RELIABILITY PAPER
Total risk evaluation framework

Zsolt Tibor Kosztyán
Department of Quantitative Methods, University of Pannonia, Veszprém, Hungary
Tibor Csizmadia
Department of Management, University of Pannonia, Veszprém, Hungary
Zoltán Kovács
Department of Supply Chain, University of Pannonia, Veszprém, Hungary, and
István Mihálcz
Department of Quantitative Methods, University of Pannonia, Veszprém, Hungary

Abstract

Purpose – The purpose of this paper is to generalize the traditional risk evaluation methods and to specify a multi-level risk evaluation framework, in order to prepare customized risk evaluation and to enable effectively integrating the elements of risk evaluation.

Design/methodology/approach – A real case study of an electric motor manufacturing company is presented to illustrate the advantages of this new framework compared to the traditional and fuzzy failure mode and effect analysis (FMEA) approaches.

Findings – The essence of the proposed total risk evaluation framework (TREF) is its flexible approach that enables the effective integration of firms' individual requirements by developing tailor-made organizational risk evaluation.

Originality/value – Increasing product/service complexity has led to increasingly complex yet unique organizational operations; as a result, their risk evaluation is a very challenging task. Distinct structures, characteristics and processes within and between organizations require a flexible yet robust approach of evaluating risks efficiently. Most recent risk evaluation approaches are considered to be inadequate due to the lack of flexibility and an inappropriate structure for addressing the unique organizational demands and contextual factors. To address this challenge effectively, taking a crucial step toward customization of risk evaluation.

Keywords Risk evaluation, Flexibility, Multi-level structure

Paper type Research paper

Nomenclature

ANP	Analytic network process	EDAS	Evaluation based on distance from average solution
Cr	Criticality factor		
CI	Consistency index	FMEA	Failure mode and effects analysis
CR	Consistency ratio		
f	Vector of risk factors	Fuzzy FMEA	Fuzzy failure mode and effects analysis

ISO	International Organization for Standardization	TREF	Total risk evaluation framework
K	Invention function	$S(f, w)$	Risk aggregation function
n	Number of risk factors	VIKOR	VIsekriterijumska optimizacija i
RAP	Risk aggregation protocol		KOmpromisno Resenje
RI	Random consistency index	w	Vector of weights
RPN	Risk priority number	$(W1) - (W3)$	Warning rules
T	Threshold vector	WS	Warning system
TOPSIS	Technique for order preference by similarity to ideal solution		

1. Introduction
To respond to exponentially growing stakeholder and societal demands, companies have had to develop solutions for complex operations, which are highly sensitive to the external and internal organizational environment. To ensure smooth operation, it is necessary to understand the hazards and risks associated as well as their mitigation. Risk evaluation is used in many application areas, and many frameworks and methods have been proposed in practice and in the scientific literature. Conventional risk evaluation approaches, however, ignore the fact that many contemporary organizational and process components or failure effects across hierarchical levels of a system are inherently complex (O'Keeffe *et al.*, 2015; Pasman *et al.*, 2014), and they are not sufficient to address the continuously changing organizational demands. Such situations call for new approaches, suggesting flexible and adaptive risk evaluation methods (Aven, 2016; Reiman *et al.*, 2015) that change to fit the environment and situational factors of the organization. As Kanes *et al.* (2017) stated, it is important to focus research on the area of flexibility in risk evaluation as a way forward for improving current risk evaluation methodologies. In light of the aforementioned, this paper proposes a multi-level risk evaluation framework, where flexibility is given a more prominent role than in the current state of the art in the application area. It helps us to think differently about risk evaluation – as a process that is recursive rather than linear, flexible rather than rigid and pluralist not binary (O'Keeffe *et al.*, 2015). Flexibility in risk evaluation is important in the following areas: rating scale, number of risk factors, risk aggregation and warning system (WS); however, their integration into one framework has yet not been developed in the literature. The lack of the flexibility of risk evaluation raises difficulties in integration, customization and adaptivity of risk management systems.

Various rating scales have been developed for risk evaluation in the literature, which can be divided into two categories as predefined or invariant scales according to the stage of evaluation. In the case of invariant scales, linguistic scales were mostly used with defined levels (e.g. 3, 4 or 5 levels), and the risk was considered an occurrence of high risk levels (Gauthier *et al.*, 2018). Risk values can also be a result of a pairwise comparison (Merrick *et al.*, 2005), where risk effects are compared by their factors. In these cases, based on the results, risk effects can be ranked and management can treat the most risky effects. However, in this case, the results of risk values are very difficult to interpret. In the implementation of predefined scales, experts use previously defined ordinal scales, for example, the traditional 10-point scales (Liu *et al.*, 2013, 2012; Liu and Tsai, 2012), probabilistic scales with values between [0, 1] (Gauthier *et al.*, 2018; Malekitabar *et al.*, 2018) or mixed scales such as the Fine–Kinney method (Kinney and Wiruth, 1976). The proposed framework allows both types of risk value to be used and makes it possible for matching the aggregation function to its rating scale in a consistent way.

Methods in the literature define the degree of risk depending on a fixed number of factors. In the failure mode and effect analysis (FMEA) method, the risk value is calculated based on

the occurrence, severity and detectability parameters (Fattahi and Khalilzadeh, 2018), whereas the Fine–Kinney (practical risk assessment) method calculates risk depending on the likelihood of occurrence, exposure to the event and possible consequence parameters (Kinney and Wiruth, 1976). Various extensions of the number of risk factors have been introduced in the literature (see e.g. Karasan *et al.*, 2018; Yousefi *et al.*, 2018). However, they are limited to a fixed number of factors. Unlike the literature, we propose a risk evaluation framework that flexibly considers the impacts of arbitrary $(n \geq 2)$ number of risk factors to cope with the deficiency mentioned earlier.

Several methods and analyses (such as MacKenzie, 2014; Azadeh-Fard *et al.*, 2015; Panchal *et al.*, 2019c, 2019a) have been proposed for aggregating risks. Summarizing available information into a single number is a difficult and sensitive issue (MacKenzie, 2014), which requires an analyst to carefully select and utilize constituent elements, factors, weights and algebraic operations (Ni *et al.*, 2010; Azadeh-Fard *et al.*, 2015). The proposed total risk evaluation framework (TREF) works with a risk aggregation protocol (RAP); this is a triplet, which includes three elements such as factors, weights and aggregation functions. The aggregation can be implemented at different levels such as factor, effect, mode, process, process area and organization. In addition, various domains such as health and safety, quality or environment can be considered in the course of aggregation. This hierarchical aggregation procedure is flexible; items such as the properties of available information, mathematical expression and details required by the users can determine the exact elements and methods that are best suited to the organization.

Warnings also play a vital role in risk evaluation (Khan *et al.*, 2015; Øien *et al.*, 2011). Conventional risk evaluation has the disadvantage of being rigorous (Kalantarnia *et al.*, 2009); repeatedly adopting a single index (Zheng *et al.*, 2012) or a list of warning indicators (Øien *et al.*, 2011) to signal warning events fails to capture all meaningful failures. There have been many efforts to develop a WS for risk assessment (Ilangkumaran *et al.*, 2015; Øien *et al.*, 2011; Zheng *et al.*, 2012), but none of them addresses warning events from both factors and levels of aggregation, that is, effect, mode and process, in order to capture comprehensive failure identification. The principal feature of the WS of TREF is its flexibility. TREF extends the literature by presenting a flexible WS considering risk values from the levels of factors, domains, effects, modes, processes and the organization. This feature is a novel tool to specify unique warning rules for each risk factor separately in each level.

In summary, the main contribution of this study to the literature is that it is the first step toward customization of risk evaluation. In other words, TREF is a flexible risk evaluation framework, which can be tailored to the specific needs of companies. TREF is composed of the following:

(1) An arbitrary number $(n \geq 2)$ and weight of risk factors can be set by domain in risk evaluation.

(2) To obtain a domain-specific risk value, the attributes of different management systems (e.g. health and safety, quality or environment) can be considered in risk evaluation.

(3) The paper develops a flexible multi-level risk aggregation protocol, which can be implemented at different levels, adapted to the organizational needs.

(4) A flexible warning system can specify unique rules for warnings in each level.

The rest of this paper is organized as follows. Section 2 provides a brief review of risk evaluation and its shortcomings. In Section 3, the proposed new risk evaluation framework, the TREF, is presented. Section 4 presents how the proposed TREF can be applied in risk evaluation. A real case serves to demonstrate the applicability and practicality of the

proposed TREF in Section 5, including comparisons with two most commonly used risk evaluation methods: traditional and fuzzy FMEA. Finally, we draw conclusions and make suggestions for future research in Section 6.

