DOI NUMBER: 10.19255/JMPM01603

RISK TOOL

KEYWORDS
Project Risk Management • Risk Matrix • Quantitative Risk Analysis • Heatmap • Color Vision Deficiency • Accessible colors

Improving RISK MATRIX DESIGN using HEATMAPS AND ACCESSIBLE COLORS

MARTIN ENTACHER
• Ph.D. - Project Manager at RiskConsult
• *entacher@riskcon.at*

PHILIP SANDER
• Ph.D. - Managing Director at RiskConsult
• *sander@riskcon.at*

• ABSTRACT •

Risk management is an integrated part of project management to identify, assess, evaluate and mitigate threats or opportunities that aid or endanger the goals of a project. Risk Matrices are one of the most popular tools to execute a proper risk management process. However, they suffer from shortcomings that can lead to non-meaningful results and oversimplified classification. Cox (2008a) and other researchers have described these shortcomings, Thomas et al. (2013) even proposed to avoid risk matrices in general. This paper presents adaptions that will prevent oversimplified classification, e.g. continuous coloring schemes and viewing risk probability and risk consequence as independent variables. Furthermore, the typical green-yellow-red color scheme is questioned due to its lack of perceptual uniformity and inaccessibility for people with color vision deficiencies. It is proposed that these changes will shift the area of application from standalone risk analysis tools to supporting tools for quantitative risk analysis which will develop to be the state of the art in project risk management.

1. INTRODUCTION

--- 1.1 Motivation ---

Project risk management is one of the major aspects of project management. Project managers need knowledge about professional risk management and they need the tools to identify, analyse, evaluate and treat (mitigate) risks. Project managers seek to meet specific goals such as cost and time. ISO 31000 defines risk as the "effect of uncertainty on objectives". This deviation from the expected can be positive or negative. Hence, risks are threats and/or opportunities that either compromise project goals or aid in achieving them. Modern project management has to deal with complex situations under uncertainty. Thus, risk management is an essential and integral part of it.

While ISO 31000 provides a generic framework for risk management, ISO/IEC 31010 offers a variety of tools to put risk management into practice. One of these tools – and possibly one of the most widespread ones – is the so-called consequence/probability matrix which is better known as risk matrix. Risk matrices are used to visualize risks in a two-dimensional space (e.g. risk probability/frequency and risk consequence/impact). Each cell in a risk matrix is associated with a color that is used for classification. The color is determined by multiplying the two variables probability and consequence. The typical green, yellow and red colors represent low, medium and high risks respectively. The risk class (i.e. color) will then determine which risks have a high priority for counter-measures or mitigation.

Risk matrices are popular due to their supposed clarity and ease of use. However, a number of papers over the last 10 years have drawn a critical picture regarding the common standard of risk matrix application (see section 1.2). While major improvements were suggested by some researchers, others have concluded that risk matrices should be avoided generally.

Based on these papers the present article will offer a variety of adaptions to develop a set of best practice risk matrices. That includes treating "probability of occurrence" and "impact" as independent variables (such as in a Monte Carlo Simulation), questioning the adequacy of the green – yellow – red color scheme and implementing continuous heatmaps instead of discrete colors.

--- 1.2 Related Work ---

In 2008 Cox wrote a critical paper (Cox, 2008a) in which he pointed out (and mathematically proved) that typical risk matrices do not fulfil their purpose of supporting good project management decisions. Therefore they have to be used with caution (Cox, 2008a; Cox and Huber, 2008). He observed poor resolution, errors in the assignment of risk ratings, suboptimal resource allocation and ambiguous inputs and outputs. To overcome these shortcomings, Cox concluded that risk matrices have to fulfill the criteria of weak consistency, betweenness and consistent coloring. Thomas et al. (2013) reviewed 30 publications from the oil & gas industry and concluded that risk matrix application produced arbitrary risk management decisions.

Another source of inaccuracies, the use of qualitative risk rating systems or – more generally speaking – inadequate scales, was described by Cox et al. (2005), Hubbard & Evans (2010) and Hubbard & Seiersen (2016). Qualitative scales can be interpreted extremely inconsistently, i.e. assessing a consequence of 4 on a scale from 1-5 will mean different things to different experts. The same can be said for the interpretation of words such "low" or "high". Unless there is a clear description, such terms are ambiguous. Quantitative scales will lead to better results. If qualitative scales have to be used, clear definitions of the meaning of words such as "high" or "low" or numerical scales (1-5) are needed. This effect is amplified by biases that experts and non-experts both have. Cognitive biases were described in great detail by Montibeller & Winterfeldt (2015) and brought into the context of risk matrices by Ball & Watt (2013).

