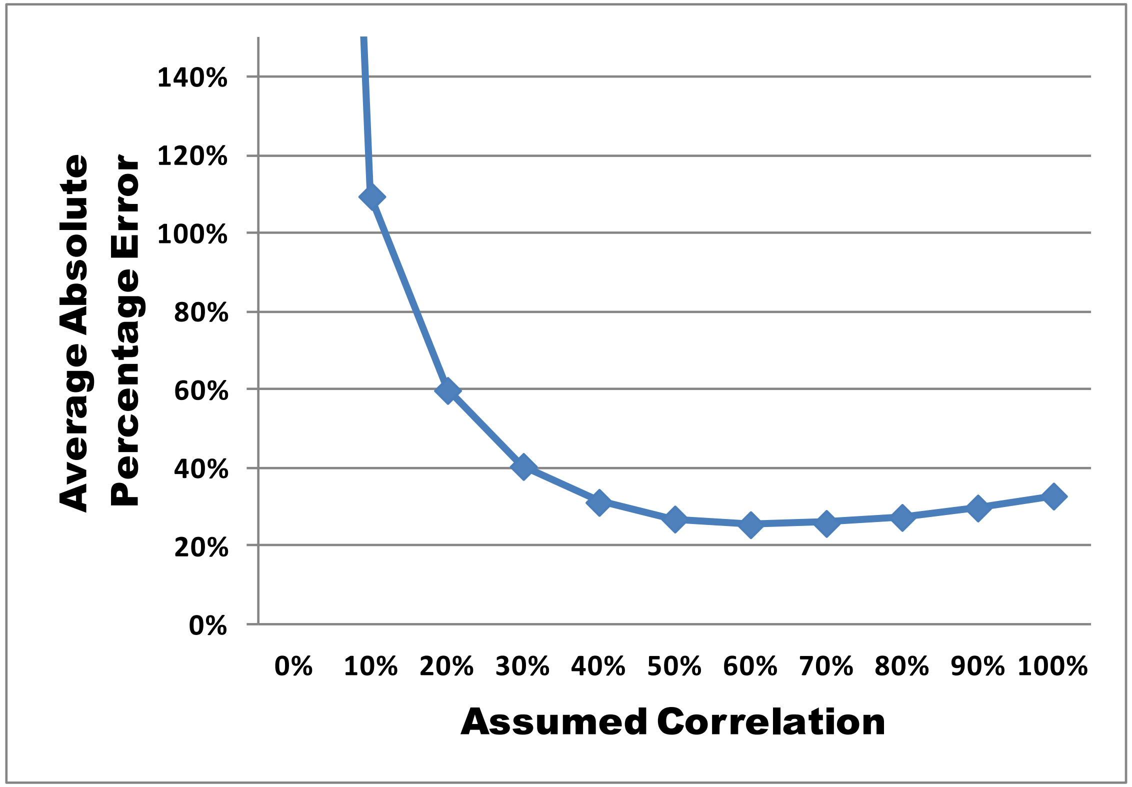
- In a 2013 ICEAA paper (Robust Default Correlation for Cost Risk Analysis) Christian Smart proposes four approaches to determining a correlation
- Of these approaches, a default correlation of ~60% should be used.

Summary of the Four Cases

- **All four cases minimize the expected value of the absolute error in the variance, but use different metrics for measuring error**
- **Case 1:**
  - **Error is measured as a percentage of the variance that results from the actual correlation, result in the limit is 25%**
- **Case 2:**
  - **Error is measured as a percentage of the variance that results from the assumed correlation, result in the limit is 63%**
- **Case 3:**
  - **Error is measured as total difference in variances, result is 50%**
- **Case 4:**
  - **Error is measured as a percentage of the variance that results from the actual correlation, with the correlation range limited to 10-90%; result is 40%**

https://www.iceaaonline.com/2013-r209/

Recommended Correlation



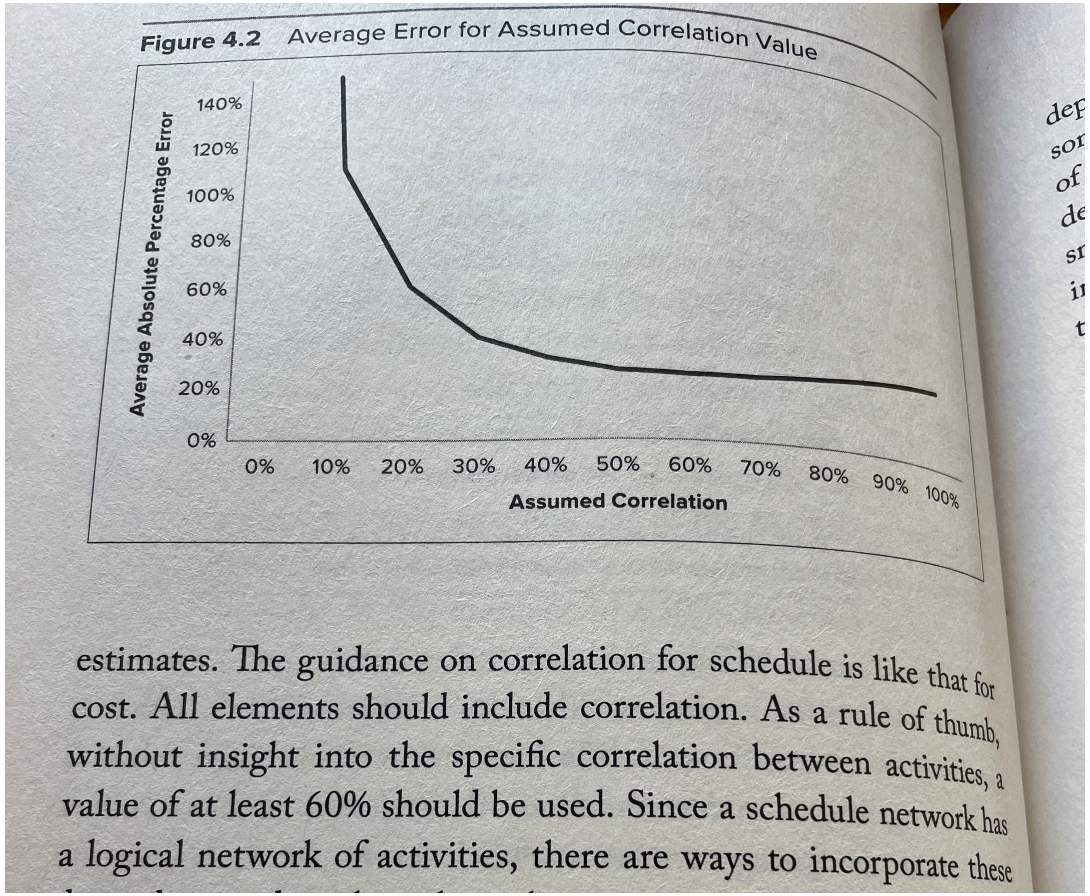


Figure 4.2 Average Error for Assumed Correlation Value

estimates. The guidance on correlation for schedule is like that for cost. All elements should include correlation. As a rule of thumb, without insight into the specific correlation between activities, a value of at least 60% should be used. Since a schedule network has a logical network of activities, there are ways to incorporate these

GALORATH

Recap....

Thus far we have established that

The use of correlation is Important in risk analysis as it will impact the variance about the mean.

A single value for ρ is sufficient and accessible

The single value should not be 0% and 60% is a good value to use.

Now, we can discuss how this correlation is applied in a SEER based risk analysis, specifically in the Monte Carlo process

Monte Carlo Basics

Key Elements of a Monte Carlo Process

Variable Being Sampled

Determine what is being sampled. This is usually cost our hours.

Model Input Ranges

Determine the inputs to the model and how they are distributed. SEER uses a range of inputs (least, likely and most) and applies a modified pert distribution to capture the range. A probability input will determine the value in input range that will be used to compute cost.

Sampling

Generate a sampling of cost using random inputs for probability. This will involve generating a set of random numbers between 1% - 99% and using those to compute an estimate. Each calculation is an "iteration." You must iterate a sufficient number of times to obtain a statistically significant result.

Compute Statistics

Computee descriptive statistics. Using the sample set generated, SEER computes a mean, standard deviation, and percentiles. This helps to quantify the understanding of the estimate range.

Key Points About Risk in SEER

1 Calculations are deterministic

That means given the same inputs, the same output is computed. Uncertainty in estimates is driven by input ranges.

2 Parameter Ranges Drive Output Range

All uncertainty is reflected in the input ranges. Flat parameter inputs (Least = Likely = Most) yield a point estimate

3 Kbases Offer Parameter Ranges

The variability in the data which is used to create knowledge bases has been captured in parameter value ranges. However key size inputs such as SLOC, weight, PCBs are not set by kbases.

4 Input Distributions are Generated

Least, Likely and Most inputs for a parameter are used to create a range and the probability input drives what value in the range is used for calculation.

Correlation

What does this mean in context of SEER Monte Carlo?

Relation Among WBS Elements

Relation among random variables

In SEER, the correlation relates to how related the WBS elements are to one another. If risks impact all elements in the same way, they would be correlated. If the outcome of one element doesn't impact others, it is uncorrelated.

Correlation drives the random probability draw

In Monte Carlo sampling correlation relates to the random probability used in each iteration. If elements are correlated, they use the same probability. If not, they use different probabilities.