2. Related work

Risk evaluation is the process of assessing the impact and likelihood of identified risks based on Chang and Wen (2010) and Hansson and Aven (2014). The main aim of risk evaluation is to determine the importance of risks and to prioritize them according to their effects on systems, processes, designs and/or services for further attention and action (Klinke and Renn, 2002). In other words, this process determines which risk source warrants a response. The need for this process is based on the fact that organizations, processes and projects face a large number of risks, each with different effects; thus, it may be impractical or even impossible to manage them all because of time and resource constraints.

Risk evaluation is used in many application areas, and many frameworks, methods and techniques have been proposed in practice and in the scientific literature. Liu *et al.* (2013) conducted an analysis of 75 papers on the subject of risk evaluation. They systematically classified the existing literature and concluded that the FMEA approaches introduced in the last decades can be divided into three categories according to their failure mode prioritization methods: multicriteria decision-making, mathematical programming and integrated approaches. In addition, it can be observed from the surveyed literature that the fuzzy rule-based system is the most popular method for prioritizing failure modes (Liu *et al.*, 2013; Panchal and Kumar, 2017; Panchal *et al.*, 2018a). Conventional risk evaluation approaches nevertheless ignore the fact that many contemporary organizational and process components or failure effects across hierarchical levels of a system are inherently complex (O'Keeffe *et al.*, 2015; Pasman *et al.*, 2014), and they are not sufficient to explain all that can go wrong. Such situations call for new approaches, suggesting the need to develop flexible and adaptive risk evaluation methods (Aven, 2016; Reiman *et al.*, 2015) that change to fit the environmental and situational factors of the organization. As Kanes *et al.* (2017) stated, it is important to focus on the area of flexible risk evaluation, as a way forward for improving current risk evaluation methodologies. O'Keeffe and his team also emphasized that a risk evaluation process should be recursive rather than linear, flexible rather than rigid and pluralist not binary (O'Keeffe *et al.*, 2015). Such a situation calls for different approaches and methods, and it is a challenge for the risk field to develop suitable frameworks and tools for this purpose (Aven and Zio, 2014; SRA, 2015). As a result of a shift in risk evaluation thinking from traditional and rigid to flexible and adaptive attributes, new risk evaluation methods should be developed where flexibility is one of the most important characteristics. Flexibility in risk evaluation can be implemented in the following areas: scale, number of factors, aggregation and WS.

Various scales have been developed for risk evaluation in the literature; they can be divided into two categories of predefined or invariant scales according to the state of evaluation. In the case of invariant scales, in the early stages of risk evaluation, scale was not used; risk evaluation was performed via percentage of occurrence (Etherton and Myers, 1990). Later, linguistic scales were used with 3–5 distinguished levels, and the assessment was made by the evaluation team's top ratings percentage (Gauthier *et al.*, 2018; ISO 12100, 2010). Linguistic scales Merrick *et al.* (2005) use the pairwise comparison instead of percentage. After the comparison, we can determine the ranking order of all alternatives and select the best ones from among a set of feasible alternatives. The main challenge of this approach is to interpret the resulting risk values. Indeed, regardless of whether we have compared risky or less risky effects, the results will fluctuate around the same value.

Another approach is to use predefined scales for all factors. Before performing the evaluations, the appropriate numeric scales were defined first in the failure analysis (Liu *et al.*, 2013). Various scoring guidelines exist; for example, Goodman as cited by Silva *et al.* (2014)

developed the 10-point scales for evaluating the failure modes with respect to each risk factor. Similarly, Lolli et al. (2015) developed an evaluation scale for assessing the three risk factors such as the widely known FMEA. In some cases, mixed scales can be found, as in Fine–Kinney (Kinney and Wiruth, 1976), where for likelihood and exposure [0.1,10] is used and for consequence [1,100] is used. Both approaches can be used in risk evaluation; however, predefined scales, in particular the FMEA method using the product formula, were the most common (Liu et al., 2013).

Methods developed in the literature define the degree of risk depending on a fixed number of factors. In the traditional FMEA method, the risk value is calculated based on the occurrence, severity and detectability parameters (Liu et al., 2013; Fattahi and Khalilzadeh, 2018). The Fine–Kinney method calculates risk depending on the likelihood of occurrence, exposure and consequence parameters (Kinney and Wiruth, 1976). Some extensions of the number of risk factors have been introduced in the literature. Karasan et al. (2018) extend the number of factors, calculating risk based on severity, probability, frequency and detectability values. Ouédraogo et al. (2011) increased the factors to five: risk perception, impact of hazard, research specificities, hazard detectability and probability of occurrence of accident. Maheswaran and Loganathan (2013) proposed four risk factors including severity, occurrence, detection and protection. Yousefi et al. (2018) considered two additional factors including cost and duration of treatment in addition to severity, occurrence and detection. These methods, however, are limited to a fixed number of risk factors. In addition, during our literature investigation, we found that authors calculate with risk factors, as they are independent (Liu et al., 2013). One of the possible causes of ignoring additional risk factors is that their dependence should be addressed. These issues call for new solutions that can address the dependence of risk factors and an arbitrary number of risk factors.

Several methods and analyses have been proposed for aggregating risk. Traditionally, FMEA uses the risk priority number (RPN) to evaluate the risk of failure. The occurrence factor measures the likelihood that a failure mode occurs. The severity is the expected consequence of failure. The ability to recognize an error before it affects customers is measured by the detection factor. Scales based on guidelines for usage (such as Fine–Kinney and FMEA) and for evaluation/aggregation require different functions, such as additive, average, product, geometrical mean, logarithmic (Malekitabar et al., 2018), but the most common is the FMEA method with product formula (Liu et al., 2013). The multiplication of these factors generates the RPN, and the aggregation is performed solely at the factor level. Detailed procedures for carrying out an FMEA have been documented in Stamatis (2003) and Tay and Lim (2006). The traditional FMEA has proven to be one of the most important early preventive methods (Liu et al., 2013, 2014; Silva et al., 2014), whereas the traditional RPN method has been criticized in the literature (see the summary in Liu et al. (2013); Lolli et al. (2015); Malekitabar et al. (2018)). Numerous alternative approaches have been proposed to overcome the shortcomings of traditional FMEA. It can be observed from one of the most recent reviews of FMEA conducted by Liu et al. (2013) that the fuzzy rule-based system is the most popular method for prioritizing failure modes. The fuzzy rule-based FMEA approach uses linguistic variables to prioritize failures in a system to describe the severity, detection and occurrence as the riskiness of failure (Tay and Lim, 2006; Petrović et al., 2014; Bowles and Peláez, 1995). However, the most commonly used membership functions are the triangular and trapezoidal (Riahi et al., 2012). An advantage of using fuzzy rule-based FMEA for risk evaluation is that the resulting evaluation becomes qualitative and has the ability to model uncertain and ambiguous information. A disadvantage of fuzzy rule-based FMEA approaches is that they can produce erroneous results if analysts do not have a sufficiently deep understanding of the system. In addition, similarly to traditional FMEA, fuzzy rule-based FMEA aggregates only at the factor level. Other aggregation techniques have also been proposed in the literature, for example, geometric mean (see e.g. Kokangl et al.,

2017; Maheswaran and Loganathan, 2013; Wang et al., 2009), median Karasan et al. (2018) and radial distance Malekitabar et al. (2018). The weighted geometric mean is also applied in the analytic hierarchy process (AHP) (Braglia and Bevilacqua, 2000) or analytic network process (ANP) (Liu and Tsai, 2012; Torabi et al., 2014; Wang et al., 2018). The AHP/ANP enables the decomposition of elements into a hierarchy and calculates weights for the risk factors. In the AHP, each element in the hierarchy is considered to be independent of all the others. However, ANP does not require independence among elements, so it can be used as an effective tool also in the case of interdependency (Saaty, 2004; Wang et al., 2018).

In addition, the authors emphasize a remarkable shift toward integrated methods for ranking failure modes when aiming at accurate risk evaluation. For instance, fuzzy evidential reasoning is integrated with grey theory (Chang et al., 1999; Liu et al., 2011), fuzzy TOPSIS (technique for order preference by similarity to ideal solution) with fuzzy AHP (Kutlu and Ekmekçioğlu, 2012) and VIKOR (VIsekriterijumska optimizacija i KOmpromisno Resenje) or EDAS (evaluation based on distance from average solution) (Panchal et al., 2019c) with fuzzy logic (Liu et al., 2012; Panchal et al., 2019b; Panchal and Srivastava, 2019) and grey techniques (Panchal and Kumar, 2016; Panchal et al., 2018b; Panchal and Srivastava, 2019). There is a trend toward using more than one method to enhance the efficacy and empirical validity of risk evaluation results (Liu et al., 2013). Recent research (Lolli et al., 2015; Liu et al., 2014) also shows a shift toward integrated methods (e.g. ANP (dos Santos et al., 2015; Zammori and Gabbrielli, 2012) has been combined with other models), so that synergies can be maximized. Liu et al. (2013); Shaker et al. (2019) conclude that objective and combination weighting methods should be applied in risk evaluation because they evaluate relative importance objectively without decision-makers. However, some doubts remain concerning the applicability of integrated methods to real-life circumstances, for example, the need to add risk factors to the determination of risk priority of failure modes (Liu et al., 2013) and the need to support the aggregation of risk levels from different domains. Considering risk effects in different domains is important because the same source of hazards often causes risks in multiple management areas with different levels of relevance (Pasman et al., 2014). Therefore, the sources of hazards describing the possible risk effects in different management system areas (e.g. ISO 9001 (2015); ISO 14001 (2015) and ISO 45001 (2017) (previously OHSAS, 18000) should be considered and developed holistically and cohesively (Abad et al., 2014; Asif et al., 2013; Bernardo, 2014; de Oliveira, 2013; Rebelo et al., 2016). Domains such as health and safety, quality or environment can be considered in risk evaluation with different weights. To conclude, priorities and demands can be different by domains, which calls for flexible risk aggregation.