Another issue is range compression. Levine (2012) and Baybutt (2016) recommended to overcome it issues by using logarithmic scaling. Various important thoughts on resolution / range compression were presented by Dujim (2015) who also proposed to use a coloring scheme based on iso-lines instead of the typical grid. Similar coloring schemes were also presented by Ni et al. (2010).

Olson & Wu (2017) provide useful advice for risk matrix application in enterprise risk management. Elmontsri (2014) reflects on risk matrix application within the National Health Service (NHS, UK). Papers focusing on the process of how to find the right risk matrix for a specific application were presented by Dillon et al. (2009), Li et al. (2017) or Oliveira et al. (2015).

2. DESIGN OF RISK MATRICES

--- 2.1 General Principles ---

In a typical risk matrix probability of occurrence and consequence are both given a range of 1-5. The product of probability and consequence will then result in a number of 1-25 which is called the risk rating. If the scales of 1-5 for probability and consequence are not defined specifically with numbers such as USD or percentages, arbitrary results are likely to occur due to the ambiguous nature of the scale. To avoid this, quantitative risk analysis should be preferred over qualitative risk analysis whenever possible. Instead of 1-5, probability will then be quantified from 0% - 100% or with a frequency (rate of occurrence) and risk consequence can be quantified in terms of cost and/or time. In typical project management applications,

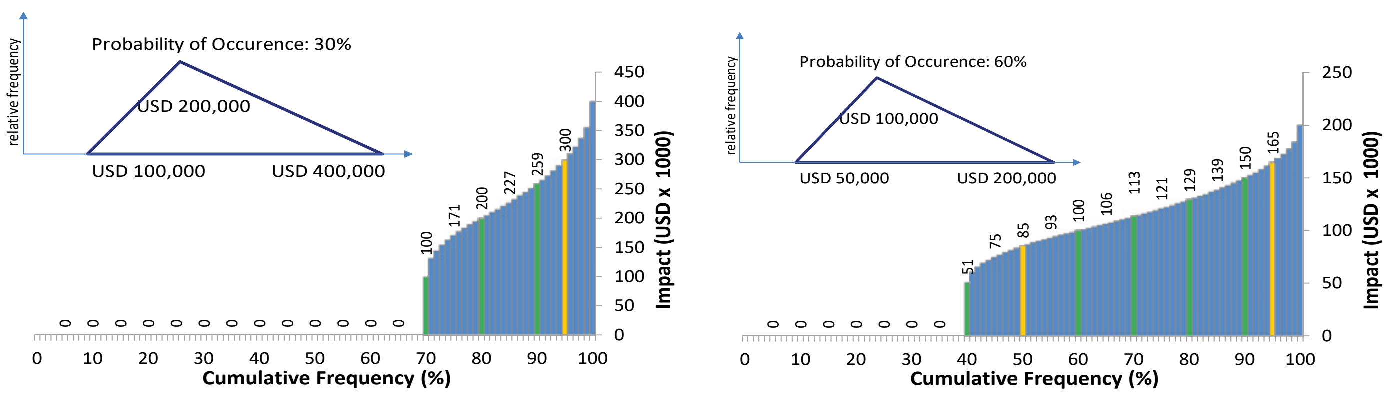

FIGURE 1. Distribution function for two single risks modeled with a triangular distribution and a 30% / 60% probability of occurrence (figure created with RIAAT, http://riaat.riskcon.at)

goals are quantifiable. Hence, deviations from goals will also be quantifiable. There is seldom a need to use an abstract scale (1-5) instead of actual numbers. In spite of this, such applications are widespread. We believe that this is due to the tempting simplicity of semi-quantitative (scales 1-5, risk ratings 1-25) risk matrices. In the case of risk matrices this simplicity has to be treated with caution because it is often an oversimplification that will lead to meaningless results. This was shown by previous researchers (see section 1.2).