Correlation

SEER Supports elements being correlation of 0% - 100% with 60% as a recommended default.

Monte Carlo

Quick Takes

Sampling of Development Hours					
Iteration	Element A	Element B	Element C		Sum
1	3,083	4,462	6,415		13,959
2	1,418	2,415	3,615		7,448
3	2,581	3,858	5,597		12,035
4	5,294	7,164	9,683		22,141
5	3,150	4,542	6,522		14,214
:	:	:	:		:
:	:	:	:		:
:	:	:	:		:
:	:	:	:		:
996	1,443	8,438	8,289		18,170
997	920	4,647	14,416		19,983
998	2,572	5,255	14,003		21,830
999	2,636	6,411	8,021		17,069
1000	9,198	4,466	12,407		26,071
Statistics Computed From The Sampling					
Mean	4,549	6,123	8,356		19,028
Median	3,741	5,069	7,300	=/=	17,617
90% Percentile	8,635	10,615	13,692	=/=	29,766

Sampling technique

Random draw picks a probability to be used to determine the value in the input range. Cost/hours/schedule is then computed using this random probability.

Iterations

Calculations are repeated for a specified number of iterations. Iterations can range from the 100s to 1000s. The more iterations, the greater the accuracy.

Distribution is created from the sampling

Given a set of 100 or more calculations, you can compute a mean, standard deviation, and estimate percentiles.

Random Probability Draws

For all or nothing correlation

Assume 100% Correlation

Random Probablity Draw - 100% Correlation			
Iteration	Element A	Element B	Element C
1	37.91%	37.91%	37.91%
2	5.01%	5.01%	5.01%
3	26.25%	26.25%	26.25%
4	70.09%	70.09%	70.09%
5	39.46%	39.46%	39.46%
:	:	:	:
:	:	:	:
996	6.30%	6.30%	6.30%
997	49.53%	49.53%	49.53%
998	50.19%	50.19%	50.19%
999	45.10%	45.10%	45.10%
1000	32.87%	32.87%	32.87%

Fully Correlated or 100% Correlation. Random probability used is the same for all elements for each iteration.

Assume 0% Correlation

Random Probability Draw - 0% Correlation			
Iteration	Element A	Element B	Element C
1	43.91%	34.41%	78.82%
2	27.95%	42.77%	16.43%
3	50.05%	74.93%	92.94%
4	54.50%	78.86%	16.93%
5	93.56%	76.43%	74.72%
:	:	:	:
:	:	:	:
996	94.49%	88.31%	72.76%
997	83.49%	71.75%	65.73%
998	59.40%	8.60%	22.09%
999	94.06%	55.05%	26.20%
1000	92.58%	74.04%	34.38%

Fully Uncorrelated or 0% Correlation. Each element uses a unique set of random probabilities for each iteration.

GALORATH

60% Correlation

Iterations are grouped
Into correlated and uncorrelated.

Correlated Iterations
Use the same probability for all elements.

Uncorrelated Iterations
Use unique probability for all elements.

60% of iterations

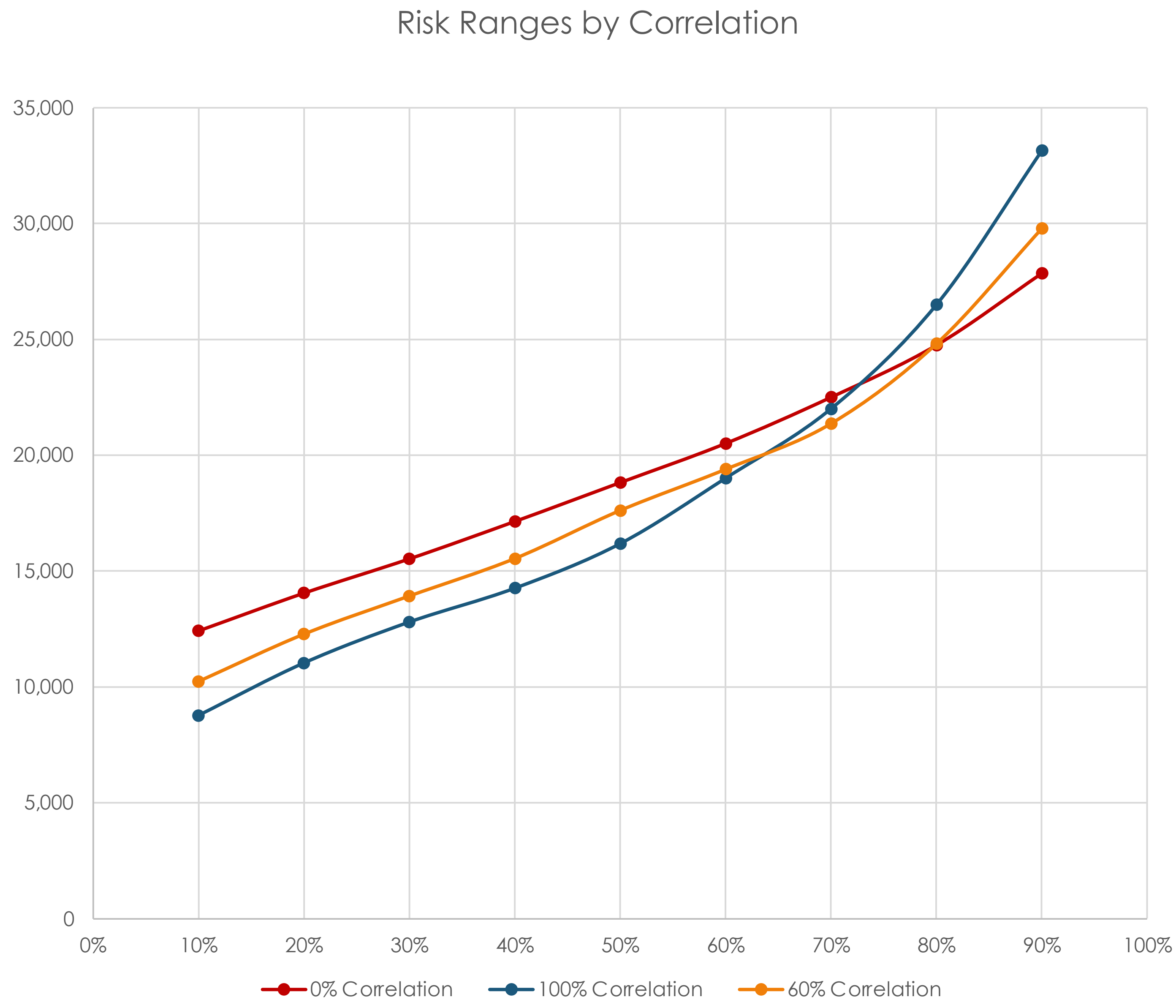
Random Probability Draw – 60% Correlation			
Iteration	Element A	Element B	Element C
1	37.91%	37.91%	37.91%
2	5.01%	5.01%	5.01%
3	26.25%	26.25%	26.25%
4	70.09%	70.09%	70.09%
5	39.46%	39.46%	39.46%
:	:	:	:
:	:	:	:
599	51.55%	51.55%	51.55%
600	51.81%	51.81%	51.81%
601	89.49%	37.26%	36.50%
602	94.35%	77.86%	30.99%
:	:	:	:
:	:	:	:
996	5.30%	78.44%	59.43%
997	1.14%	41.49%	90.07%
998	26.05%	51.81%	89.05%
999	27.53%	63.79%	57.08%
1000	91.11%	38.00%	84.04%

Estimate Ranges

Comparing different correlation

The fully correlated case has the widest range of estimates.

60% correlation falls (mostly) between the 0% and 100% cases.



Monte Carlo Report

Reports can be copied or printed

Estimates by Confidence Levels

Hours, Cost and/or Schedule are shown by **confidence level**. The confidence level can be interpreted as the probability that the actual result will be at or below the estimate. Think of it as a probability of not exceeding the estimate. (SEER-H has an option to provide Base Year vs Then Year cost)

Summary Statistics & Assumptions

The **mean** and **standard deviation** for the estimate. The coefficient of variation (CV) can be computed as the StdDev/Mean which is a relative measure of the spread. Assumptions on iterations and correlation are also included.

WBS Allocation by Confidence Level

Selectable confidence level can be used to show how a rollup estimate will flow down into the elements. This is applicable at the rollup level only.

