Eric Druker

Graham Gilmer

David Hulett

Using Stochastic Optimization to Improve Risk Mitigation

March, 26 2015

Table of Contents

List of Figures	iii
Abstract	iv
Introduction and Traditional Risk Prioritization Methods	1
Using Stochastic Optimization to Prioritize Risks	1
Method One: Automated Risk Prioritization	2
Method Two: Automated Step-Wise Risk Prioritization	3
Method Three: Automated Knapsack Risk Prioritization	4
Comparison of Run Times	5
Conclusion & Future Research	6
Bibliography	13

List of Figures

Figure 1	"Traditional Risk Cube"	1
Figure 2	"Risk 1 Exhibiting High Correlation to Final Cost"	2
Figure 3	"Risk 2 Exhibiting High Correlation to Final Cost"	3

Figure 4	"Traditional Risk Correlation Chart Based on Correlation"	3
Figure 5	"Analysis Run-Time for Automated Risk Prioritization (Minutes)"	8
Figure 6	"Analysis Run-Time for Automated Step-Wise Risk Prioritization (Minutes)"	8
Figure 7	"Analysis Run-Time for Automated Knapsack Risk Prioritization (Minutes Except Where Noted)"	9

Abstract

Today's risk analysts have several tools to help them identify and mitigate future sources of cost and schedule risk. Traditionally, the risk cube method has been used to provide probability-weighted metric for each risk's severity by multiplying its likelihood and consequence factors together. Unfortunately, this methodology ignores secondary and tertiary impacts of risks, in particular when they could drive cost by creating a new critical path within the project plan. Integrated cost & schedule risk analysis provides greater insights by integrating the risks into the schedule. Using traditional sensitivity metrics such as Pearson's correlation, analysts are able to identify risks contributing to cost and schedule growth. While stronger than the risk cube methodology, analysts using this method are unable to measure a risk's contribution to a particular confidence level of either cost or schedule and cannot uncover when the impact of removing a set of risks may be greater than sum of the impacts of removing them individually. This paper will show how stochastic optimization – the optimization of simulation models – can be used to better identify risks, and combination of risks, that when mitigated will best reduce project cost and schedule risk.

Introduction and Traditional Risk Prioritization Methods

Traditionally, risk managers have relied on the risk cube, a 5 x 5 matrix representing the likelihood and consequence of each risk, to develop a probability-weighted metric for ranking risks for mitigation. Unfortunately, this method is of limited value due to a couple of shortcomings. First, the ranking's usefulness is largely dependent on the quality of the scale used to establish consequence. Should the bins for consequence factors be too wide, or the threshold for the highest consequence bin be set too low, risks with vastly differing impacts could be grouped into the same factor. If they have similar likelihoods of occurrence they will not be differentiated in expected value. Second, and most importantly, both likelihood and consequence factors are typically developed by subject matter experts focusing only on the area of the project directly impacted by the risk. More often than not, the downstream – secondary and tertiary – impacts of risks can exceed the localized impact by orders of magnitude. For instance, risks with small individual consequences may lead to enormous cost and schedule growth on the project should their occurrence create a new critical path that drives standing army level of effort (LOE) costs for the length of the delay. Additionally, the occurrence of one risk may increase the likelihood of another risk occurring. Unfortunately, both of these shortcomings mean that while the risk cube provides a concise quick-look assessment of risk, it should not be used to rank risks on anything but the most simplistic projects. For these reasons, many analysts are now using sensitivity results from integrated cost and schedule risk analysis models to rank risks.

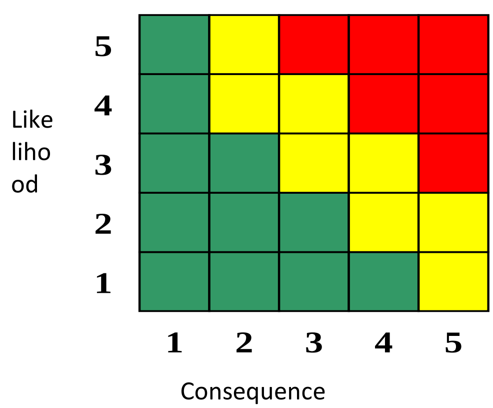

Figure 1: Traditional Risk Cube

To address the challenges associated with the risk cube method, many analysts have transitioned to ranking risks using the sensitivity analysis metrics associated with most simulation-based cost and schedule risk analysis models. In these models, uncertainty distributions are applied to various aspects of the project including task durations, resource costs, resource utilizations, and – most relevantly – risk probabilities of occurrence and impacts. The model is run thousands of times, with a sample selected from each distribution for each iteration of the simulation. These results are compiled and a range of projected costs and end dates for the project are calculated.