To overcome this weakness, there is a need to avoid semi-quantitative scales. It will be shown in this paper that risk matrices cannot only be used for qualitative risk analysis, but also as a visualization tool for quantitative risk analysis which is the preferred method in project risk management. In applications for non-profit organizations or for assessing more abstract high-level risks (e.g. political), quantifying consequences will sometimes be impossible. In such situations it is necessary to stick to qualitative (low, high) scales. Describing the meaning of each word ("low") on such scales is the key factor for a successful and unambiguous application. The subsequent sections will focus on developing a qualitative risk matrix / heatmap that can be seen as a derivative of the semi-quantiative risk matrix (consequence/probability matrix) which was so heavily criticized by previous researchers. An example of an application for quantitative risk analysis will be shown at the end of this paper. The changes that are made to the original risk matrix are the same for both applications, qualitative and quantitative.

--- 2.2 Probability and Consequence as Independent Parameters ---

Figure 1 shows two examples for the quantitative analysis of single risks which are defined by a probability of occurrence (30% / 60%) and a consequence of USD 100,000 – 400,000 / USD 50,000 – 200,000. The product of the two variables will lead to a deterministic risk rating of USD 60,000 for both risks. For the first risk, the distribution function shows us that the risk consequence will be 0 in 70% of all cases and 100,000 or more in 30% of all cases, but it will never be 60,000. The second risk can actually become 60,000, but there is a 40% chance that it will be zero and a >58% chance that it will be higher than 60,000.

In spite of a wide-spread simplification and/or belief, risk is not per se the product of probability and consequence. In fact, the two variables are independent. The risk will occur or it will not occur. Only if it occurs, there will be a consequence. The typical approach of risk matrices to use the product of probability and consequence will lead to an equalization of low probability – high consequence events and high probability – low consequence events by assigning the same color. Such events are however quite different in their nature. Keeping the probability and consequence independent will account for that. However, distinguishing between low probability – high consequence events and high probability – low consequence events seems to destroy the very nature of risk matrices, i.e. assigning a risk rating / a color (green, yellow, red) based on the product of probability and consequence. As said before, risk matrix application can be taken to a new level when used as a supporting tool for quantitative risk analysis. In this support function, oversimplified classification is not needed anymore. This conclusion will open the door for alternative visualizations where probability and consequence are not multiplied, but kept independent.

--- 2.3 Ranges and System Limits ---

A major flaw of risk matrices that was pointed out in section 1.2 is range compression. Thomas et al. (2013) showed an actual example of a real project where a USD 50 billion blowout with a low probability of occurrence would lead to a lower risk rating than other risk events with higher probabilities and consequences in the range of a few million USD. Such results are obviously artefacts produced by range compression where risk categories do not reflect the ranges needed for the specific purpose, i.e. a USD 50 billion blowout must not have a risk rating of 5 (probability 1, consequence 5) on scale from 1-25 if there is even the slightest possibility of this risk to occur. Meaningful ranges have to be chosen to fit the individual project's needs. Most of the time, non-linear scales will prove more useful than linear ones. For very large ranges, logarithmic scales are an option.

To produce meaningful and comparable results, a risk analysis identifies and evaluates uncertainties for a base with a clear scope and boundaries (system limits) that fits the specific task. This requirement is especially hard to fulfill for companies seeking to establish company-wide risk management standards on a qualitative basis. Typically, risk analysis in a production plant will have different requirements and scopes than a research department or project-oriented teams. There is no single risk matrix that will make them comparable.

--- 2.4 Resolution and Classification ---

Cox (2008a) showed why risk matrices need to fulfill weak consistency, betweenness and consistent coloring. The examples in this paper will have continuous coloring either with iso-contours or independent colors for probability of occurrence and consequence. Resolution in a continuous heat map is infinitely high, hence these risk matrix principles do not apply anymore in their original meaning.