Warnings play a vital role in risk evaluation (Khan et al., 2015; Øien et al., 2011). Conventional risk evaluation has the disadvantage of having rigor (Kalantarnia et al., 2009), repeatedly adopting a single index (Zheng et al., 2012) or a list of warning indicators (Øien et al., 2011) to signal warning events and failing to capture meaningful failures. There have been many efforts to develop the WS of risk evaluation. Ilangkumaran et al. (2015) proposed a hybrid technique (Liu et al., 2015; Panchal et al., 2019a) for assessing work safety in hot environments including a warning rating and safety grade at the risk factor level. Øien et al. (2011) have developed a set of risk indicators that can provide warnings about potential major accidents. Zheng et al. (2012) proposed an early warning rating system for hot and humid environments calculating safety indexes at the factor and subfactor levels. In addition, Xu et al. (2002) suggested two levels of warnings. In the scientific literature, the risk hierarchy is occasionally mixed with risk level; for example, Chen et al. (2012); Manuele (2005) use the action levels as risk hierarchies, and no real hierarchy levels are used. This summary shows that methods developed in the literature do not address warning events originating from multi-levels such as factor, effect, mode and process in order to specify unique warning rules for each risk factor separately in each level.

In conclusion, based on the aforementioned works, the risk evaluation practice of companies is adapted in line with the limits of the existing rigorous methods. Therefore, we

propose a flexible risk evaluation framework, which can be tailored to the specific needs of companies. We found that a fixed number of risk factors mostly from one domain were included only in risk evaluation. Our proposed framework makes it possible for users to set an arbitrary number of risk factors, which can be different in different risk domains. Additional advantages of TREF are that it considers the importance of risk factors by domain and can handle the dependence of risk factors. In addition, most methods aggregate risk value at the factor level and the same is true for signal warnings. Instead, TREF includes a flexible RAP and a flexible WS. The flexible RAP makes it possible for users to use different risk factors, weights and aggregation functions at different levels, adapted to the needs of companies. Flexible WS can specify unique warnings for each single risk factor separately.

3. Mathematical background
The TREF includes risk evaluation, risk assessment and schedules of corrective/preventive actions. This paper focuses only on the evaluation of the risk effects based on a novel formula.

3.1 Formal description of calculating risk values

Definition 1. Let $f = [f_1, f_2, \ldots, f_n]^T$, $(n \geq 2)$ be the vector of risk factors, and let $w = [w_1, w_2, \ldots, w_n]^T$, $(n \geq 2)$ be the weight vector of risk factors $(w_i \in \mathbb{R}^+)$. Denote $r = S(f, w)$ as a **resulting risk value**, where S is a monotonous aggregation function. Denote (f, w, S) as the **risk aggregation protocol** (RAP).

Remark 1. Usually we assume that risk factors f_i and f_j, $(i \neq j)$ are independent of each other. However, the proposed RAP does not require their independence.

The proposed RAP can integrate the traditional FMEA, Fuzzy FMEA and the Fine–Kinney risk evaluation methods. RAP generalizes these three types of methods; therefore, they can be considered special cases of the proposed RAP.

Example 1. In the case of traditional FMEA, $n = 3$, $w_i := 1$, $f_i \in \{, 2, \ldots, 10\}$, $i := 1, \ldots, n$, $S := \prod_{i:=1}^{n} f_i$.

Example 2. In the case of Fuzzy FMEA, $n = 3$, $w_i := 1$, $f_i := \tau_i(x)$, $\tau_i(x) : I \to [0, 1]$ is the so-called membership function, $i := 1, \ldots, n$, $S(f, 1) := \prod_{i:=1}^{n} \int_I f_i \mathrm{d}x = \prod_{i:=1}^{n} \int_I \tau_i(x) \mathrm{d}x$.

Example 3. In the case of the Fine–Kinney approach, $n = 3$, $w_i := 1$, $f_1 \in \{0.1, 0.2, \ldots, 10.0\}$ (likelihood of occurrence), $f_2 \in \{0.5, 1.0, \ldots, 10.0\}$ (exposure factor), $f_3 \in \{1.0, 2.0, \ldots, 100.0\}$ (factors of possible consequences), $S := \prod_{i:=1}^{n} f_i$.

To include weights, AHP/ANP can be integrated into the traditional FMEA, Fuzzy FMEA and Fine–Kinney methods. In addition, the proposed RAP allows us to consider arbitrary (more or less than three) risk factors.

Example 4. In the case study, $n \in \{2, 3, 4, 5, 6\}$, $w_i \in \mathbb{R}^+$, $f_i \in \{1, 2, \ldots, 10\}$, $\sum_{i:=1}^{n} w_i = 1$, $i := 1, \ldots, n$, and we used four types of functions:

(1) $S_1(f, w) = \prod_{i:=1}^{n} f_i^{w_i}$ is the weighted geometric mean of risk factors.

(2) $S_2(f, 1) = \max(\{f_1, \ldots, f_n\})$ is the maximum value of risk factors.

(3) $S_3(f, w) = \text{Median}(\{f, w\})$ is the weighted median of risk factors.

(4) $S_4(f, w) = \sqrt{\sum_{i:=1}^{n} w_i f_i^2}$ is the weighted radial distance of risk factors.

In the case of $w_i = 1/n$, the aggregation functions S_1, S_3 and S_4 produce the unweighted geometric mean, unweighted median and unweighted radial distance of risk factors.

> *Remark 2.* One of the main advantages of the proposed RAP is that the number of risk factors can be different, for example, for each risk domain and for each risk mode. In the extreme case, TREF allows us to ignore the evaluation of a risk mode if it does not have an effect on a domain. To compare different kinds of resulting risk values that are based on different numbers of risk factors, the choice of an adequate aggregation function is one of the most important topics. According to Calvo *et al.* (2002) and Beliakov *et al.* (2008), an aggregation function should produce symmetric result for symmetric inputs and should be continuous if the inputs are continuous. Malekitabar *et al.* (2018) and others criticized FMEA because production as aggregation is a very (left-side) asymmetric function; therefore, the risks are often underestimated. In contrast to Malekitabar *et al.* (2018), who suggest radial distance as an aggregation function, we suggest a weighted median if the resulting risk factors are to be compared or aggregated. The (weighted) median maintains the symmetry as much as possible not only for two factors, which were investigated by Malekitabar *et al.* (2018), but also for more than two factors (see Figure 1), where two (see Figure 1)(a–c) and six (see Figure 1)(d–f) are aggregated. Factors followed discrete uniform distributions $f_i \in 1, 2, \ldots, 10$.

Nevertheless, the (weighted) median and (weighted) geometric mean can be robust to outliers, which is commonly referred to as a good feature of an aggregating function, except for the aggregation of risks, where the high value of a risk factor may require intervention. This problem can only be solved with a multi-level WS, where not only the aggregated risk values but also the risk factors are considered (see Section 3.2).

> *Remark 3.* It is important to note that the proposed risk aggregation protocol does not require existing (predefined) scales (see Section 2). Scale values can be a result of a pairwise comparison (see e.g. Merrick *et al.*, 2005).

Applying the risk aggregation protocol iteratively, the risk values can be specified in a higher hierarchy level.

> *Definition 2.* Let $(R^{(N)}, W^{(N)}, S)$, $(R^{(N-1)}, W^{(N-1)}, S)$ be risk aggregation protocols. Denote $\text{TRPN}_i^{(N)} = R_i^{(N)} = S(R_i^{(N-1)}, W_i^{(N-1)})$ as the **total risk priority number** i in the hierarchy level N.

The proposed framework provides the opportunity to evaluate a failure mode through different types of aspects. The effects of a failure mode can have impacts on different domains such as quality, environment, health and safety as well as others. These effects are usually evaluated separately if companies follow different types of standards (such as ISO 9001, 14001, 45001). However, in addition to the separated evaluations, aggregated risk values can also be important for the company, which can include these aspects. Considering these aspects, a weighted risk value can be calculated regarding the iterative formula.