Monte Carlo Development Risk

Display outputs in Base Year Cost

Confidence Level	*Dev Cost*	*Dev Hours*
10%	5,881,911	29,336
20%	6,771,014	33,766
30%	7,423,481	37,016
40%	8,044,484	40,118
50%	8,662,516	43,199
60%	9,674,954	48,243
70%	11,037,077	54,984
80%	12,565,811	62,667
90%	15,518,125	77,313
Mean	9,792,672	48,812
StdDev	3,896,260	19,389

Currency: US Dollars
Exchange Rate: 1.0000

MONTE CARLO ASSUMPTIONS
Based on 1,000 iteration sampling
Parameters Fully Correlated
Work Elements 60% Correlated

'ork Element Allocation Of Cost and Hours at 50%

10%
20%
30%
40%
✓ 50%
60%
70%
80%
90%

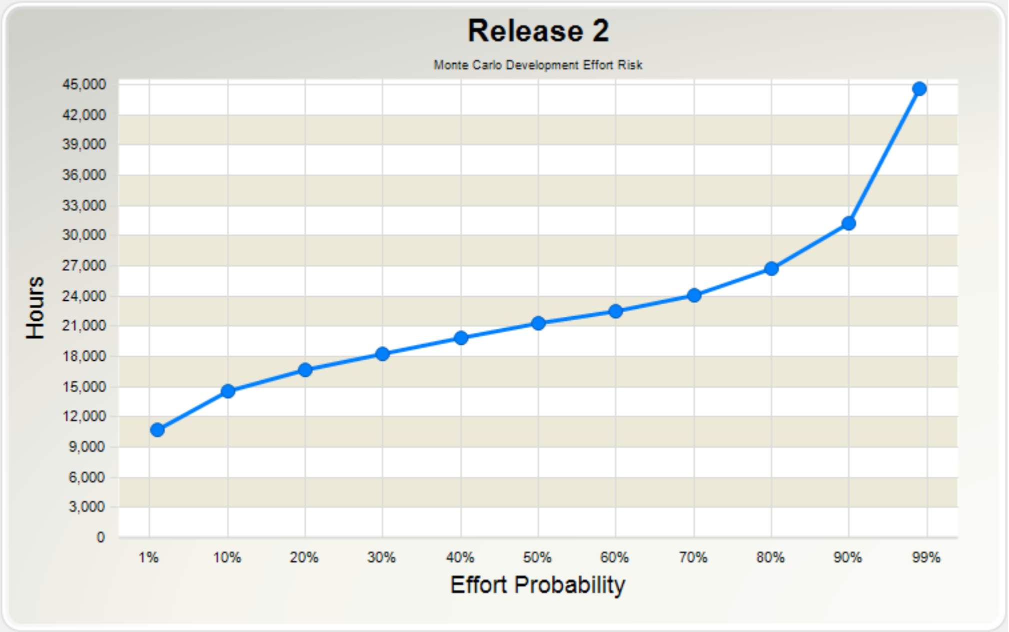
	Dev Cost	*% of Total*	*StdDev*
1.2: Equipment Configuration	8,662,516	100%	3,896,260
1.2.1: Board, Receiver Module	5,552,388		3,435,088
1.2.1.1: Board, Receiver	523,326	6%	165,214
1.2.1.2: RF Module	4,881,917	56%	3,316,513
1.2.1.3: RF Machined Housing	45,887	1%	11,401
1.2.1.4: Rcv Chassis	101,257	1%	27,918
1.2.2: Digital Processing	2,977,465		576,473
1.2.2.1: Converter & Noise Reduction	473,497	5%	155,545

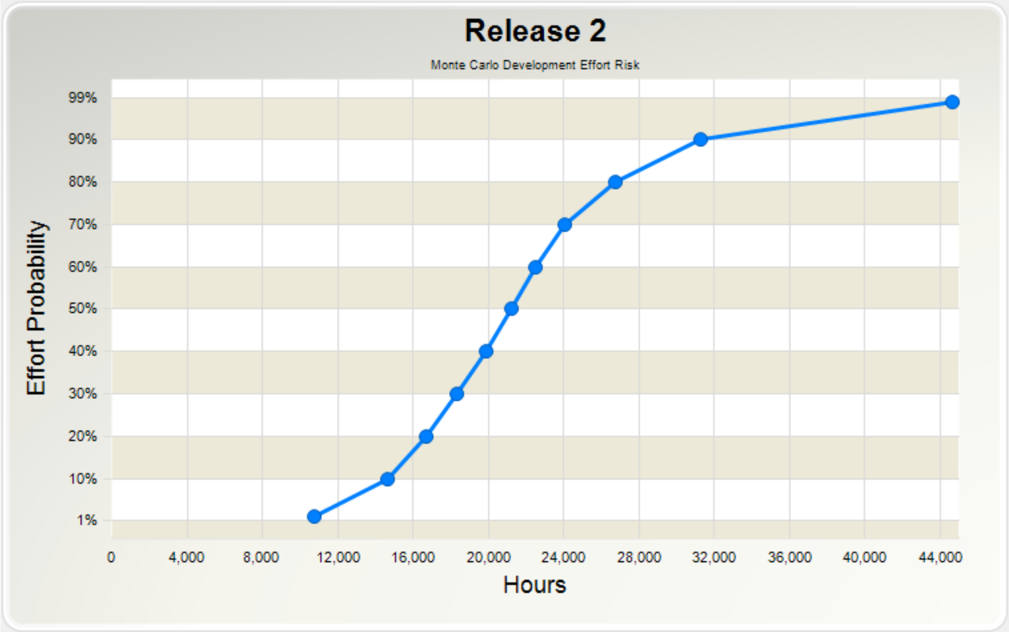
G A L O R A T H

Monte Carlo Charts

Visually Depict Risk

Risk Charts will use the Monte Carlo data when MC is enabled. There are different charts for cost, effort and schedule. Charts can be shown as an S-curve by clicking >> in the lower left corner