To calculate sensitivity metrics, most models capture the samples from each distribution for each iteration of the simulation and then correlate these to the final cost and schedule. In the examples below, the X-axis represents the occurrence of the risk (0 means the risk did not occur in that iteration, 1 means the risk did occur in that iteration) and the Y-axis represents the final cost for the project. To rank risks, a regression line is drawn across this data and the correlation between the risk occurrence and final cost is calculated. In figure one, we see a risk (Risk 1) that is highly correlated[1] to the total cost of the project. In figure two, we see a risk (Risk 2) that exhibits almost no correlation with the total cost of the project. More over, we can also see the ghost of risk 1 in the chart as exhibited by the discontinuous jump in total cost whether or not risk 2 occurs. In this case, the top set of data points represents where risk 1 occurs, further driving home its impact on cost.

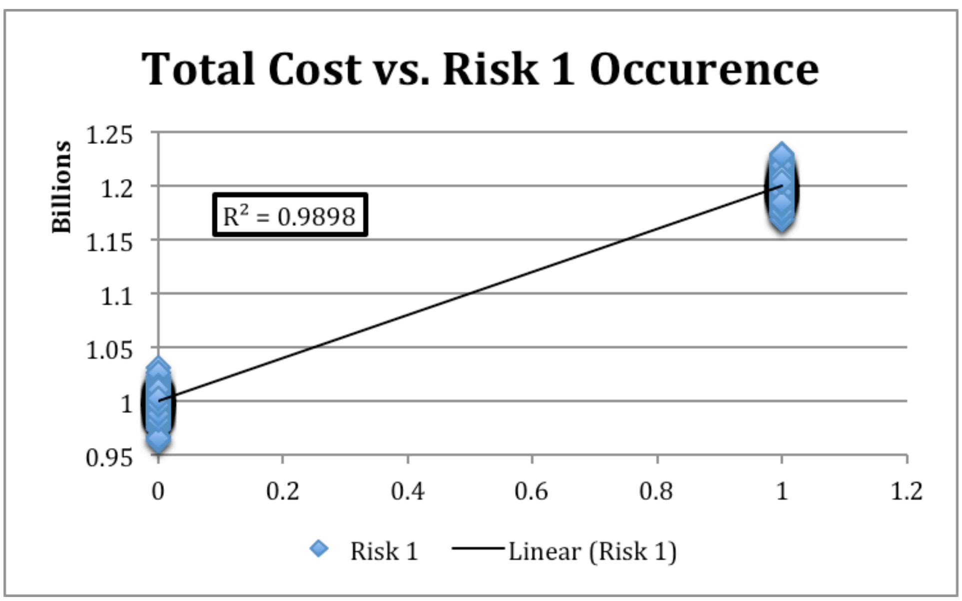

Figure 2: Risk 1 Exhibiting High Correlation to Final Cost

[1] In figures 1 and 2 we are showing R^2 as a substitute for actual correlation. In most models, r (sqrt(R^2)) is captured to allow opportunities to be modeled beside risks. Additionally, there are many nuances for how correlation is captured in models – the simplest of these is presented for illustration