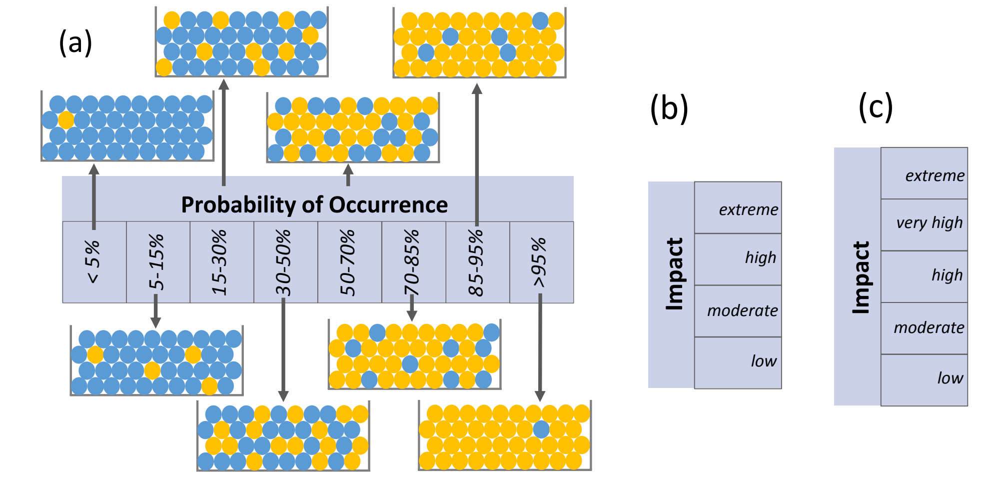

FIGURE 2. Scales for probability of occurrence (a) and consequence (b)+(c). Qualitative risk consequence scales needs additional text explanation to reduce ambiguous interpretations.

The risk matrix will however not be continuous in general. A grid is used to categorize risks in qualitative risk analysis. For the examples in this paper, probability of occurrence will be divided into eight categories (see **Figure 2a**). A finer resolution is needed at the lower end of the scale (e.g. for high consequence risks) because their relative distance is very high (a factor of 2 lies between 5% and 10%). The middle range will be coarser whereas the resolution increases again towards the end. This increase is based on the author's experience that risk assessors want to differentiate between highly probable and almost certain events (e.g. change orders in construction). This scale is only valid for single occurring risks. To take multiple occurring risks into account, other scales have to be used.

Typical risk matrices are symmetrical, i.e. the resolution is identical for probability and consequence. For certain applications, this might be a good choice. However, in this paper a different approach is presented. It was stated that the probability of occurrence will be divided into eight categories. This rather fine resolution was chosen because of the exact meaning of the scale. A probability of occurrence of e.g. 15-30% is unambiguous. For the consequence however, an example with a purely qualitative scale is chosen, e.g. with a scale that goes from "low", "moderate", "high", ("very high",) to "extreme" (**Figure 2b, Figure 2c**). This qualitative scale has to be described carefully to avoid ambiguous ratings, a fine resolution makes this task more difficult. To reduce ambiguity and to avoid a tendency towards the center (which is typical for uneven numbers of categories) a scale with four categories (**Figure 2b**) will be used in this paper.

--- 2.5 Accessible Colors---

Risk Matrices have typically utilized green, yellow and red colors (g-y-r) to classify low, intermediate and high risks. This color scheme is firmly established, questioning it will not be easily accepted. However, many engineering disciplines have started to banish the famous rainbow color map which suffers from the same downsides as the g-y-r scheme. The major flaws of such color maps were pointed out by a number of researchers, e.g. Rogowitz & Treinish (1998) or Borland & Taylor (2007).

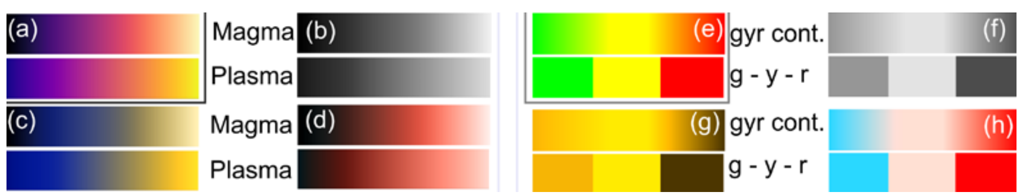

FIGURE 3. (a) Perceptual uniform color maps, (b) meaningful achromatic version, (c) deuteranopia, (d) tritanopia – (e) green-yellow-red color maps, (f) misleading achromatic version, (g) deuteranopia, (h) tritanopia. (Converted pictures were created with eye.side 2011.1.02, www.eyesyde.de)

Figure 3 (e) shows two variations of the g-y-r scheme, continuous (preferred in this paper) and discrete (most widespread risk matrix application). **Figure 3** (f) shows the achromatic version of the same colors. They are actively misleading due to lack of perceptual ordering, e.g. green is more intense than yellow. This is an observation that many risk managers have made after greyscale printing of risk matrices, an example of misleading risk rating colors is given in **Figure 4**. Besides printing, these colors are misleading for people with color vision deficiencies. **Figure 3** (g) and (h) show the same colors as seen by people with deuteranopia (green blindness) and tritanopia (blue blindness) respectively.