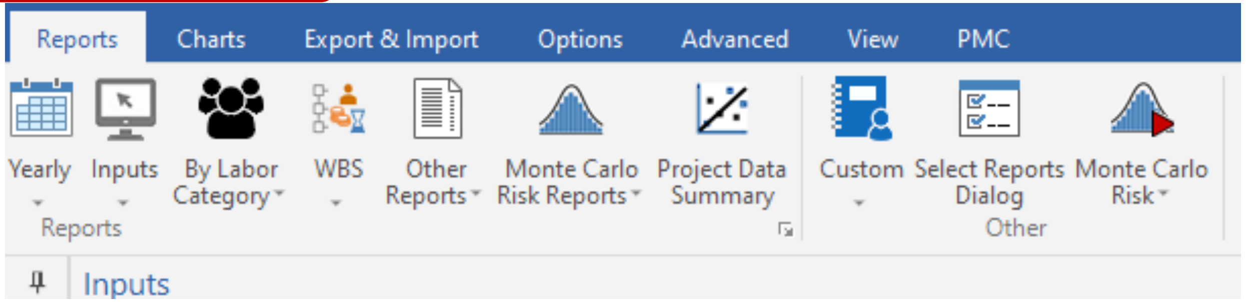


GALORATH

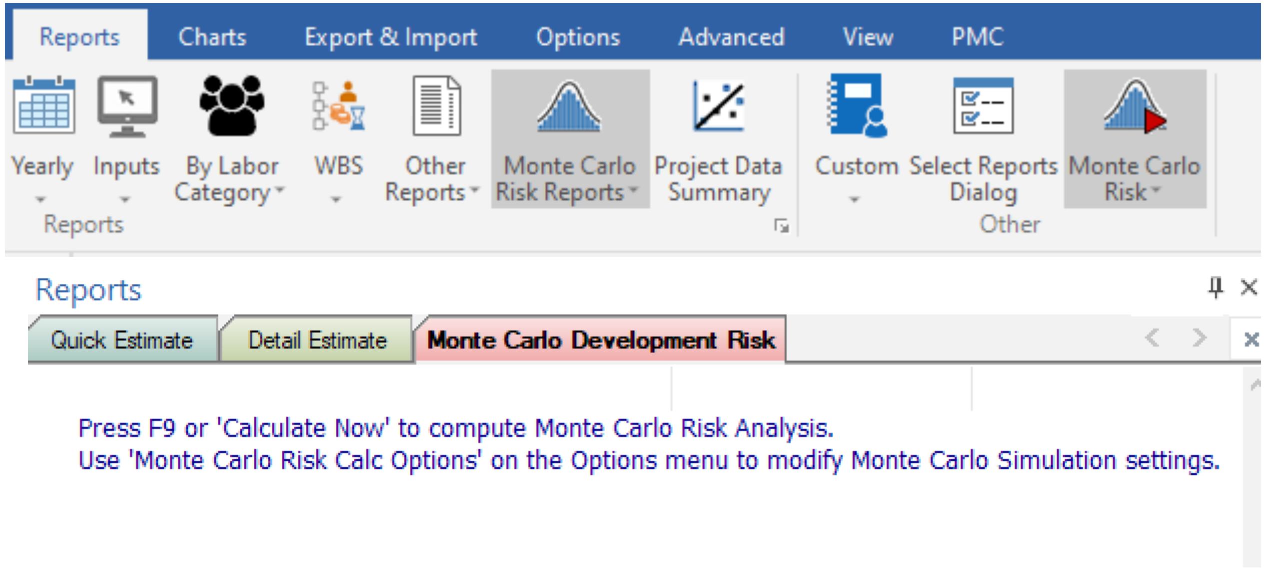
SEER Mechanics

Turn on a Monte Carlo Report or Chart to Activate

Not Enabled



Enabled

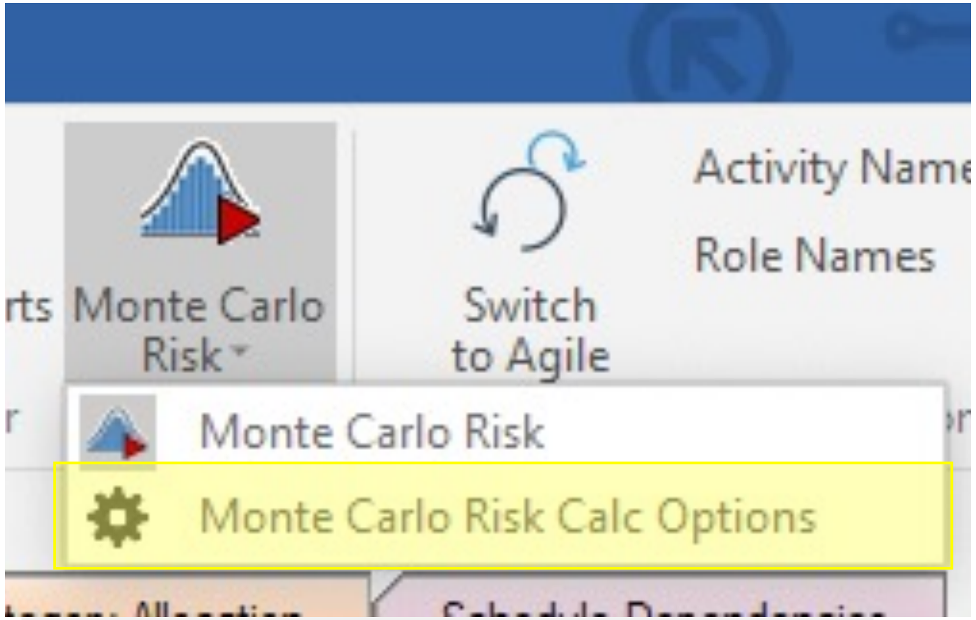


In the old days you would have to jump hoops to enable the Monte Carlo. This is because it took time to compute and you didn't always want to trigger it. Now the sampling and sorting is much faster, so we have made it more accessible.

When enabled you will see a risk report and/or a chart in the output area. You sometimes need to press F9 to trigger a recalc

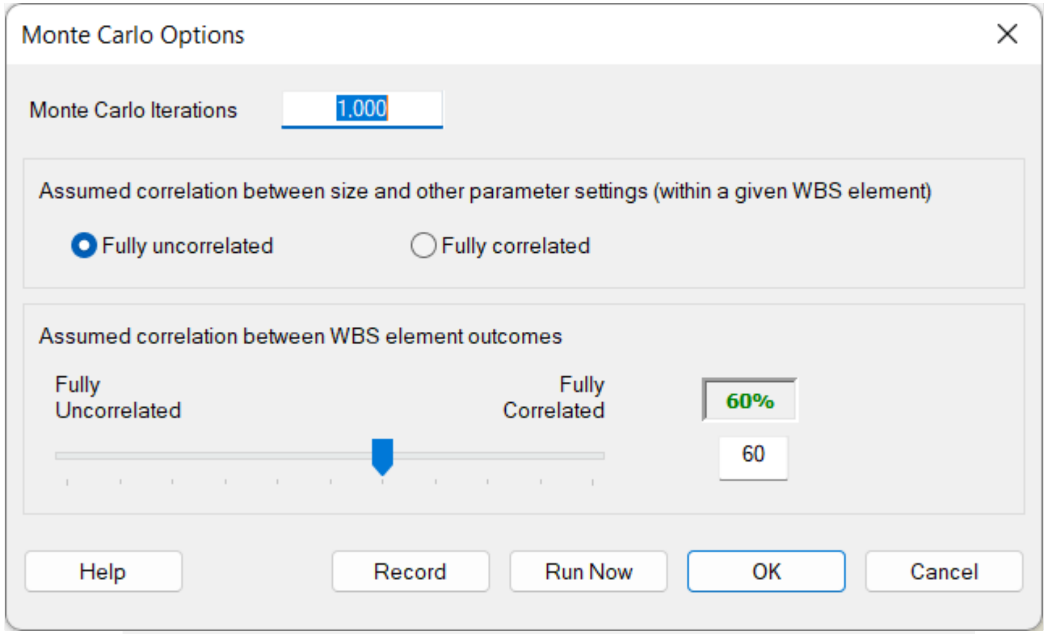
Monte Carlo Options

SEER for Software



Accessing Options

Select Monte Carlo Risk Options from the drop down on the Monte Carlo Risk button. Found on the report, chart or options ribbon.

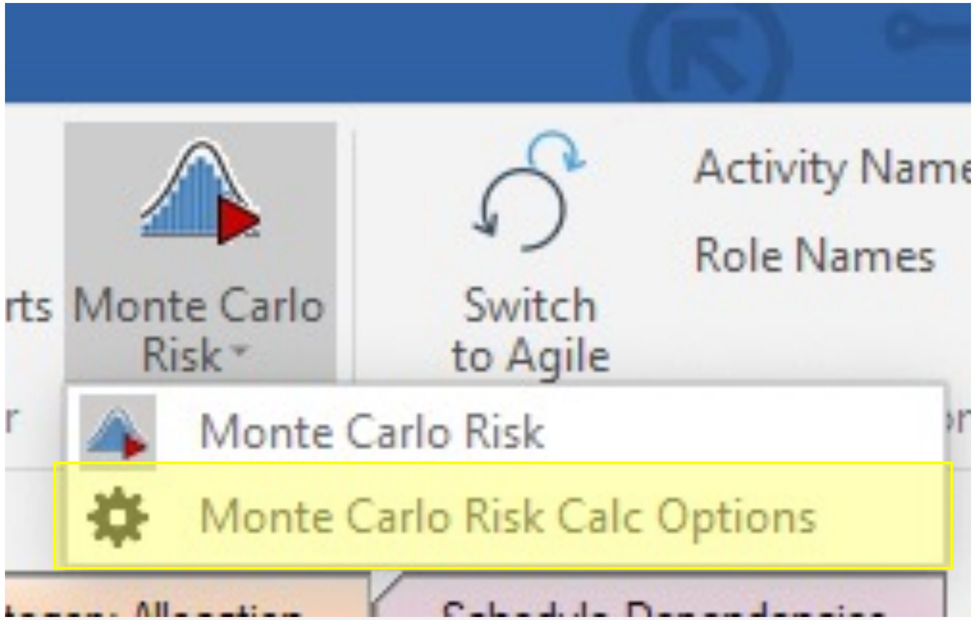


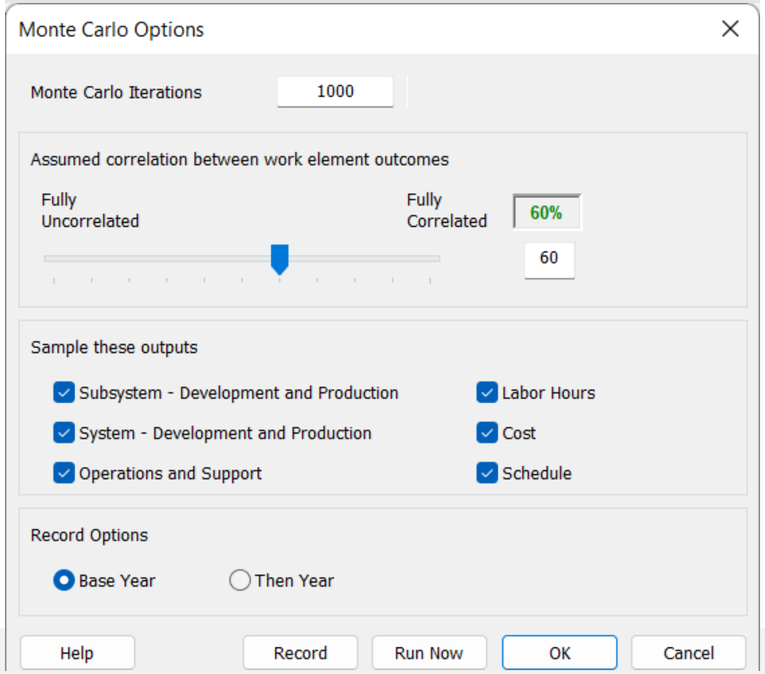
Setting Options

Iterations can be set between 100 – 10000. Specify **correlation** between WBS elements. SEER-SEM has a unique option specifying correlation between size and other parameters.

Monte Carlo Options

SEER for Hardware





Accessing Options

Select Monte Carlo Risk Options from the drop down on the Monte Carlo Risk button. Found on the report, chart or options ribbon.

Setting Options

Iterations can be set between 100 – 10000. Specify **correlation** between WBS elements. Select which outputs to include in the sampling. Choose between BY and TY for cost outputs.

Record Option

For the analytical DIYers

Create a record of the sampling

The record option will create a text file with the calculation for each iteration. This includes the random probability draw. This data can be used to check the math and help understand nature of the Monte Carlo sampling.

There is one row for each iteration that includes the probably used and the computed outputs. This example includes Dev Hours, Dev Cost, and Dev Schedule.