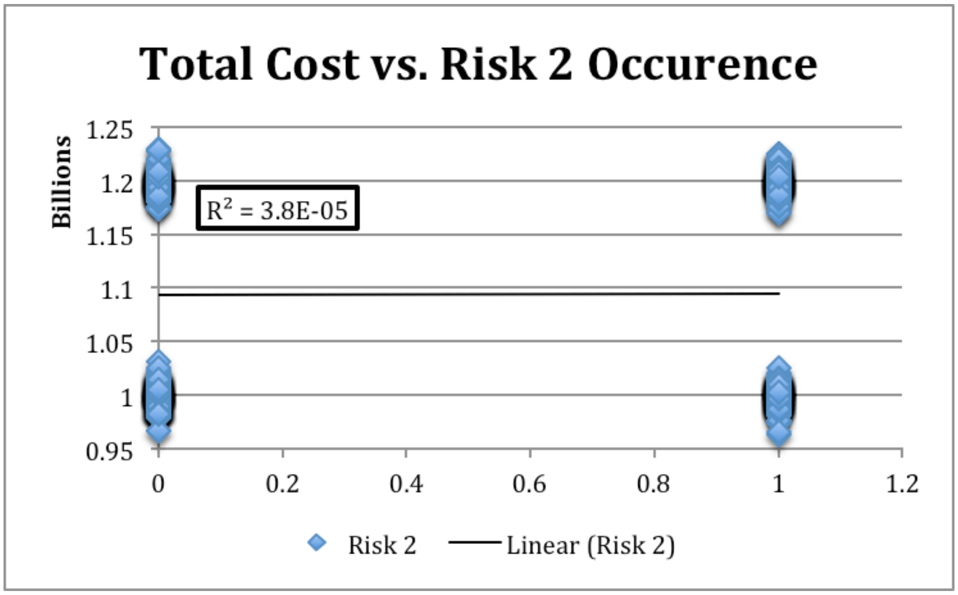

Figure 3: Risk 2 Exhibiting Low Correlation to Final Cost

Each risk's correlation with cost is then plotted using a bar chart (colloquially known as a tornado chart). Those risks with higher correlation can be reasonably assumed to be contributing more to the final cost of the project than those risks with lower correlation. This example can be extended not just to risks, but also to uncertainties; and not just to the projected cost, but all to the projected finish date in order to provide a complete picture of risk.

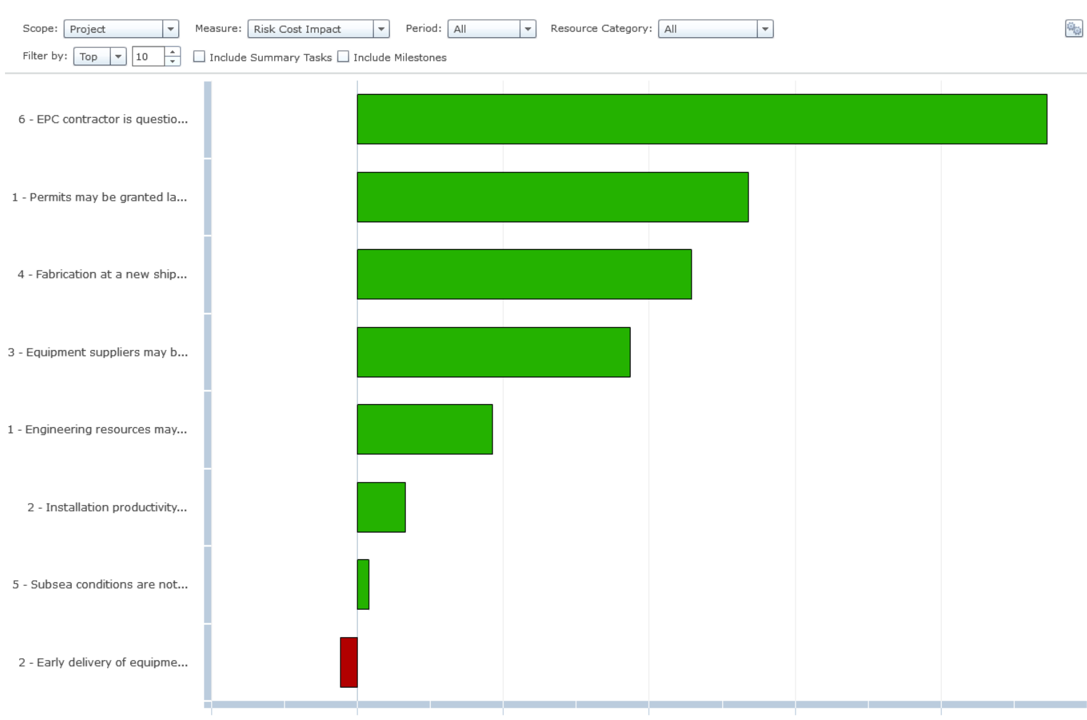

Figure 4: Traditional Risk Prioritization Chart Based on Correlation

Unfortunately, this methodology also has its limitations. First, the correlation metric is difficult for decision makers to understand. Project managers and company executives are used to seeing results in dollars and days; presenting a unit-less metric not easily converted to more tangible metrics presents a challenge to say the least. Furthermore, this approach for prioritizing risks ranks them on their impact assuming that none are mitigated. Although the highest correlated risk is typically the best place to start mitigating, once this risk is removed the risk rankings are almost certain to change. For example, if removing a risk results in a new critical path to the schedule, risks which were ranked low in terms of impact but which may now fall on the new critical path, could leapfrog all the way to the top of the rankings. Conversely, two risks may be ranked individually low, but be the only two risks on a critical path. While mitigating one or the other may provide limited value, mitigating both could provide substantial cost and schedule savings. Neither of these situations can be accounted for using traditional correlation metrics.