Two out of various excellent alternatives, magma and plasma color maps, are presented in **Figure 3** (a). They were developed by van der Walt & Smith (2015) and are perfectly perceptual uniform in their original form (a) and when converted to greyscale (b). Even with color vision deficiency filters, they show a continuous increase of color intensity. In spite of being uncommon, they are a better choice for risk management applications for everyone, with or without color vision deficiency.

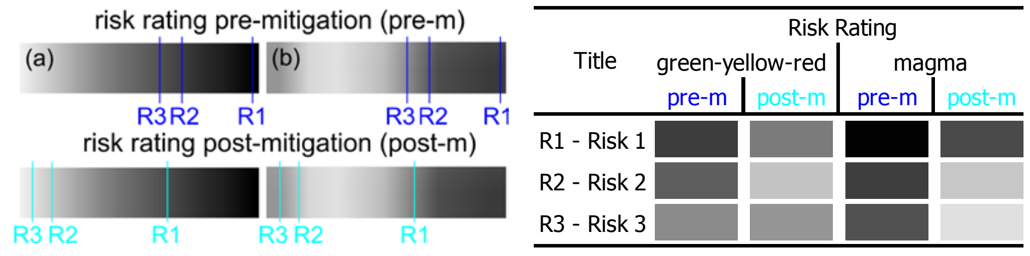

FIGURE 4. Comparison of magma (a) and green-yellow-red (b) color map. Risk rating colors of the latter are not meaningful due to lack of perceptual order.

3. RISK MATRIX EXAMPLES

With the design principles of section 2 at hand, a variety of new risk matrices can be created. Different grids/resolutions should be used for specific tasks, for demonstration purposes however all examples with have the 8 x 4 grid that was introduced in section 2. **Figure 5** shows eight examples of new risk matrix designs. Heatmaps are used instead of discrete coloring. It eliminates the resolution problem and indicates that no category is 100% green, yellow or red, every category summarizes individual characteristics of a heatmap section.

Figure 5a shows the most intuitive evolution of the risk matrix that will easily find its way into practice. Qualitative probability of occurrence scales are replaced with a quantitative one, continuous coloring is used while the familiar green-yellow-red color scheme remains. **Figure 5b** goes one step further