File Edit View

```
1/31/2023  4:54:12 PM   : WBS Description: Redstone Scheduler; Element Type: Program
1/31/2023  4:54:12 PM   : Parameters Fully Correlated
1/31/2023  4:54:12 PM   : WBS Elements 60% Correlated
1/31/2023  4:54:12 PM   : Iteration:    ProbSize:   ProbTech:   DevHrs:     DevCost:    DevSched:   MaintHrs:   MaintCost:
1/31/2023  4:54:12 PM   : 1   37.910000   37.910000   3082.656547 423865.275256    9.899339   0.000000    0.000000
1/31/2023  4:54:12 PM   : 2   5.010000    5.010000    1417.788744 194945.952292    7.610876   0.000000    0.000000
1/31/2023  4:54:12 PM   : 3   26.250000   26.250000   2580.619999 354835.249917    9.320852   0.000000    0.000000
1/31/2023  4:54:12 PM   : 4   70.090000   70.090000   5294.006532 727925.898136    11.941726  0.000000    0.000000
1/31/2023  4:54:12 PM   : 5   39.460000   39.460000   3149.621971 433073.021004    9.971682   0.000000    0.000000
1/31/2023  4:54:12 PM   : 6   77.060000   77.060000   6139.402107 844167.789694    12.576650  0.000000    0.000000
1/31/2023  4:54:12 PM   : 7   18.880000   18.880000   2246.594730 308906.775351    8.893484   0.000000    0.000000
1/31/2023  4:54:12 PM   : 8   11.120000   11.120000   1843.791549 253521.338055    8.318173   0.000000    0.000000
1/31/2023  4:54:12 PM   : 9   97.560000   97.560000   13630.73142 1874225.571505   16.608972  0.000000    0.000000
1/31/2023  4:54:12 PM   : 10  70.930000   70.930000   5384.950553 740430.701011    12.013089  0.000000    0.000000
1/31/2023  4:54:12 PM   : 11  97.230000   97.230000   13178.79328 1812084.076593   16.415284  0.000000    0.000000
1/31/2023  4:54:12 PM   : 12  58.890000   58.890000   4269.268460 587024.413197    11.075443  0.000000    0.000000
1/31/2023  4:54:12 PM   : 13  69.950000   69.950000   5279.089486 725874.804308    11.929944  0.000000    0.000000
1/31/2023  4:54:12 PM   : 14  50.350000   50.350000   3644.779499 501157.181171    10.478293  0.000000    0.000000
1/31/2023  4:54:12 PM   : 15  32.200000   32.200000   2837.718040 390186.230455    9.625546   0.000000    0.000000
1/31/2023  4:54:12 PM   : 16  83.190000   83.190000   7131.488826 980579.713585    13.252301  0.000000    0.000000
1/31/2023  4:54:12 PM   : 17  6.540000    6.540000    1542.662875 212116.145327    7.831251   0.000000    0.000000
1/31/2023  4:54:12 PM   : 18  54.570000   54.570000   3941.007011 541888.464075    10.769284  0.000000    0.000000
1/31/2023  4:54:12 PM   : 19  70.180000   70.180000   5303.633539 729249.611654    11.949319  0.000000    0.000000
1/31/2023  4:54:12 PM   : 20  65.610000   65.610000   4846.939951 666454.243286    11.578689  0.000000    0.000000
1/31/2023  4:54:12 PM   : 21  5.520000    5.520000    1461.515993 200958.448972    7.689449   0.000000    0.000000
1/31/2023  4:54:12 PM   : 22  3.210000    3.210000    1238.197865 170252.206389    7.270269   0.000000    0.000000
1/31/2023  4:54:12 PM   : 23  54.580000   54.580000   3941.735595 541988.644320    10.769981  0.000000    0.000000
1/31/2023  4:54:12 PM   : 24  51.920000   51.920000   3752.508506 515969.919630    10.585852  0.000000    0.000000
1/31/2023  4:54:12 PM   : 25  74.310000   74.310000   5778.926068 794602.334369    12.313420  0.000000    0.000000
1/31/2023  4:54:12 PM   : 26  76.820000   76.820000   6106.266505 839611.644478    12.552884  0.000000    0.000000
1/31/2023  4:54:12 PM   : 27  60.580000   60.580000   4405.975025 605821.565880    11.198390  0.000000    0.000000
1/31/2023  4:54:12 PM   : 28  73.360000   73.360000   5663.266685 778699.169155    12.226683  0.000000    0.000000
1/31/2023  4:54:12 PM   : 29  84.010000   84.010000   7291.754234 1002616.207171   13.355537  0.000000    0.000000
1/31/2023  4:54:12 PM   : 30  32.550000   32.550000   2852.717474 392248.652672    9.642750   0.000000    0.000000
1/31/2023  4:54:12 PM   : 31  10.120000   10.120000   1784.216638 245329.787691    8.226245   0.000000    0.000000
1/31/2023  4:54:12 PM   : 32  88.740000   88.740000   8424.115823 1158315.925611   14.045600  0.000000    0.000000
1/31/2023  4:54:12 PM   : 33  53.370000   53.370000   3854.554218 530001.204966    10.685886  0.000000    0.000000
```


GALORATH

Potential Next Steps

What Else is Possible

Correlation at the Rollup Level

Assigning different correlation values at the rollups. Some subsystems might have more correlation than others.

Tail Dependency

Correlation does not capture tail dependency, but can be modeled with correlation.

Conclusions

Use of Correlation is Recommend and A Single Correlation Value Will Work Well

Correlation among WBS elements should be used when doing risk analysis.

01 **Events that cause cost growth often impact many or all elements.**

02 **Specifying a Single Value is Realistic**

03 **Multiple Model Estimates Can Be Combined (e.g. hardware and software)**

04 **Tools Are Equipped to Do Use a Single Value**

Prevents Underestimation

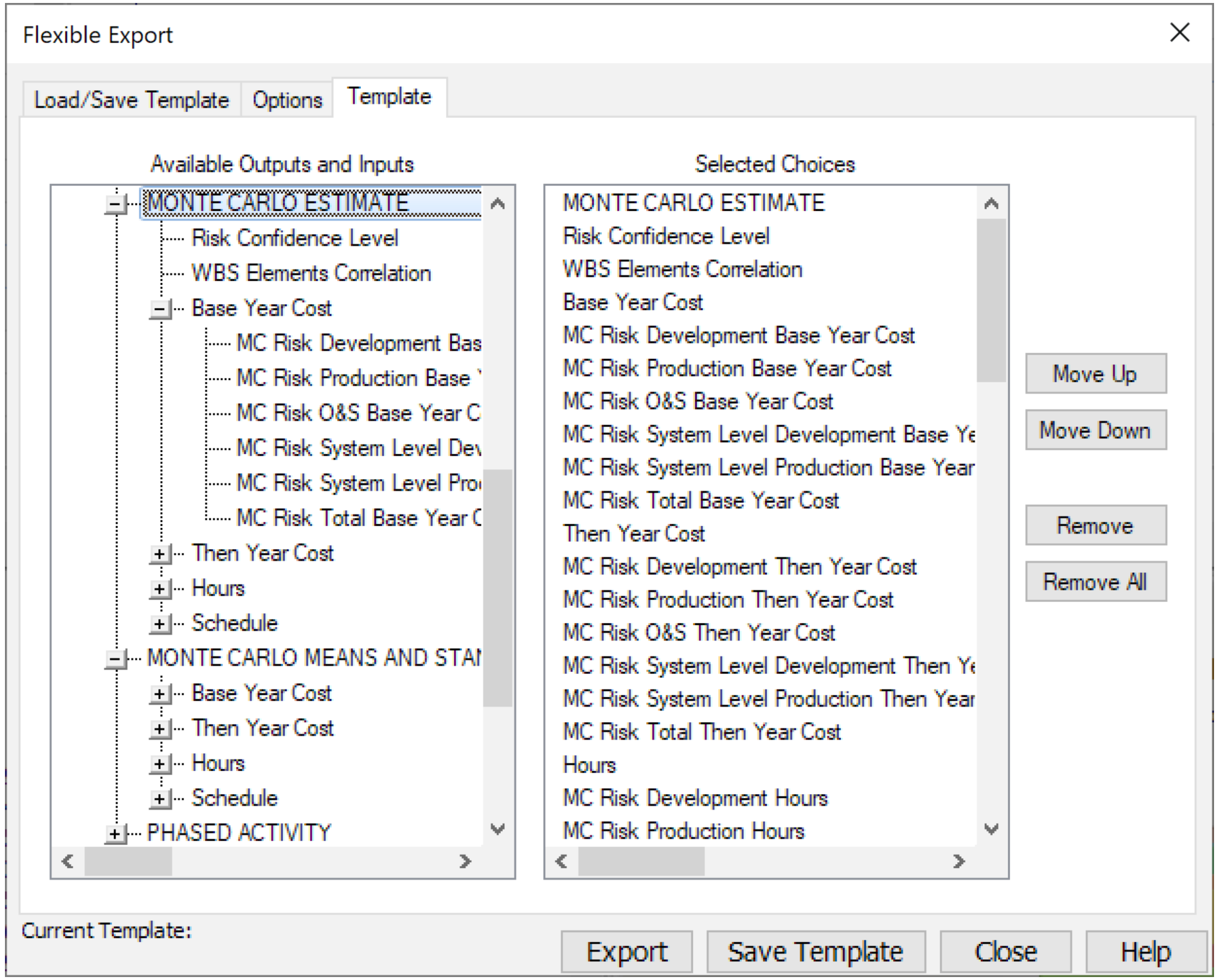
Not considering correlation leads to an under estimation of the uncertainty.