In order to provide analysis of maximum value to risk analysts, project managers, and executive decision makers we need to develop a risk prioritization metric that 1. Works in tangible figures such as dollars and days, and 2. Accounts for combinations of risks which, when removed together, could result in a substantial reduction of cost and schedule risk.

Using Stochastic Optimization to Prioritize Risks

Optimization involves changing variables, according to constraints, in order to maximize an objective function that they feed into. The variables are those items which one has the ability to change, constraints are the bounds within which the variables must operate, and the objective function is a calculated value, or set of values, based on a set of mathematical or logical rules and the variables used in the model.

In the field of project management, we have many variables we can use to affect the project: work sequencing, cost phasing, contract type, and most important for this paper, risk mitigation. All of these variables must operate within a certain set of constraints: we only have a limited budget for each year and our projects must finish within a certain schedule constraint. At the end of the day, our goal as project analysts is to identify those actions, living within our constraints, which will have the greatest impact on mitigating cost and schedule growth to the project.

For this particular paper, we will focus on three different metrics for using stochastic optimization to identify risks, and combinations of risks, that when mitigated will result in the greatest possible reduction in project cost and schedule risk. In the process we will develop metrics that are in tangible metrics such as dollars and days.

One of the greatest challenges in producing each of the three metrics is the run-time to perform the analysis. Run-time performance for each metric for various size integrated cost & schedule risk analysis models (with size being measured by the number of lines in the schedule) will be provided.

Method One: Automated Risk Prioritization

The first challenge the research team focused on was building a set of metrics that, unlike the unit-less correlation measure, were tangible and presented in days and dollars. Conversations with several colleagues in the oil/gas industry provided a methodology, currently in use, that provides analysis in these units. In the integrated cost & schedule risk analysis model, the user selects several reporting metrics: 1. The confidence level at which they would like the results reported, 2.Where in the hierarchy they would like the results reported. For example, the user could preference that they want to see risks prioritized according to their impact on the 50th percent confidence level for the project's end date. Or, to provide more granularity, they could request risks prioritized according to their impact on the 80th percentile for costs in 2017 for the detailed design phase of the project.

Once the measurement calculation is selected, the model is then simulated to get the baseline results for the selected metric. At this point, the first risk is removed, the model re-simulated, and the desired metric collected once again. Lastly, the first risk is added back into the model and the second risk is removed. This process is repeated until the model has been calculated without each risk. The risks are then ranked in order of priority based on how much the desired metric (cost or schedule) was reduced when that risk was removed.

The resulting metric provides us a tangible quantification of each risk's contribution to the cost and schedule for the desired component of the project. However, it still lacks in that it is only quantifying impacts on a model with n-1 (where n is the starting number of risks) risks. This methodology cannot account for situations where the removal of two risks may result in a greater savings than sum of their impacts when removed individually.

Method Two: Automated Step-Wise Risk Prioritization

In order to account for aforementioned situation where removing multiple risks may result in a savings greater than the sum of their impacts when removed individually, we have developed a new method for prioritizing risks. In this methodology, we begin by applying the previous method of removing each risk individually and ranking them according to their impact on our target cost or schedule metric. At this point, the methods diverge. In the step-wise risk prioritization, we will now forget about all but the highest ranked risk. Then, with the remaining n-1 risks, we will once again remove each of them individually in order to show their impact on cost and schedule *with the first risk already removed*. Once we've completed this we re-rank the n-1 risks to see which removal had the greatest impact on cost and schedule risk reduction. At this point, we permanently remove the highest-impact risk and continue the process with the remaining n-2 risks.

This fundamentally differs from the previous algorithm in that it captures instances where removal of a critical set of risks has a discontinuous jump in cost and schedule savings. This methodology provides project managers with an actionable plan for mitigating risks in that they can go mitigate the risks in order of priority and immediately see the reduction they should expect in cost and schedule risk.

Method Three: Automated Knapsack Risk Prioritization

The challenge with the previous method is, if a project manager can only mitigate a certain number of risks, the method cannot for sure tell them which combination of risks will provide the maximum value. In other words, the project manager may optimally remove risks A and B if they only have the ability to mitigate two risks but may optimally remove risks B, C, and D if they can mitigate three.

This methodology is a take on the well-studied combinatorial optimization question known as the knapsack problem. This problem is defined as follows: "Given a set of items, each with a mass and a value, determine the number of each item to include in a collection so that the total weight is less than or equal to a given limit and the total value is as large as possible."[2] In our frame of reference, the problem is redefined as "given a set of risks, each with a probability of occurrence and cost/schedule impact, determine the optimal combination of risks that, when removed, will reduce cost and schedule risk as much as possible.