Similarly to this bottom-up calculation mode, risk values can be specified for all hierarchy levels.

At the hierarchy level, we can say:

(1) The *risk factors* are aggregated: $\text{TRPN}_d^{(1)} = R_d^{(1)} = S(R_d^{(0)}, W_d^{(0)}) = S(f, w)$, and they specify the *risk domain values d*.
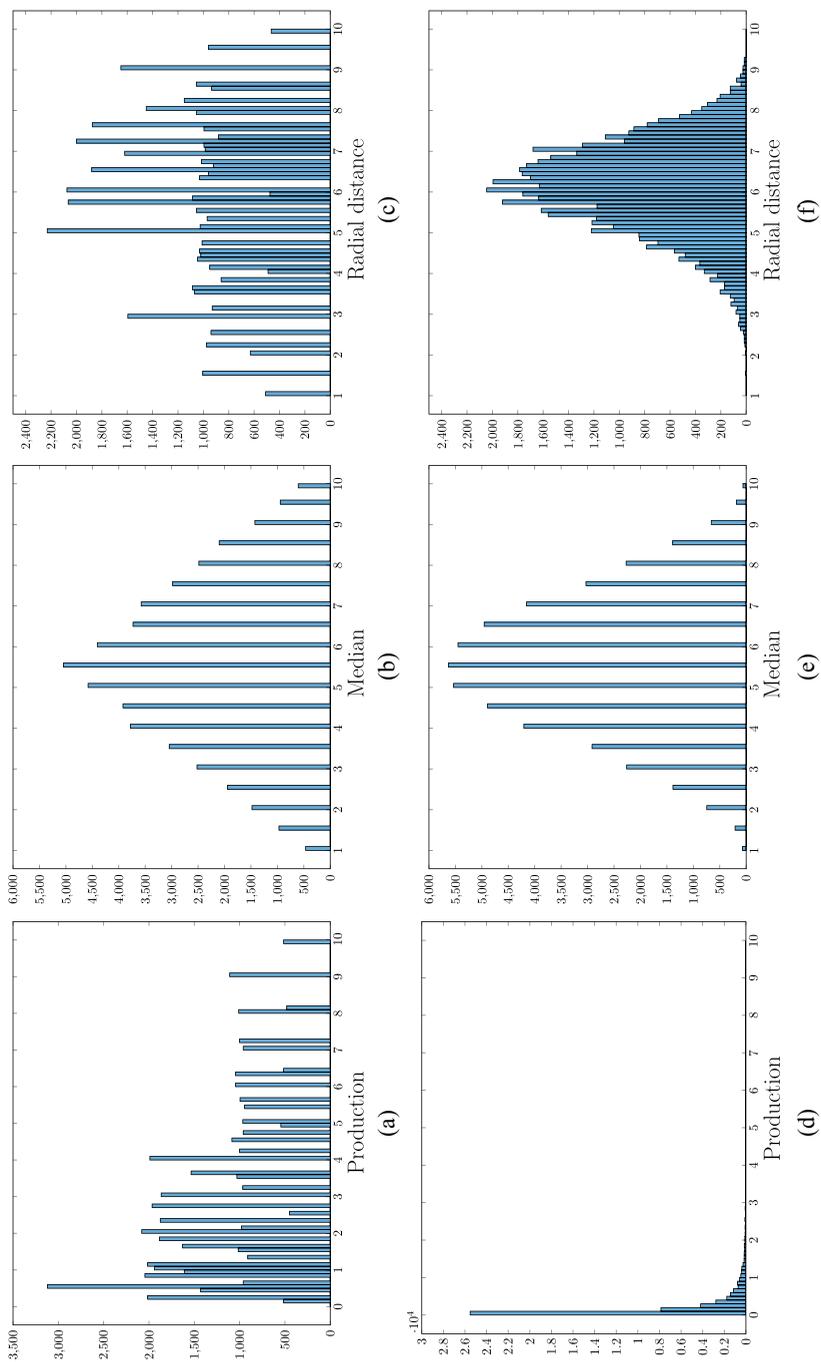
(2) The *risk domain values* are aggregated: $\text{TRPN}_e^{(2)} = R_e^{(2)} = S(R_e^{(1)}, W_e^{(1)})$, and they specify the *risk value for risk effect e.*

(3) The *risk values of risk effects* are aggregated: $\text{TRPN}_m^{(3)} = R_m^{(3)} = S(R_m^{(2)}, W_m^{(2)})$, and they specify the *risk value for failure modes m.*

(4) The *risk values of failure modes* are aggregated: $\text{TRPN}_p^{(4)} = R_p^{(4)} = S(R_p^{(3)}, W_p^{(3)})$, and they specify the *risk value for process p.*

(5) The *risk values of processes* are aggregated: $\text{TRPN}_a^{(5)} = R_a^{(5)} = S(R_a^{(4)}, W_a^{(4)})$, and they specify the *risk value for the process area a.*

(6) The *risk values of process areas* are aggregated: $\text{TRPN}_M^{(6)} = R_M^{(6)} = S(R_M^{(5)}, W_M^{(5)})$, and they specify the *risk value for the main process M.*

(7) The *risk values of main processes* are aggregated: $\text{TRPN}^{(7)} = R^{(7)} = S(R^{(6)}, W^{(6)})$, and they specify the *risk value in the organizational level.*

In the proposed framework, an arbitrary number of factors (but at least two) can be specified for each single risk domain. The compatibility between the TREF and the traditional FMEA or fuzzy FMEA can also be realized if severity, occurrence and detection are considered as risk factors. Similarly, compatibility between TREF and Fine–Kinney can also be realized if the applied risk factors are likelihood of occurrence, exposure and consequence. Beyond the traditional methods, additional risk factors can be specified, for example, control, information and range (see Section 5.) The number of risk factors follows the company needs and the nature of the risk domains. The TREF also proposes an extra factor, criticality, to allow the risk evaluation team to specify corrective/preventive actions. Risk factors can be evaluated and aggregated by different domains (e.g. quality, environment, health and safety) that cover different facets of corporate systems. The potential failure modes can be identified, as can their causes and effects (see Figure 3). Organizational level processes (e.g. production, sales) can be divided into subprocesses, sub-subprocesses and so on Certa *et al.* (2017). This decomposition may go on to an arbitrary depth in the hierarchy depending on the nature and complexity of organizational processes. Nevertheless, hierarchy level is only one option to use the proposed iterative formula. If an acyclic graph or chain of causes, failure modes and effects are specified (see Figure 3), the risk values for common causes and common effects can also be specified.

3.2 The warning system
The WS signals to the risk evaluation team where critical failures are, and this team can see the general conditions of the processes. The WS considers risk values at all levels. As with the calculation of TRPNs, the specification of the WS follows the bottom-up conception. Corrective/preventive actions are scheduled if a risk factor is not lower than a threshold *W*1, but also corrective/preventive actions are scheduled if the aggregated value is not lower than a threshold *W*2. The TREF proposes an extra factor, criticality, to allow the risk evaluation team to specify corrective/preventive actions *W*3, even if the aggregated risk value is lower than the specified threshold. If its value is 1, corrective or preventive actions should be specified. However, if its value is 0, corrective or preventive actions can be specified because both the risk factors and/or the aggregated risk value can be higher than the thresholds. The criticality factor produces another flexibility for the team to override the evaluation and specify preventive tasks for the events that are not risky but that may be potentially risky events (e.g. nonquantifiable risks and difficultly quantifiable customer expectations, or even their possible changes) and should be evaluated independently from other risk factors.