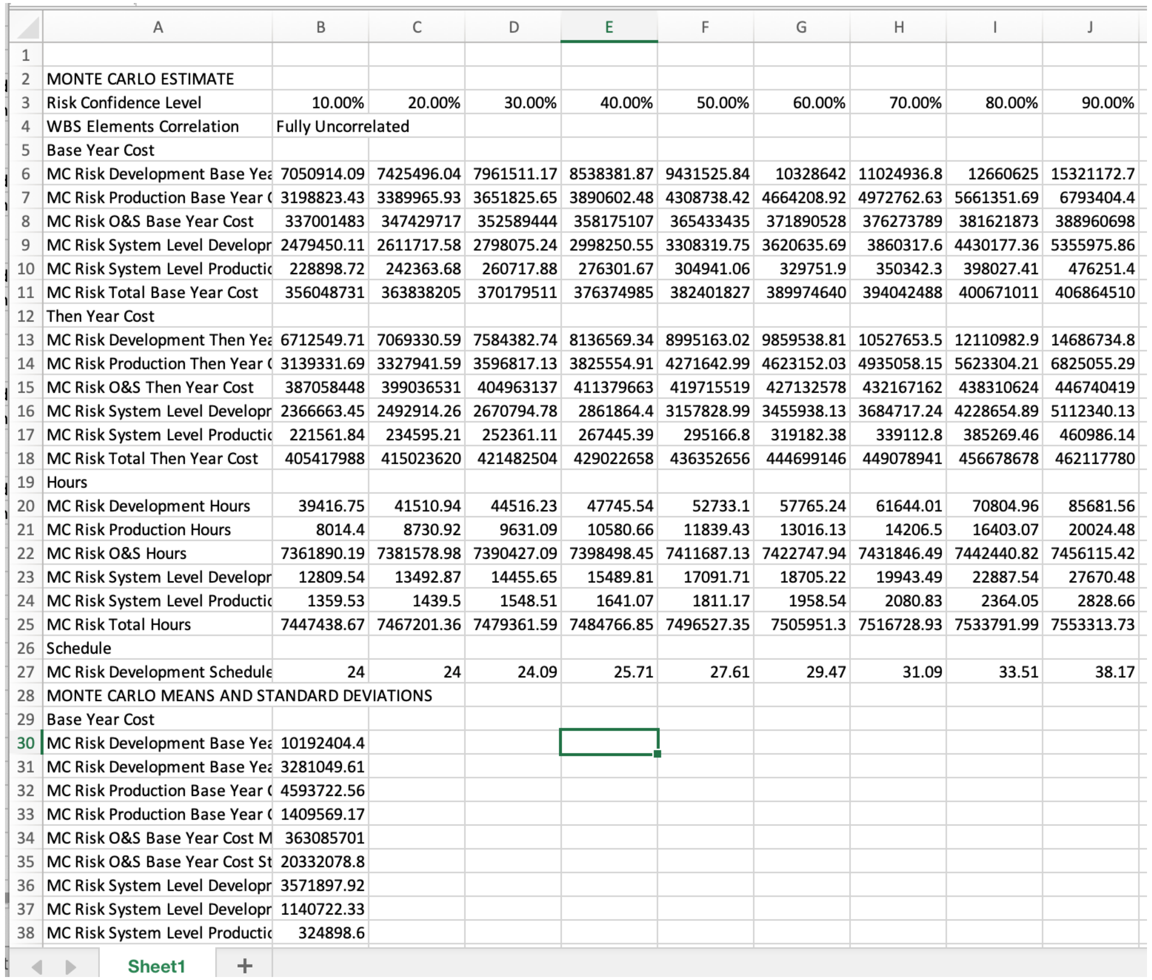
A Single Value is an Effective Approach

This can be specified, but the single best value will likely be between 20%-60%.

Backup Slides

Flex-Ex of Monte Carlo Results





	A	B	C	D	E	F	G	H	I	J
1										
2	MONTE CARLO ESTIMATE									
3	Risk Confidence Level	10.00%	20.00%	30.00%	40.00%	50.00%	60.00%	70.00%	80.00%	90.00%
4	WBS Elements Correlation	Fully Uncorrelated								
5	Base Year Cost									
6	MC Risk Development Base Yea	7050914.09	7425496.04	7961511.17	8538381.87	9431525.84	10328642	11024936.8	12660625	15321172.7
7	MC Risk Production Base Year (	3198823.43	3389965.93	3651825.65	3890602.48	4308738.42	4664208.92	4972762.63	5661351.69	6793404.4
8	MC Risk O&S Base Year Cost	337001483	347429717	352589444	358175107	365433435	371890528	376273789	381621873	388960698
9	MC Risk System Level Developr	2479450.11	2611717.58	2798075.24	2998250.55	3308319.75	3620635.69	3860317.6	4430177.36	5355975.86
10	MC Risk System Level Productic	228898.72	242363.68	260717.88	276301.67	304941.06	329751.9	350342.3	398027.41	476251.4
11	MC Risk Total Base Year Cost	356048731	363838205	370179511	376374985	382401827	389974640	394042488	400671011	406864510
12	Then Year Cost									
13	MC Risk Development Then Yea	6712549.71	7069330.59	7584382.74	8136569.34	8995163.02	9859538.81	10527653.5	12110982.9	14686734.8
14	MC Risk Production Then Year (	3139331.69	3327941.59	3596817.13	3825554.91	4271642.99	4623152.03	4935058.15	5623304.21	6825055.29
15	MC Risk O&S Then Year Cost	387058448	399036531	404963137	411379663	419715519	427132578	432167162	438310624	446740419
16	MC Risk System Level Developr	2366663.45	2492914.26	2670794.78	2861864.4	3157828.99	3455938.13	3684717.24	4228654.89	5112340.13
17	MC Risk System Level Productic	221561.84	234595.21	252361.11	267445.39	295166.8	319182.38	339112.8	385269.46	460986.14
18	MC Risk Total Then Year Cost	405417988	415023620	421482504	429022658	436352656	444699146	449078941	456678678	462117780
19	Hours									
20	MC Risk Development Hours	39416.75	41510.94	44516.23	47745.54	52733.1	57765.24	61644.01	70804.96	85681.56
21	MC Risk Production Hours	8014.4	8730.92	9631.09	10580.66	11839.43	13016.13	14206.5	16403.07	20024.48
22	MC Risk O&S Hours	7361890.19	7381578.98	7390427.09	7398498.45	7411687.13	7422747.94	7431846.49	7442440.82	7456115.42
23	MC Risk System Level Developr	12809.54	13492.87	14455.65	15489.81	17091.71	18705.22	19943.49	22887.54	27670.48
24	MC Risk System Level Productic	1359.53	1439.5	1548.51	1641.07	1811.17	1958.54	2080.83	2364.05	2828.66
25	MC Risk Total Hours	7447438.67	7467201.36	7479361.59	7484766.85	7496527.35	7505951.3	7516728.93	7533791.99	7553313.73
26	Schedule									
27	MC Risk Development Schedule	24	24	24.09	25.71	27.61	29.47	31.09	33.51	38.17
28	MONTE CARLO MEANS AND STANDARD DEVIATIONS									
29	Base Year Cost									
30	MC Risk Development Base Yea	10192404.4								
31	MC Risk Development Base Yea	3281049.61								
32	MC Risk Production Base Year (	4593722.56								
33	MC Risk Production Base Year (	1409569.17								
34	MC Risk O&S Base Year Cost M	363085701								
35	MC Risk O&S Base Year Cost St	20332078.8								
36	MC Risk System Level Developr	3571897.92								
37	MC Risk System Level Developr	1140722.33								
38	MC Risk System Level Productic	324898.6								

Sheet1

Monte Carlo reports can also be copied to the clipboard

Random Probability and SEER

Sampling Uses Random Probability

Each Iteration draws a number between 1%-99%. This can cause differences in Monte Carlo runs. These differences will grow smaller as the number of iterations increase.