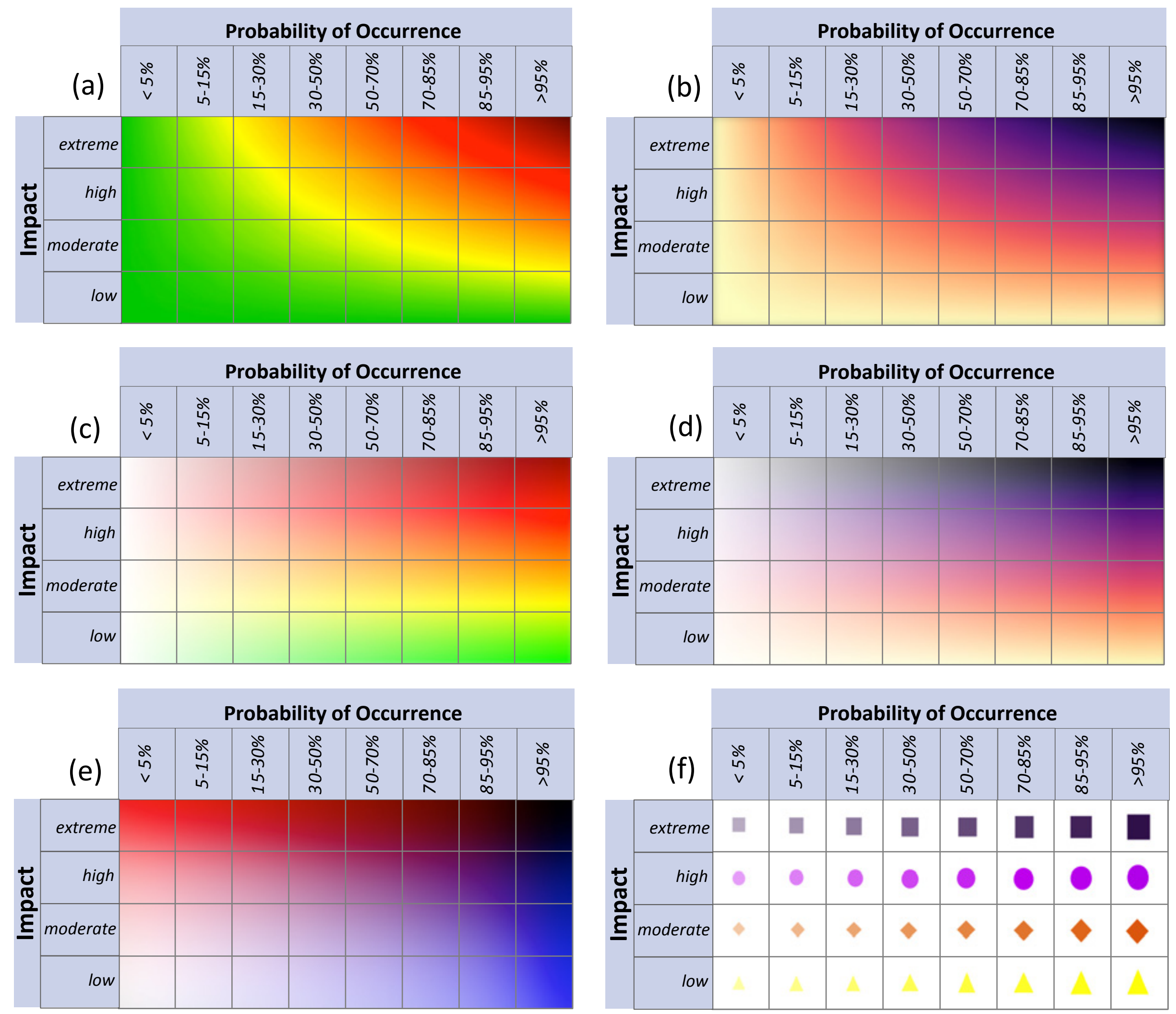

FIGURE 5. Iso-contour g-y-r (a), iso-contour magma (b), transparency g-y-r (c), transparency magma (d), independent axes (e), symbols (g).

by introducing the more advanced magma color map that has proven to be valuable in section 2.4. While being superior in theory, it might be rejected for practical applications.

The next step is to not use the product of probability of occurrence and consequence, but to keep the axes of the matrix independent. By doing this, every point becomes unique, classification becomes more challenging. This blurs the original "ease of use" of the risk matrix, but produces more meaningful results that avoid comparing apples and oranges. **Figure 5c** uses a green – yellow – red coloring to indicate risk consequence. Probability of occurrence is indicated with transparency. Low probability – high consequence events are no longer green, they are still red but with very high transparency which results in light rose or even white. **Figure 5d** shows the same design with the more advanced magma color map.

Instead of using transparency, two individual color maps – one per variable – were merged in **Figure 5e** (white to blue for probability of occurrence, white to red for consequence). Again, every pixel has an individual color which makes classification difficult. Yet, the intensity increase is clearly visible starting from the bottom left to the top right.

Finally, a risk matrix with symbols is shown in **Figure 5f**. Color and symbol indicate consequence, increasing size and decreasing transparency indicate probability of occurrence. This risk matrix is not suited for visualizing various risks in a plot, its purpose is to provide meaningful symbols for risk registers, tables, etc.

4. DISCUSSION

--- 4.1 Strengths and limitations of risk matrices ---

Compared to the traditional risk matrix, three major changes were introduced: First, the use of perceptual uniform colors seems unfamiliar at first, but excels compared to classic green – yellow – red color schemes. Second, continuous coloring reminds us that no two grids in a risk matrix are truly comparable. Third, this effect is further emphasized by separating probability of occurrence from consequence by either separate color maps that are merged to one (**Figure 5e**) or transparency effects (**Figure 5c, d**).

These changes make risk matrices accessible for people with color vision deficiencies and generally result in stronger and more meaningful visualization. But they will also weaken classification capacities which might lead to a reduced area of application. Recent research that was described in section 1.2 of this paper raised strong doubts about the way risk matrices are used in the industry or about risk matrix use in general. The changes introduced in this paper were developed to overcome these shortcomings of traditional risk matrices to re-enable their use.