Formally, corrective/preventive actions can be prescribed in three ways:

$$K_i^{(N)} = \begin{cases} 1, & R_i^{(N-1)} \geq T_i^{(N-1)} \\ 0, & otherwise \end{cases} \tag{1}$$

Definition 3. Let $(R^{(N)}, W^{(N)}, S)$ and $(R^{(N-1)}, W^{(N-1)}, S)$ $(N \geq 1)$ be risk aggregation protocols. Additionally, denote $\mathrm{Cr}^{(N-1)} \in \{0,1\}$ as the criticality value in hierarchy level $N-1$. Let $T^{(N)}, T^{(N-1)}$ be threshold vectors, where $\forall i,j, T_i^{(N-1)}, T_j^{(N)} \in \mathbb{R}^+$. Denote the invention function in level N for factor i.
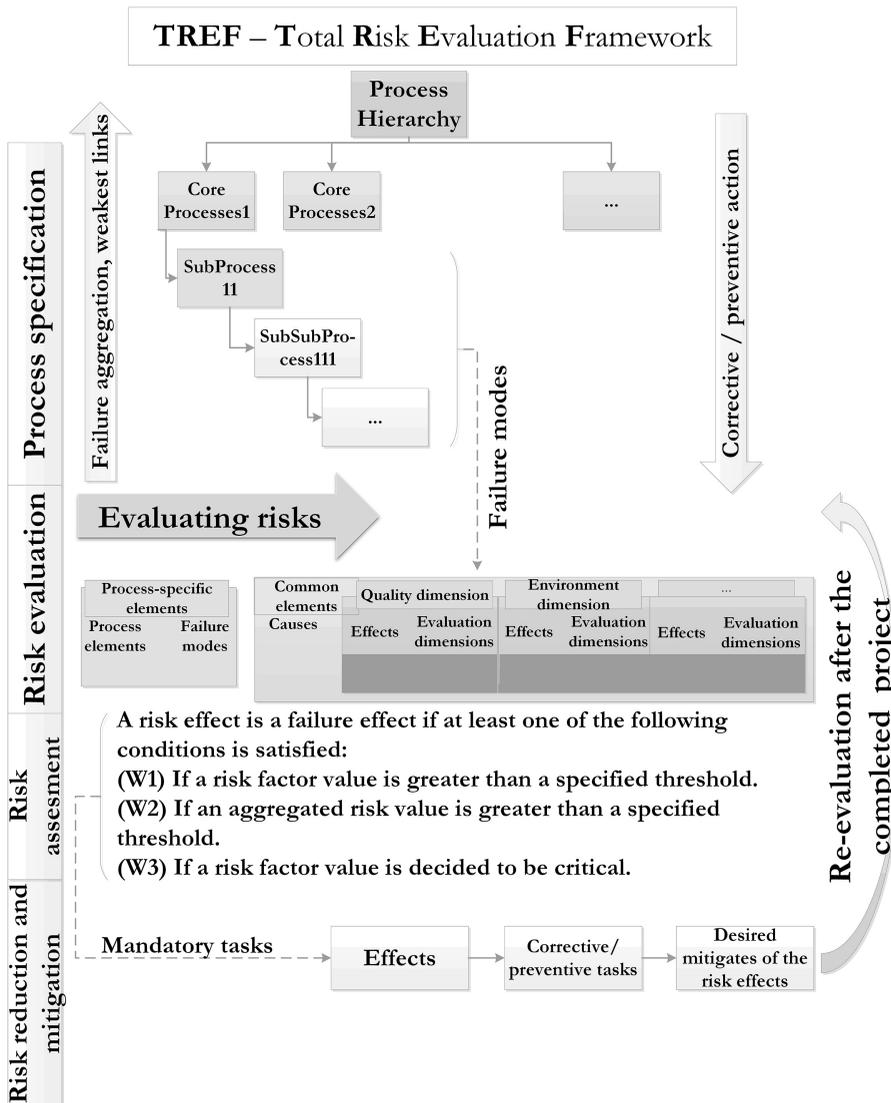
Figure 2.
The proposed total risk
evaluation framework
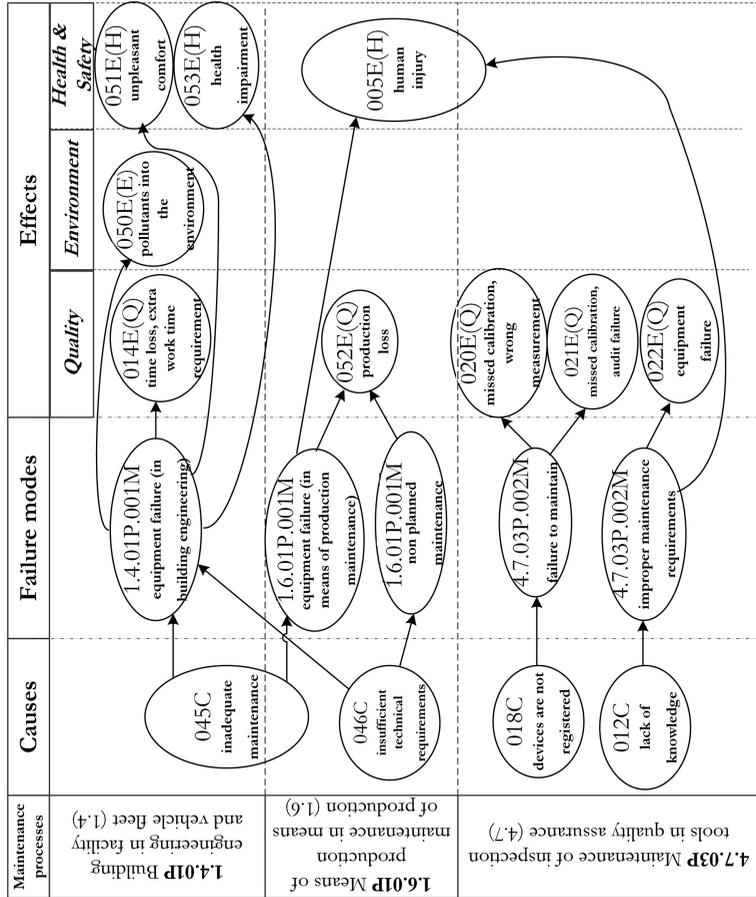
Figure 3.
The TREF graph for
evaluating the risks of
maintenance
processes: the chain of
causes, failure modes
and effects

A **warning event** has occurred if

(W1) $\sum_i K_i^{(N-1)} \geq n^{(N-1)}$ (at least $n^{(N-1)}$ of risk factors are not lower than the specified
threshold);

(W2) $\sum_j K_j^{(N)} \geq n^{(N)}$ (at least $n^{(N)}$ aggregated risk values are not lower than the specified
threshold);

(W3) $Cr^{(N-1)} := 1$ (a risk factor is decided as critical).

The thresholds and the rule of thresholds can be specified as arbitrary, based on the company experts. Generally, warning thresholds are specified based on former experiences, but standards can also provide a threshold. (In our case study, because the company had to follow more than one standard requirement, the minimum value of the experts' opinions was the threshold.) In addition, the dependence of risk factors can also be addressed by specifying different thresholds for each single risk factor separately.

Definition 4. We can say that a (risk) effect is a **failure effect** if at least one of the conditions $W1$–$W3$ is satisfied.

If TRPNs are calculated for the total process tree (see Figure 2), thresholds should be specified for all levels.

4. The process of the application of the proposed framework
The proposed framework has four stages (see Figure 2). The first stage (process specification) is detailed in Section 4.1 Phase 1. The next step, the risk evaluation, is detailed in Section 4.1 Phase 2 and Phase 3, while stage 3 (risk assessment) is detailed in Section 4.2. The risk reduction and mitigation (stage 4) are detailed in Section 4.3.

4.1 Process of risk evaluation
4.1.1 Phase 1: logic planning. The process hierarchy, including the core processes, subprocesses and their subprocesses and so on (see Figure 2), the process-specific elements and failure modes and the chain of causes and risk effects based on their domains should be specified before the proposed TREF is used. This process hierarchy helps us to obtain the failure modes. The proposed framework shows us where we should concentrate efforts to correct and improve processes.

Instead of the process hierarchy, a (process) graph can also be specified (see, e.g. Figure 3). To apply the proposed recursive risk evaluation process, the only requirement for the logic network is that the graph be acyclic. In addition, in the process hierarchy, further arbitrary acyclic graphs (furthermore, the TREF graph (see, e.g. Figure 3) can be specified to better fit changing company needs.

4.1.2 Phase 2: defining factors and scales. At the beginning of Phase 2, there are two options. Either scales for evaluating risk factors are identified (1) or invariant scale methods are used (see e.g. Merrick *et al.*, 2005; Gauthier *et al.*, 2018), for example, performing pairwise comparisons of risks for separated risk factors (see, e.g. Merrick *et al.*, 2005) (2). In the first case, first the adequate scales for risk factors should be specified. When specifying scales, linguistic categories can also be specified, and instead of scale values, membership functions can also be defined (see Example 2). In case of an invariant scale, there is no need to use a predefined scale; however, after pairwise comparison, the results should be categorized. In all cases, each effect is evaluated through $n + 1$ risk factors, where the total risk priority number

(TRPN) is calculated by these factors (see risk reduction and mitigation in Figure 2). Since one of the main benefits of the proposed TREF is flexibility, the evaluation of at least two factors (e.g. occurrence, severity) is required. However, more than three factors can be specified, and the number of evaluation risk factors is flexible.

4.1.3 Phase 3: calculation of risk values and thresholds through the process hierarchy/ process graph. Based on the proposed iterative bottom-up calculation method (see Definition 2), through the process hierarchy or an acyclic process graph, risk values can be calculated for each hierarchy level.

Contrary to traditional FMEA and fuzzy FMEA, TREF allows the specification of more than one effect to be assigned to a cause (see Figure 2). However, different failure modes and risk effects may have the same causes (common causes) (see Figure 3). The only restriction is to avoid cycles in the process hierarchy.