SEER Results Need to be Repeatable

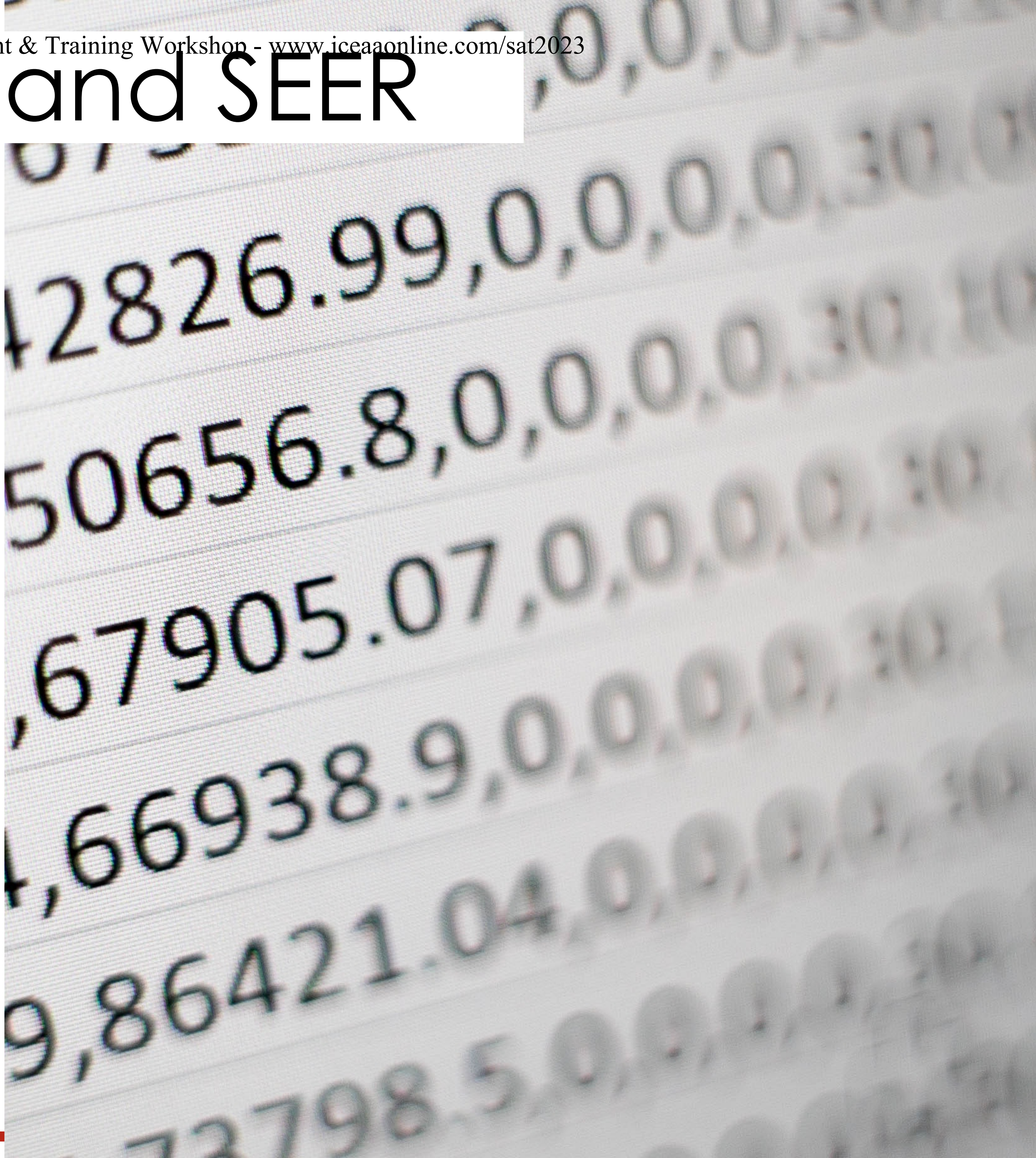
A cost model gives different answers each run is not good. Using a different random number sequence for each MC run would cause this to happen.

SEER Uses a Random Seed

Random Seed is a method to initialize the random numbers generated for sampling. It will ensure that results are repeatable for a given WBS element.

Random Seed Derived From UUID

Each SEER Element has a UUID. The element seed is derived from the UUID. Two identical elements in SEER will have different UUIDs, and thus slightly different Monte Carlo results.

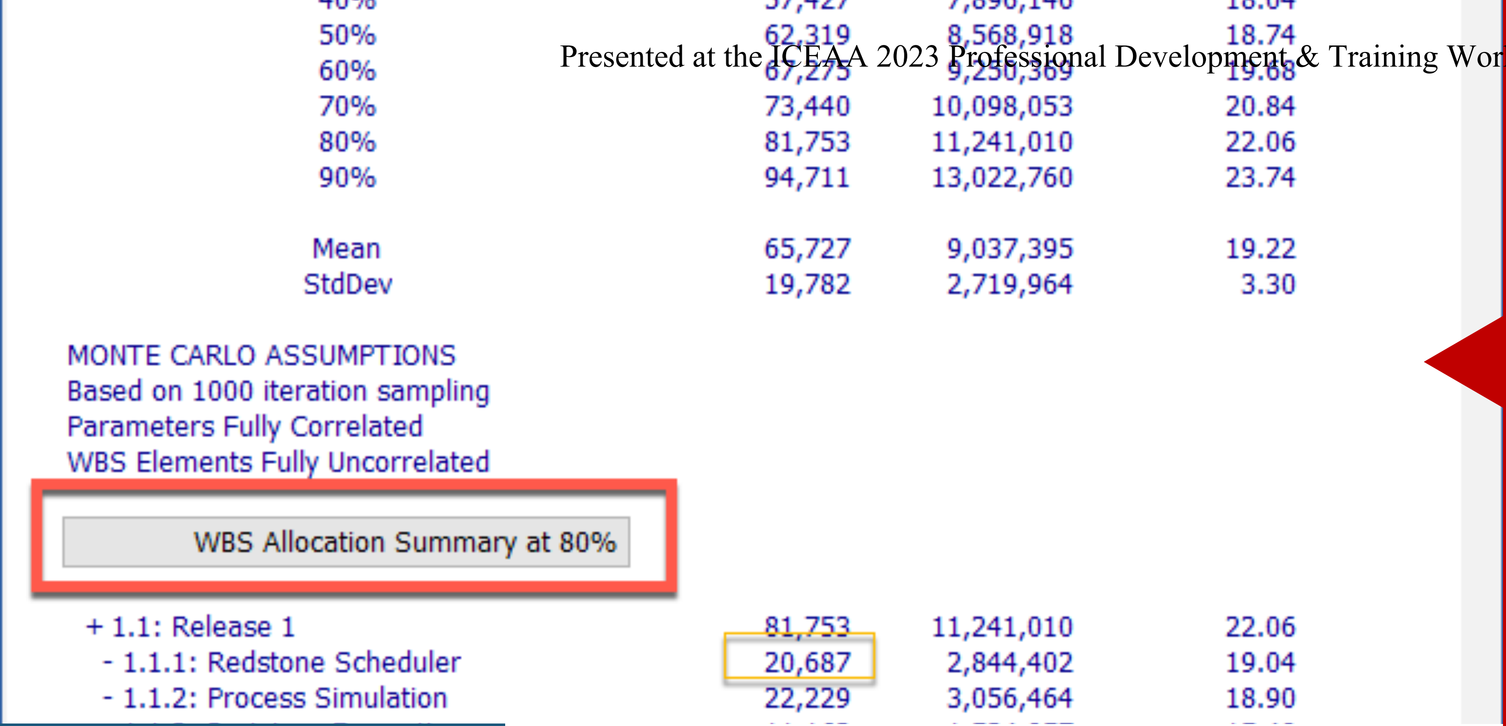


Reconciling Monte Carlo Result & The Point Estimate

The point estimate (generally) won't match the Monte Carlo result

The point estimate has all the great detail (monthly, yearly, by role, by activity) but you don't have a confidence level for it. What you can do is understand the confidence level of the point estimate.

> With the power comes the responsibility
>
> **- Alan Morgan**



01

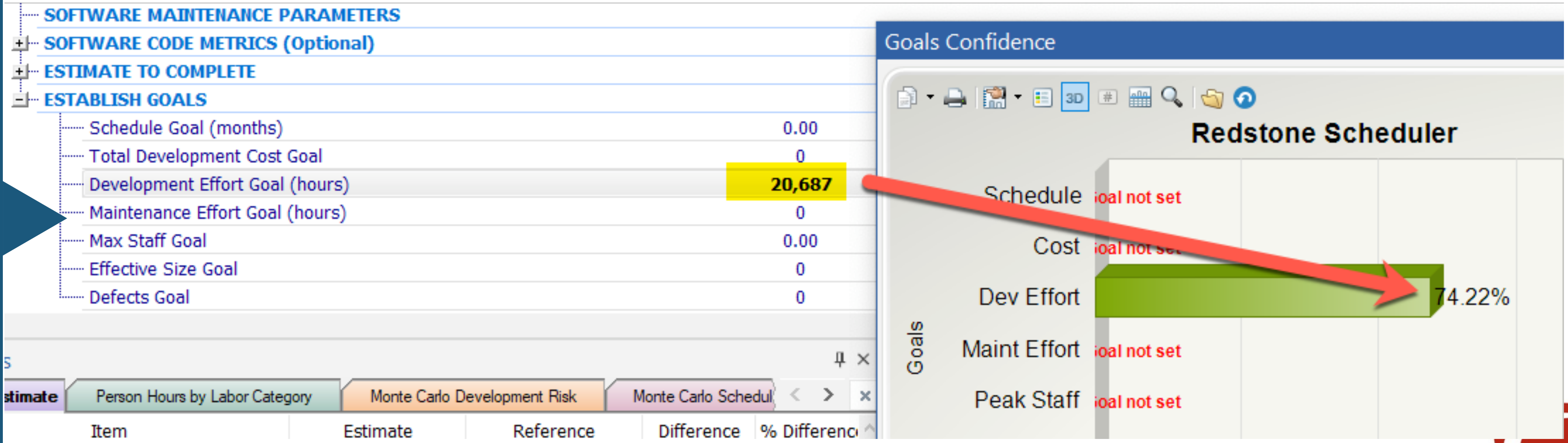
Identify Target Element Hours/Cost

Find the confidence level and view the allocation for that element. For instance, say you want 80% confidence. Set that and note the hours needed for the specific WBS element.