Risk matrices allow for classification of risks with little effort. This ease of use can (and often does) lead to quick but meaningless results. To avoid this, risk scenarios will have to be carefully developed and described within pre-defined system limits, ideally with more than just a risk title. Ratings should be quantitative whenever possible. This is always the case for probability of occurrence (or frequency) and mostly for consequence. In rare applications (e.g. non-profit) where qualitative scales are used, a detailed description of the scale is needed. Finally, scales and rules for classification cannot be developed universally, they have to be validated and refined for every application. Even then, classification results can and must still be questioned instead of using blind faith. With these principles in mind, risk matrices can be valuable supporting or standalone risk management tools.

--- 4.2 Application for quantitative risk analysis ---

Besides the ability to improve qualitative risk rating, the introduced changes will make risk matrices a powerful visualization tool for the results of a quantitative risk analysis process (Sander et al., 2015). **Figure 6** shows a colormap similar to the ones developed before. While the colormaps in **Figure 5** contained a subjective element (i.e. the ratio between low, moderate, high, extreme that will determine color intensity), the color intensity for fully quantitative application has no ambiguities, it is simply determined by the product of the two axes. The figure – together with other visualizations such as range consequence or tornado diagrams – will allow to separate important from unimportant risks and to visualize the effect of mitigation measures. The left side of **Figure 6** shows risks prior to mitigation, the right side shows the same risks after mitigation. White indicates that risks were not mitigated, blue indicates that mitigation measures were activated.

5. CONCLUSIONS

In spite of their popularity, severe shortcomings of common risk matrix applications were identified recently and summarized in this paper. Based on these findings, we have suggested three adaptions that will prevent oversimplified classification (e.g. viewing high probability – low consequence and low probability – high consequence risks as equal) and improve risk matrix applications. The adaptions are: (1) Use continuous coloring (heat maps), (2) replace green-yellow-red color schemes with perceptually uniform color schemes to improve clarity as well as accessibility for people with color vision deficiencies and (3) view probability and consequence as independent variables. With these changes at hand, a new area of application for risk matrices is to support quantitative risk analysis by visualizing results for the mitigation process. ◆

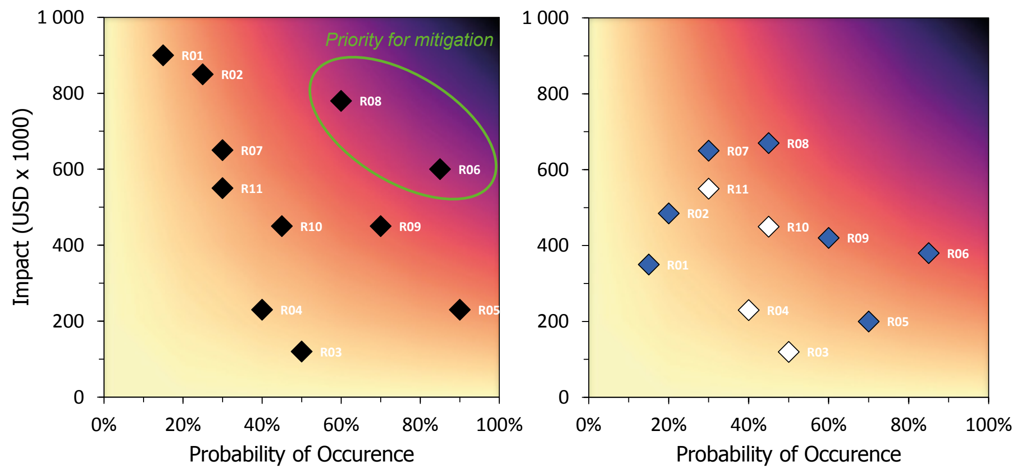

FIGURE 6. Pre-mitigation (left) and post-mitigation risk matrix (right). Mitigation measures are activated for blue risks.