On the one hand, weights can be calculated by using ANP method, which can follow the process hierarchy. Applying weights gives a general view of the process risks, which are weighted by their importance. On the other hand, using weights is only optional. If there is no information about the importance of risk factors, the equal weights can be used. The other relevant example of unweighted aggregation uses the maximal value of the risk factors. The maximal value can also produce valuable information about risky processes (see S_2 in Example 4). This value presents the weak links (the worst/most risky processes).

In addition to calculating risk values or before performing the task, the thresholds must be specified for all levels (see risk assessment in Figure 2).

4.2 Monitoring risk values – operating the warning system
While the calculation of risk values and the thresholds should be calculated by the bottom-up iterative formula, the operation of the monitoring system can follow both the bottom-up and the top-down approaches.

4.2.1 Bottom-up approach. At the 0-th hierarchy level, risk factors are evaluated. A warning event has occurred if a risk factor is not lower than the threshold $W1$ or a criticality value is set to be 1 $W3$. For maintenance, this monitoring system shows which risk effect (in which domain) of process mode caused a failure mode and which factors are not lower than a threshold; therefore, a **specific corrective/preventive action** must be prescribed to mitigate the value of the risk factor. If a specific corrective/preventive action is not prescribed but the aggregated risk value is not lower than a threshold, **general corrective/ preventive actions** should be prescribed $W2$ to mitigate the aggregated risk values. General corrective/preventive actions should contain the set of specific tasks, which mitigates the values of risk factors. This bottom-up approach can be extended to the higher hierarchy levels, where general activities in a hierarchy level N should contain specific tasks to mitigate risk factors or risk values in the lower hierarchy.

4.2.2 Top-down approach. The top-down or *managerial approach* can be specified if in addition to the aggregating risk values the number of failure effects are calculated for all hierarchy levels. If there is a warning event on hierarchy level N, a general corrective/ preventive action is specified, which, similarly to the bottom-up, may (but in this case not necessarily) contain a (detailed) corrective/preventive action to mitigate risk factors. The number of failure effects in every level helps management to drill down and specify the set of corrective/preventive actions. While the bottom-up approach goes from the lower hierarchy level, specific corrective/preventive actions are specified to mitigate the risk factors, and general corrective/preventive actions are usually specified as a set of specific corrective/preventive actions. The top-down or managerial level starts at the top level of a hierarchy. Aggregated risk values give a general view of the risks; however, to reduce the number of failure effects, general corrective/preventive actions should be specified. Nevertheless, these general corrective/preventive actions may (but not necessarily) contain

specific corrective/preventive actions. For example, purchasing a new piece of equipment can be a general activity, which can solve several specific problems.

4.3 Schedule of corrective/preventive actions
After specifying the set of corrective/preventive actions:

(1) The forecasted effect of corrective/preventive actions should be specified (see, e.g. Bowles, 2003; Carmignani, 2009).

(2) Corrective/preventive actions should be organized as a maintenance project to minimize system shutdowns (see, e.g. Kosztyán, 2018).

The proposed TREF includes the schedule of corrective/preventive actions, which is a kind of flexible, discrete time/cost/quality trade-off problem; a future paper will focus on this scheduling problem. After completing risk mitigation projects, the improved risk effects will be re-evaluated (see the re-evaluation arrow in Figure 2), and if necessary, a new maintenance project will be organized.

5. Case study
5.1 Case study background
This section presents risk evaluation with different methods, including traditional FMEA, fuzzy FMEA and TREF, in the area of electric motor manufacturing. This case was chosen because of its substantive significance (Ragin, 1999). We have used a single-case design approach, where the case is selected because it is critical; that is, its conditions allow our method to be tested (Dubé and Paré, 2003; Yin, 2013). Some features of the company and processes must be clarified before the numerical illustration. This Hungarian subsidiary of a multinational corporation operates in the high-technology automotive industry. In the last decade, the market for high-precision drive systems has grown substantially. This company is a global leader in high-quality electric motors that are installed in critical applications such as surgical power tools, race cars and high-precision industrial applications. In so-called high-added-value manufacturing, the reliability of products plays a crucial role in their long lifespans. To improve the reliability of processes, a risk evaluation was conducted. The company has integrated quality management (ISO 9001), environmental management (ISO 14001) and health and safety management (ISO 45001) systems.

Company processes are divided into three categories (main processes): *support* (e.g. IT and process planning, finance, legal, facility and vehicle fleet, environmental protection); *management* (e.g. personnel development, risk management, management review, operative planning, process evaluation, audits, continuous improvement); and *customer order-related processes* (e.g. procurement, material management, production planning, production, quality assurance, warehousing). In this study, maintenance activities were selected as illustrative examples of the proposed model in Figure 2. They allow us to present the evaluation of each domain and all risk factors. Maintenance activities do not occur in separated functional units but are integrated with the core functions of the company.

Maintenance includes series of actions taken to maintain or restore the functionality of facilities/equipment. *Maintenance activities* occur in three processes: building engineering in facilities and the vehicle fleet (1.4.01P), means of production maintenance (1.6.01P), and maintenance of inspection tools in quality assurance (4.7.03P). In each case, potential failure modes, their causes and effects (on all three domains, i.e.: quality, environmental, health and safety) and the evaluation of risk factors were first identified by the risk evaluation team.

Figure 3 shows the logical connections among five failure modes, four identified causes and nine possible effects. The risk evaluation team, including the *system manager*, the *process*

manager and an *academic expert*, first identified five potential failure modes. The column marked "Processes"indicates the three maintenance processes: building engineering, means of production maintenance and inspection tool maintenance. The column marked "Causes"indicates the four causes: 045C for inadequate maintenance and 046C for insufficient technical requirements are common causes of two failure modes, and the remaining two causes are 018C for devices not registered and 012C for lack of knowledge. The column marked "Failure modes" indicates the type, that is, 1.4.01P.001M: *equipment failure* in building engineering; 1.6.01P.001M: *equipment failure* in means of production maintenance; 1.6.01P.002M: *nonplanned maintenance*; 4.7.03P.001M: *failure to maintain* inspection tools; 4.7.03P.002M: *improper maintenance requirements* for inspection tools. The "Effects by domains" column indicates the three domains based on the company's integrated management system: quality, environment and health and safety. The nine effects are 014E(*Q*): time loss, extra work time requirement; 005E(*H*): human injury; 020E(*Q*): missed calibration, incorrect measurement; 021E(*Q*): missed calibration, audit failure; 022E(*Q*): equipment failure; 050E(*E*): pollutants into the environment; 051E(*H*): discomfort; 052E(*Q*): production loss; and 053E(*H*): health impairment.

For example, failure mode *equipment failure* (1.4.01P.001M) is caused by *insufficient technical requirements* (046C) and *inadequate maintenance* (045C), and it affects *quality* (time loss (014E(*Q*)), *environment* (*pollutants released into the environment* (050E(E))) and *health and safety* (*discomfort* (051E(*H*)) and *health impairment* (053E(H))). As can be seen from the identifiers, causes and effects are not assigned to the processes or failure modes; there is a common database for the whole company. For example "operator failure," "mistyping" might occur in many processes, domains. This allows a smaller size data set with codes that are easier to memorize.

5.2 Applied methods
To check the applicability of TREF, it was necessary to compare it with the most frequently used risk evaluation methods, traditional FMEA and fuzzy FMEA (Liu *et al.*, 2013). The use of traditional FMEA with fuzzy FMEA at first sounds illogical because both are not used together. Fuzzy FMEA was developed to help those who were not experts in FMEA with linguistic terms. We develop our fuzzy FMEA method by working backward for this test as an example to test the usability of the TREF. We used sigmoid and bell/splay functions as membership functions (Johanyák and Kovács, 2004), and calculations were conducted via a weight method. Defuzzyfication relied on the multiplication of membership functions.

5.2.1 Implementation of the proposed TREF. For the TREF, we have used three additional risk factors in the case study, namely, control (*C*), information (*I*) and range (*R*), for a total of six factors. The first three are the same as those used in traditional FMEA and fuzzy FMEA: severity (*S*), occurrence (*O*) and detectability (*D*). This shows that the TREF is flexible and can include any number of risk factors ($n \geq 2$). The risk evaluation team agreed on the values of severity, occurrence, detection, control, information and range by using Tables AI–AIII.

Table I.
Result of the pairwise comparison for the domains (quality, environment, health and safety).
CR = 0.0598, critical value: = 0.1,
$I := \{Q, E, H\}$

Head	CI	RI	$W^{(1)}$
Objectives	0	0.58	1
Quality	0.0986	1.24	0.4545
Environment	0.1175	1.24	0.4545
Health and safety	0.1170	1.24	0.0909

The next step is to evaluate the importance of each risk factor in all domains to generate their weights. According to ANP, the reciprocal matrix determined through pairwise comparison for the three domains is shown in Table I.