02

Goals

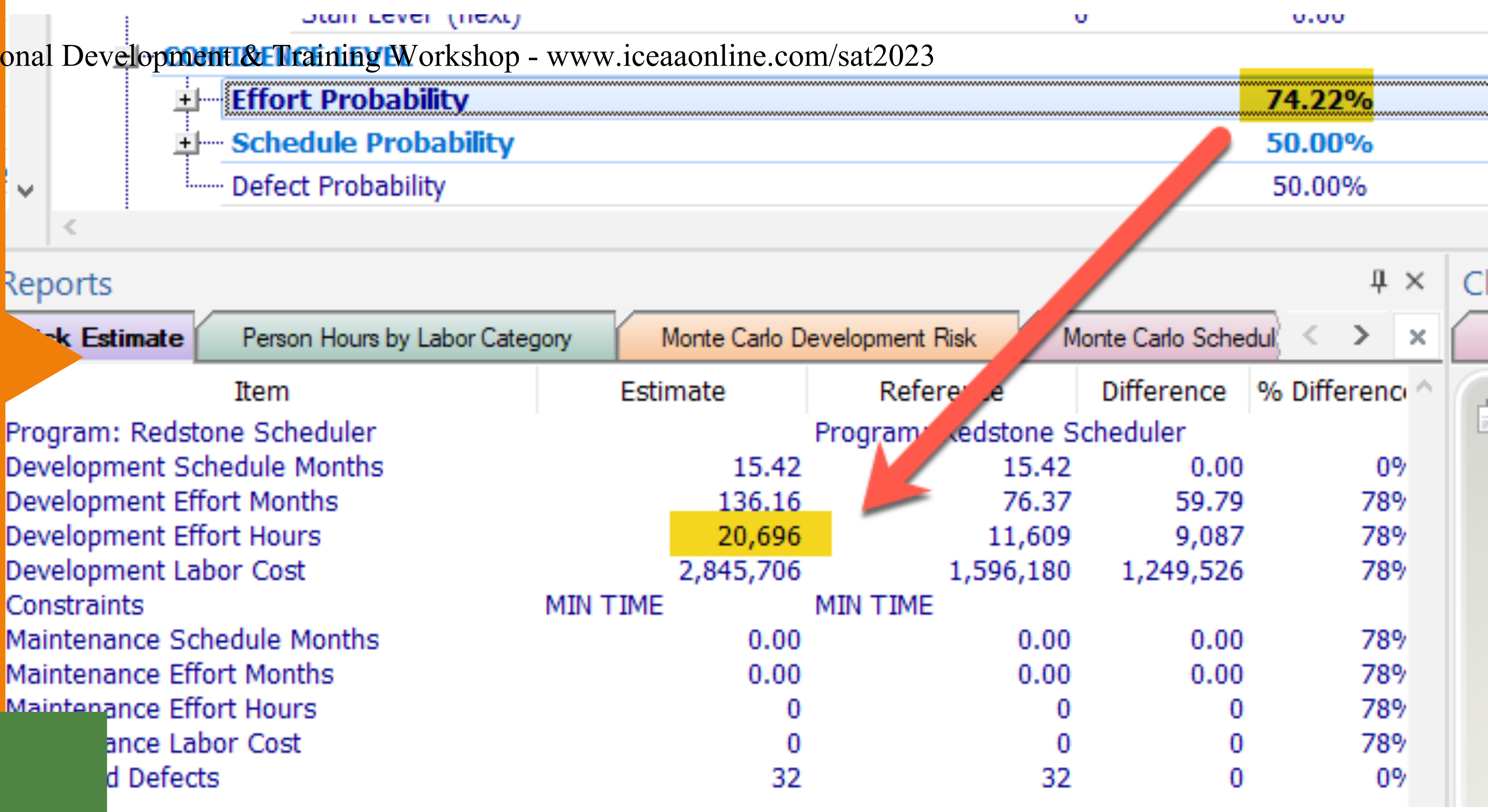
At the element level, set the effort goal. Goals Confidence chart shows the probability input needed.

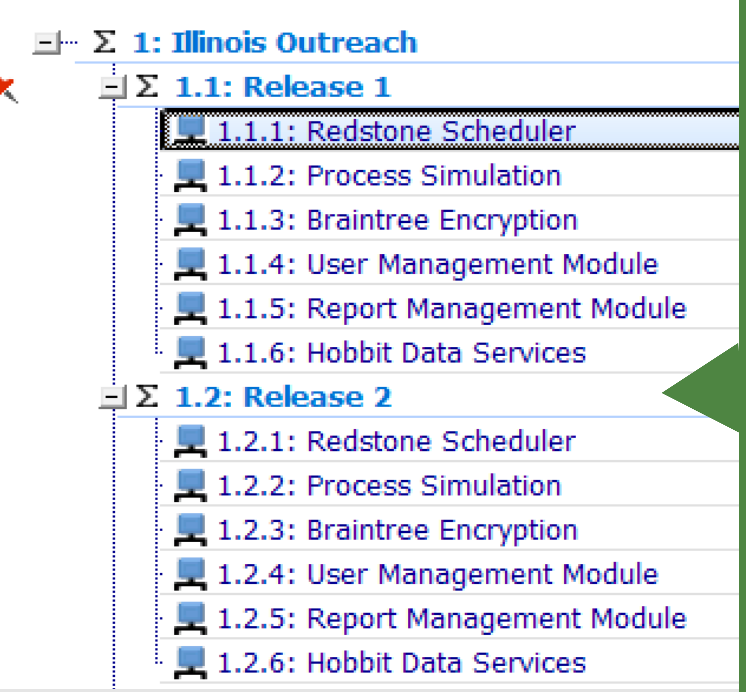


03

Apply Element Probability

Enter the probability found on the goals result into the probability input.





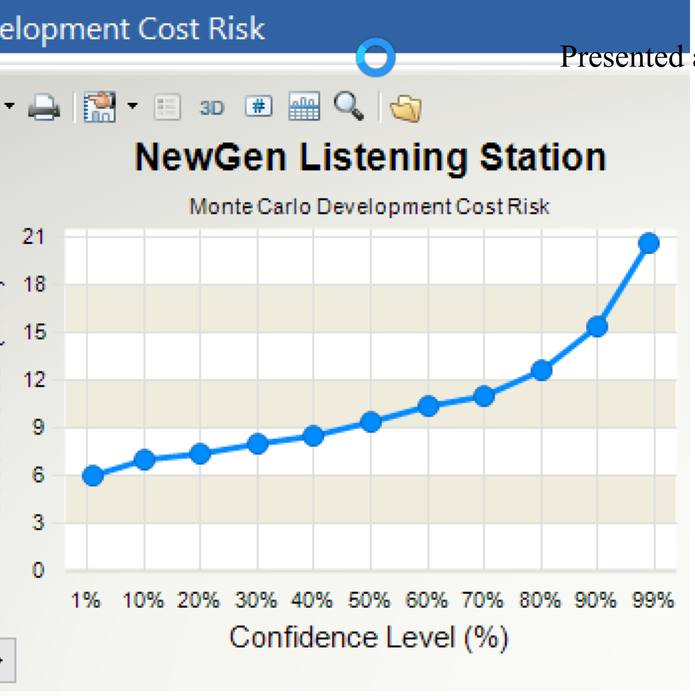
04

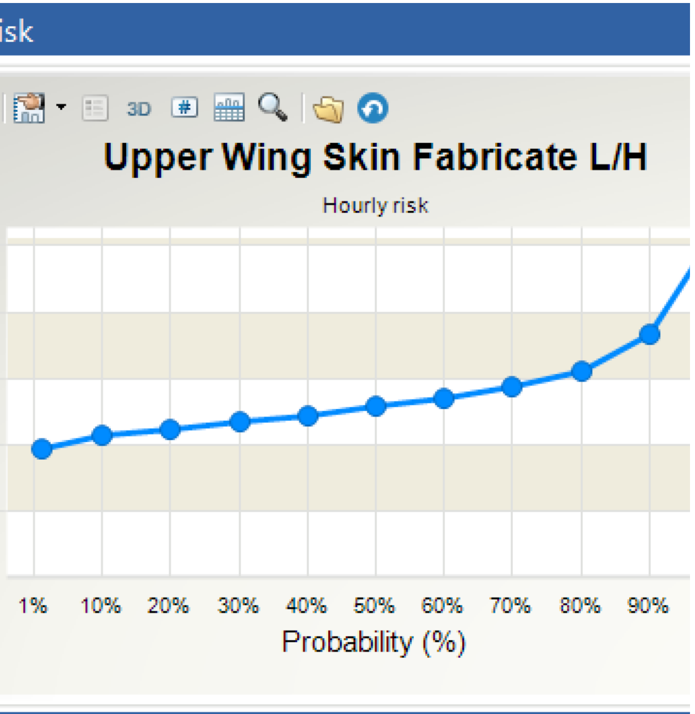
Repeat!!!

This needs to be done for each element in your WBS/Rollup.

The result is a point estimate that will roughly reflect the Monte Carlo result. It won't be exact,

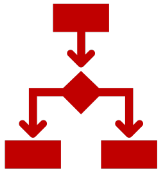
Only SEER-SEM has the Goals feature. A similar process can be done in the other products by using a manual search for the element level probability.





Why this is important

To better understand what your results mean



Better Decisions

Risk and Monte Carlo analysis are about supporting informed decision making. By assigning confidence to your estimate, it can help with difficult decisions: how much to bid, how much to spend, how much are you willing to lose to win new business.

Learn from the Experts

The government has put lots of brain share into understanding risk and the quantitative methods that can be used quantify risk. They understand that adding up point estimates will probably underestimate the program. The risk features in the SEER tools are designed to support this robust analysis.