• AUTHORS •

MARTIN ENTACHER is Project Manager at RiskConsult where he focuses on supporting owners and contractors of infrastructure projects in the field of cost and risk management, contract management and project controls. He graduated from Vienna University of Technology (MSc) and obtained a PhD in the field of mechanized tunneling at Montanuniversität Leoben. His research is focused on tunnelling and risk management and has led to multiple international publications, patents and research awards. Martin worked as a construction manager in ground engineering and tunneling at Züblin / Strabag before joining the team of RiskConsult

PHILIP SANDER is Managing Director at RiskConsult and SSP BauConsult and has focused for many years on the successful management of large-scale construction projects. He supports owners and contractors in the areas of cost and risk management, construction and organization processes and software development. Philip studied civil engineering at the Technical University of Dresden (MSc) and received his PhD at the University of Innsbruck (Probabilistic Risk Analysis for Construction Projects). He started his professional career at Grinaker LTA in South Africa. After gaining experience at Hochtief in Essen, Germany and at CPM in Stuttgart, Germany, he moved to Tyrol, Austria in 2006 to join the team of SSP BauConsult and RiskConsult.

• REFERENCES •

Ball, D. J., & Watt, J. (2013). Further Thoughts on the Utility of Risk Matrices. Risk Analysis, 33(11), 2068–2078.

Borland, D., & Taylor, R. M. (2007). Rainbow Color Map (Still) Considered Harmful. IEEE Computer Graphics and Applications, 27(April), 14–17.

Cox, L. A. (2008a). What's wrong with risk matrices? Risk Analysis, 28(2), 497–512.

Cox, L. A. (2008b). Some limitations of "risk = threat x vulnerability x consequence" for risk analysis of terrorist attacks. Risk Analysis, 28(6), 1749–1761.

Cox, L. A. (2009). Improving risk-based decision making for terrorism applications. Risk Analysis, 29(3), 321–335.

Cox, L. A., Babayev, D., & Huber, W. (2005). Some limitations of qualitative risk rating systems. Risk Analysis, 25(3), 651–662.

Duijm, N. J. (2015). Recommendations on the use and design of risk matrices. Safety Science, 76, 21–31.

Elmontsri, M. (2014). Review of the strengths and weaknesses of risk matrices. Journal of Risk Analysis and Crisis Response, 4(1), 49–57.

Gardoni, P., & Murphy, C. (2014). A Scale of Risk. Risk Analysis, 34(7), 1208–1227.

Hubbard, D., & Evans, D. (2010). Problems with scoring methods and ordinal scales in risk assessment. IBM Journal of Research and Development, 54(3), 2:1-2:10.

Hubbard, D. W., & Seiersen, R. (2016). How to Measure Anything in Cybersecurity Risk. Wiley, 304p.

Huber, B., & Cox, T. (2008). Optimal Design of Qualitative Risk Matrices to Classify Quantitative Risks. Proceedings of the Annual Meeting of the Society of Risk Analysis; Boston, Massachusetts.

Levine, E. S. (2012). Improving risk matrices: the advantages of logarithmically scaled axes. Journal of Risk Research, 15(2), 209–222.

Li, J., Bao, C., & Wu, D. (2018). How to Design Rating Schemes of Risk Matrices: A Sequential Updating Approach. Risk Analysis.

Montibeller, G., & von Winterfeldt, D. (2015). Cognitive and Motivational Biases in Decision and Risk Analysis. Risk Analysis, 35(7), 1230–1251.

Ni, H., Chen, A., & Chen, N. (2010). Some extensions on risk matrix approach. Safety Science, 48(10), 1269–1278.

Oliveira, M. D., Bana e Costa, C. A., & Lopes, D. F. (2015). Designing and exploring risk matrices with MACBETH. International Journal of Information Technology & Decision Making, 14, 1–37.

Olson, D. L., & Wu, D. D. (2017). Enterprise Risk Management Models. Springer Texts in Business and Economics Enterprise, 216p.

Rogowitz, B. E., Treinish, L. A. (1998). Data visualization: The end of the rainbow. IEEE Spectrum, 35(12), 52-59.

Sander, P., Moergeli, A., Reilly, J. (2015). Quantitative Risk Analysis – Fallacy of the Single Number. Proceedings of the World Tunnel Congress, Dubrovnik, Croatia.

Thomas, P., Bratvold, R. B., & Bickel, E. (2013). The Risk of Using Risk Matrices. SPE Annual Technical Conference and Exhibition, 56–66.

Van der Walt, S., & Smith, N. (2015). Mpl Colormaps. SciPy, July 6 – 12, Austin, Texas.

Weber, M. (2006). Some Safety Aspects on the Design of Sparger Systems for the. Process Safety Progress, 25(4), 326–330.