Values in the table were generated according to Saaty (1987, 2004). The CI comes from the matrix of comparisons, RI is the random consistency index and w = weight. The CR is the consistency ratio, which can be calculated as follows: $CR = \sum w CI / \sum w RI$. Weights were calculated using geometric means. The consistency ratio (CR) was calculated by using the information in Table I. Based on the risk evaluation team's pairwise comparisons, the importance of the quality and environment domains are judged to be the same, while health and safety is considered less important. Table II shows the (0-th level) weights ($W_{i,j}^{(0)}$) of the six risk ($i = 1, \ldots, 6$) factors in three domains ($j = 1, 2, 3$).

In the case of the quality domain, detection has the greatest weight, while in the case of the environment and health and safety domains, severity has the greatest weight. Table II also shows that "Range" is the second-most important risk factor in the environment domain.

The effects are evaluated using the method proposed in Section 3. Each effect's TRPN value was obtained by calculating the $S_1 - S_4$ risk aggregating functions. Figure 4 shows the TRPN calculations and two kinds of warnings, that is, $W1$ and $W3$. For example, according to $S_1 - S_4$ risk aggregation functions, TRPN for the failure mode (1.4.01P.001M) 051E(H) effect can be calculated as follows:

$$\left(f_{\cdot,3}, W_{\cdot,3}^{(0)}, S_1 \right) : \text{TRPN}_{S_1}^{(1)} \left(f_{\cdot,3}, W_{\cdot,3}^{(0)} \right) = \prod_{i:=1}^{6} f_{i,3}^{W_{i,3}^{(0)}} = 2.25$$

$$\left(f_{\cdot,3}, 1/6, S_1 \right) : \text{TRPN}_{S_1}^{(1)} \left(f_{\cdot,3}, 1/6 \right) = \prod_{i:=1}^{6} f_{i,3}^{1/6} = \sqrt[6]{\prod_{i:=1}^{6} f_{i,3}} = 2.49$$

$$\left(f_{\cdot,3}, 1, S_2 \right) : \text{TRPN}_{S_2}^{(1)} (f_3, 1) = \max_i f_{i,3} = 5.00$$

$$\left(f_{\cdot,3}, W_{\cdot,3}^{(0)}, S_3 \right) : \text{TRPN}_{S_3}^{(1)} \left(f_3, W_{\cdot,3}^{(0)} \right) = \text{Median}(\{w_1 f_1, \ldots, w_n f_6\}) = 2.00$$

$$\left(f_{\cdot,3}, W_{\cdot,3}^{(0)}, S_4 \right) : \text{TRPN}_{S_4}^{(1)} \left(f_3, W_{\cdot,3}^{(0)} \right) = \sqrt{\sum_{i:=1}^{6} w_i f_i^2} = 3.14$$

Figure 4 shows the TRPN of each effect. The value of range is not lower than the critical value (threshold); therefore, corrective/preventive actions have to be specified to mitigate both (051E(H), 053E(H)) range effects (see $W1$ in Section 3.2). Figure 4 also shows that despite

Factors (f), Weights ($W^{(0)}$)	Quality	Environment	Health and safety
f_1 = Occurrence	0.1612	0.1364	0.2265
f_2 = Severity	0.2459	0.4462	0.4461
f_3 = Detection	0.4259	0.0435	0.0833
f_4 = Control	0.0943	0.0798	0.1325
f_6 = Information	0.0361	0.0400	0.0352
f_6 = Range	0.0366	0.2540	0.0765
CR	0.0796	0.0948	0.0943

Table II.
Results of the pairwise comparisons of the risk factors. Critical value: = 0.1

	Occurrance	Severity	Detection	Control	Information	Range	Criticality	TRPN					Failure effect
								Geometrical	Un-weighted Geom.	Weighted Median	Radial	Max Value	
Maximum:	2	2	3	4	3	5	1	2.47	2.99				
Geom. avrg.	2	2	2.5	4	3	5	1	2.43	2.9				
Average:	2	2	2.67	4	3	5	0.8	2.43	2.91				
Warnings:	0	0	0	0	0	2	1	0	0				
	2	2	2.66666667	4	3	5	0.8	2.433	2.907				
weight:	0.2265	0.4461	0.0833	0.1325	0.0352	0.0765	1	(f,w,S1)	(f,1,S1)	(f,w,S3)	(f,w,S4)	(f,1,S2)	
Critical value	3	2	5	4	4	4	0	3	3	3	3	3	
Effect 051 Unpleasant comfort, colds	2	2	1	4	3	5	0	2.25	2.49	2	3.14	5	Yes
053 Health damage	2	2	3	4	3	5	1	2.47	2.99	2	3.27	5	Yes

Health & Safety

W1: Risk factor value is greater than a specified threshold

W2: Aggregated risk value is greater than a specified threshold

W3: Risk factor value are decided to be critical

Figure 4.
The evaluation of TRPN for failure mode (1.4.01P.001M) effects (051E(*H*) and 053E(*H*))

average TRPNs (TRPN$_{051E,H}$ and TRPN$_{053E,H}$) that are lower than the specified threshold, 053E(H) is critical (see W3 in Section 3.2), and the risk evaluation team specified corrective/preventive actions to avoid this risk effect.

To use the proposed TREF as a module in an expert system, different levels of aggregation should be performed. According to risk aggregation function (S_1), the weighted geometric mean of total risk priority numbers was calculated for process levels, failure modes, common causes and common effects. Since the effect (discomfort 051E(H)) was judged to be four times less important than health damage (O53E(H)) by the risk evaluation team, we weighted the geometric mean value (the value input into the oval in Figure 4), which is used to calculate the TRPN$^{(2)} = 2.426$. Failure mode 1.4.01P.001M has two other effects, 014E(Q) ($j = 1$) and 050E(E) ($j = 2$), which were evaluated from the quality (Q) and environmental (E) points of view (see Table BI in Appendix B). These values are TRPN$_1^{(1)} = 2.66$, TRPN$_2^{(1)} = 2.48$ (see Table BI) and TRPN$_3^{(1)} = 2.36$ (see Figure 4). This value (the average TRPN for the quality/environment/health and safety effects of failure mode 014P.001M) represents a general view of failure modes. The weighted average TRPN for failure mode 1.4.01P.001M is:

$$\text{TRPN}_1^{(2)} = \left(\text{TRPN}_1^{(1)}\right)^{W_1^{(1)}} \cdot \left(\text{TRPN}_2^{(1)}\right)_2^{W_2^{(1)}} \cdot \left(\text{TRPN}_3^{(1)}\right)^{W_3^{(1)}}$$
$$= 2.66^{0.4545} \cdot 2.48^{0.4545} \cdot 2.36^{0.0909} = 2.55 \tag{2}$$

These values are lower than a critical value (threshold); however, to detect the number of failure effects, we had to calculate both the maximum values of TRPNs and the number of failure effects (see the results in Figure 4 and Table BI). It is important to note the proposed multi-level approach detected more (in this case, three) failure effects, which would not have been possible when calculating RPNs for only one aspect. Moreover, Figure 4 and Table BI show that the traditional RPN, which is based only on the occurrence (O), severity (S) and detection (D) factors, cannot detect the critical range (R) within these effects (014E(Q), 051E(H) and 053E(H)).

Since there was no information about the importances of the processes, unweighted versions of $S_1 - S_4$ formulas are used. For example, TRPN$_{S_1,1.4.01P}^{(3)} = 2.55$, TRPN$_{S_1,1.6.01P}^{(3)} = 2.78$, TRPN$_{S_1,4.7.03P}^{(3)} = 2.44$, processes (e.g. TRPN$_{S_1,1.4P}^{(4)} = 2.64$) and process areas (e.g. TRPN$_{S_1,1P}^{(5)} = 2.56$). However, we can use another method of aggregation: We calculated the TRPNs of all maintenance processes by using unweighted S_1 formula (geometric mean) TRPN$_{S_1,\text{MAINTENANCE}}^{(3)} = \sqrt[3]{2.55 \cdot 2.78 \cdot 2.44} = 2.59$, common causes (e.g. TRPN$_{O45C}^{(3)} = \sqrt[2]{2.55 \cdot 2.78} = 2.66$) and common effects (e.g. TRPN$_{S_1,005E(H)}^{(3)} = 2.67$).

In addition to the general view, the maximum values of TRPNs and risk factors were calculated for failure modes, processes, process areas and main processes. We obtained 6 W1,

Method	Risk aggregation protocol	Warning system	W1	W2	W3	Failure effects
FMEA	$(\{O, S, D\}, , 1, \prod)$	W2	n/a	0	n/a	0
Fuzzy FMEA	$\{\tau(O), \tau(S), \tau(D), 1, \prod\}$	W2	n/a	0	n/a	0
TREF	$(\{O,S,D,C,I,R\}, 1, S_1)$	W1–W3	6	3	1	5
TREF	$(\{O,S,D,C,I,R\}, W_Q^{(0)}, S_1)$	W1–W3	6	0	1	5
TREF	$(\{O,S,D,C,I,R\}, W_Q^{(0)}, S_3)$	W1–W3	6	2	1	7
TREF	$(\{O,S,D,C,I,R\}, W_Q^{(0)}, S_4)$	W1–W3	6	3	1	5

Table III.
Comparison of risk
evaluation methods

8 *W*2, 1 *W*3 warnings; thus, we had to specify at least $6 + 8 + 1 = 15$ corrective/preventive actions.