This solution is optimal in that it finds all different combinations of risks that will allow for mitigation however it is unlikely to be feasible using existing technology in all but the most basic of cases.

Comparison of Runtimes

Each of these three methodologies provides the user with insights beyond those produced with traditional risk cube or correlation-based methods. That being said, each requires multiple simulation runs of integrated cost and schedule risk analysis models. To evaluate the feasibility of each type of analysis, the team calculated the number of simulation runs required to prioritize the top 10 risks. For Method Three, Automated Knapsack Risk Prioritization, the time required to prioritize groups of risk sized from 1 to 10 were calculated. The team assumed that 1,000 iterations were used for each simulation run. Runtimes were calculated for 100, 1000, 3000, and 6,000 line schedules containing 10, 50, 100, and 150 risks. Booz Allen's Polaris tool was used as a basis for calculating simulation run-time.

Method One, Automated Risk Prioritization was the simplest of the three problems. Since each risk was only removed once and since combinations of risks were not considered, the number of

[2] http://en.wikipedia.org/wiki/Knapsack_problem

simulation runs was equal to the number of risks. Based on Polaris run-time benchmarking, this resulted in analysis times as listed below.

Number of Tasks in Model	Risks: 10	50	100	150
100	0	0.7	0.3	0.5
1000	0.2	0.9	1.9	2.8
3000	1	4.9	9.7	14.6
6000	3	15	30	45

Figure 5: Analysis Run-Time for Automated Risk Prioritization (Minutes)

Method Two, Automated Step-Wise Risk Prioritization adds a layer of difficulty as it now accounts for the combined effects of mitigating multiple risks. Since each risk is removed up to 10 times (based on whether or not it is selected by a priority before the top 10 risks are removed) the number of simulation runs required to prioritize risks is

$$\sum_{i=n-9}^{n} n$$

where n is the number of risks in the model. Based on Polaris run-time benchmarking, this resulted in analysis times as listed below.

Number of Tasks in Model	Risks: 10	50	100	150
100	0.2	1.4	2.9	5.7
1000	1	8.5	17.8	35.6
3000	5.4	44.3	93	185.4
6000	16.5	136.5	286.5	571.5

Figure 6: Analysis Run-Time for Step-Wise Risk Prioritization (Minutes)

Method Three, Automated Knapsack Prioritization adds significantly more difficulty than either method two or three as it is now a combinatory problem. For this method, the number of simulation runs required to prioritize risks is

$$\sum_{i=1}^{10} \frac{n!}{(n-1)!}$$

where n is the number of risks in the model. Based on Polaris run-time benchmarking, this resulted in analysis times as listed below. Note that the runtime has increased dramatically to the point where it is no longer feasible on models with more than a handful of risks.

	Risks	10	50	100	150
Number of Tasks in Model	100	0.1 hours	76.7 years	110,821 years	7 million years
	1000	0.3 hours	477.1 years	689,555 years	45 million years
	3000	1.7 hours	2,487.5 years	4 million years	233 million years
	6000	5.1 hours	7,667.1 years	11 million years	718 million years

Figure 7: Analysis Run-Time for Knapsack Risk Prioritization (Minutes Except Where Noted)

Conclusion & Future Research

This paper has presented three methodologies for prioritizing risk for mitigation. Of the three, two have analysis times that are reasonable for the value they provide. As a recommendation, we believe that risk analysts should consider using method two, step-wise risk prioritization to better inform decision makers on how they can best mitigate cost and schedule risk. This method has the advantage of producing results that are in tangible dollars and days units while also allowing analysts to see when mitigating a group of risks may provide a discontinuous jump in savings. With even the most complex models running within the span of a single day (with most models capable of running in under an hour) this methodology provides value that far exceeds the time required to run it.

This analysis assumes that when a risk is mitigated it is completely eliminated on the project – that its likelihood and consequence both drop to zero. Applying another gradient to the analysis, allowing likelihood and consequence factors to be partially reduced, would likely provide a more realistic view of project manager’s ability to reduce risk.

Accounting for risks is a first step, but risks are only one of the many components that drive cost and schedule risk on projects. Future research should begin by focusing on how underlying cost and schedule uncertainty can be rolled into the analysis.

Bibliography

No.	Description
1	Smith, C, Herzog, H. (2012). Using Optimization Techniques to Enhance Cost and Schedule Risk Analysis. International Society of Parametric Analysis International Symposium