This case study shows that the TREF is a flexible risk evaluation framework. First, the same source of hazards caused risks in multiple management areas, such as automotive customer, special environmental concerns and data handling of risky processes, and each effect was evaluated by various criteria for the three domains. In addition, TREF can address an arbitrary number of risk factors; we used $6 + 1$ risk factors, namely, severity (S), occurrence (O), detection (D), control (C), information (I) and range (R), with criticality as $+1$. Finally, different risk factors had different weights in the case of the three domains; for example, "range" was the second-most important risk factor in the environment domain.

5.2.2 Comparison of applied risk evaluation methods. FMEA and fuzzy FMEA are special cases of the proposed TREF, where only three (i.e. (O,S,D)) factors are considered (see Table III). In the case of FMEA, the values are represented as determined by the risk evaluation team, and for fuzzy FMEA, the input values were modified for 5–5 linguistic terms (described by the membership function: τ), which covers the factor range.

Both traditional FMEA and Fuzzy FMEA can be extended considering the importance of the weights of factors. In Section 2, we noted that these weights can be obtained by AHP/ANP or any pairwise comparison methods. All of these methods predict a failure mode if the aggregated risk value is not lower than a threshold *W*2.

As presented in Section 2, the most frequently used aggregation function is the production ($\prod$); however, other relevant aggregation (see, e.g. S_1, S_3, S_4) functions can be used.

The proposed TREF does not require weights (see Row 3 in Table III), but thresholds for all factors *W*1 and direct settings *W*3 can also be applied.

Figure 5 shows all quality-related risk effects on the quality domain of failure modes in process 4.7.03P (see Figure 3). Figure 5 shows the setting thresholds for factors highlighting four risk factors (occurrence, control, information and range) related to four failure effects (015, 016, 020 and 021). TREF does not require the complete filling of risk factors: it allows an arbitrary number of risk factors (see 017 and 018 failure effects).

Table III shows the detected warnings and predicted failure modes of the hierarchical warning system.

FMEA and fuzzy-FMEA methods only focus on aggregated values; therefore, *W*1 and *W*3 cannot be interpreted by these traditional methods. Nevertheless, if thresholds are already specified for the first three factors, 1 *W*1 warning event can be detected, and therefore, 1 failure effect (016) can also be predicted by the traditional FMEA and fuzzy-FMEA methods. If the three extra factors are considered, five additional *W*1 warning events can be detected. Therefore, three more failure effects (015,020,021) can be predicted (see Figure 5). In addition, based on expert decision, the criticality of effect 019 is set to 1. Therefore, one more warning *W*3 is detected, and one more failure effect is specified.

The four applied aggregation functions have signaled eight warnings *W*2. Two warnings and two additional failure effects (017,018) were detected by weighted median (S_3), and three warnings and three failure effects (016, 020 and 021) were denoted by the weighted radial distance (S_4) and unweighted geometric mean (S_1).

The six detected *W*1 warnings indicated specific corrective/preventive actions to mitigate the value of risk factors, while *W*2 and *W*3 warnings required general activities.

5.3 Advantages of the proposed framework

This empirical case study demonstrates the following advantages. The main benefit of the proposed TREF is its flexibility, which enables different company demands to be considered.

First, the conventional RPN method used by traditional and fuzzy FMEA includes three risk factors, that is, severity (S), occurrence (O) and detection (D). Using TREF, each effect can

		Occurrance	Severity	Detection	Control	Information	Range	Criticality	Geometrical	Un-weighted Geom.	Weighted Median	Radial distance	Failure effect
Maximum:	7		5	3	5	5	5	1	2,81	3,45			
Geom. avg.	2,6		2,71	2,02	3,48	2,61	3,25	1	2,45	2,74			
Average:	2,9		2,9	2,1	3,6	2,8	3,3	0,1	2,25	2,51			
Warnings:	1		0	0	2	2	1	1	0	3	2	3	
Domain	2,64		2,64	1,91	3,27	2,55	3,00	0,09	2,25	2,51			
quality									TRPN (6)				
									(f,w,S1)	(f,1,S1)	(f,w,S3)	(f,w,S4)	
weight:		0,161	0,246	0,426	0,094	0,036	0,037	1,000					
Effect	Critical (threshold) value	5	5	8	4	4	4	0	3	3	3	3	3
015	Final inspection of the product is not possible	4	2	1	5	2	5	0	1,88	2,71	1,00	2,72	Yes
016	Inadequate inspection	7	2	2	5	3	4	0	2,78	3,45	2,00	3,72	Yes
017	Conform product is measured nonconform	2	2	3	3	2	3	0	2,51	2,45	3,00	2,60	Yes
018	Nonconform product is measured conform	2	3	3	3	2	3	0	2,77	2,62	3,90	2,83	Yes
017	Conform product is measured nonconform	2	2					0	2,00	2,00	0,00	1,28	No
018	Nonconform product is measured conform	2		2	3	2		0	1,71	2,21	2,00	1,83	No
019	Selection	4	2	2	2	2	3	1	2,27	2,40	2,00	2,47	Yes
020	Missed calibration, wrong measurement	2	5	2	4	5	3	0	2,81	3,26	2,00	3,35	Yes
021	Missed calibration, audit failure	2	5	2	4	5	3	0	2,81	3,26	2,00	3,35	Yes
022	Lack of knowledge	2	3	2	4	3	3	0	2,43	2,75	2,00	2,59	No

W1: Risk factor value is greater than a specified threshold

W2: Aggregated risk value is greater than a specified threshold

W3: Risk factor value are decided to be critical

Figure 5.
The risk
evaluation table

be evaluated through an arbitrary ($n \geq 2$) number of risk factors, where the TRPN is calculated by these factors. We used $6 + 1$ risk factors in the case study, namely, severity (S), occurrence (O), detection (D), control (C), information (I) and range (R), with criticality as $+1$. Flexibility is useful, as the traditional RPN, which is based only on occurrence (O), severity (S) and detection (D) factors, cannot detect the critical range (R) within the effects (014E(Q), 051E(H) and 053E(*H*)) (see Figure 4 and Table BI). Additional flexibility in the TREF occurs because it includes an extra risk factor (criticality) to allow the risk evaluation team to specify corrective/preventive actions even if the TRPN is lower than the specified threshold. As Figure 4 shows, despite average TRPNs ($\text{TRPN}_{051E,H}^{(2)}$ and $\text{TRPN}_{053E,H}^{(2)}$) that are lower than the specified threshold, 053E(*H*) is critical. This TREF property allows the risk evaluation team to specify corrective/preventive actions for that case.

Second, failure effects were evaluated by three domains, that is, quality, environment and health and safety. Each effect was evaluated by the different criteria of the three domains. The case study also demonstrated that the three domains had different weights and the different risk factors had different weights in the three domains. We expected that the traditional three FMEA factors (S, O and D) were the most important; however, the range was the second-most important risk factor in the environment domain.

Third, TREF allows for different levels of aggregation. Average TRPNs were calculated for process levels, failure modes, common causes and common effects. By aggregating failure effects, an average TRPN was calculated for the failure modes (e.g. $\text{TRPN}_{1.4.01P}^{(4)}$), process (e.g. $\text{TRPN}_{1.4P}^{(5)}$) and process area levels (e.g. $\text{TRPN}_{1P}^{(6)}$). In addition, the proposed TREF allows for another method of aggregation (i.e. horizontal). In that case, the TRPNs of all maintenance processes ($\text{TRPN}_{\text{MAINTENANCE}}^{(3)}$), common causes (e.g. $\text{TRPN}_{O45C}^{(3)}$) and common effects (e.g. $\text{TRPN}_{005E(H)}^{(3)}$) were calculated. This property is useful as it allows for flexible (i.e. hierarchical and horizontal) aggregation by TREF users. In addition to these advantages, calculating the maximum values of TRPNs and maximum values of risk factors provides an evaluation of the failure effects on the higher level of hierarchy.

Finally, the WS has signaled a number of warning events and failure effects. Considering the three additional risk factors such as range, information, control, five additional $W1$ warning events and three additional failure effects (015,020,021) were denoted. Based on expert decision, additional warning $W3$ and failure effects (019) were specified by the criticality factor. In addition, eight warnings $W2$ were signaled by the aggregation functions. Depending on the aggregation function, two additional failure effects (017,018) were predicted by the weighted median (S_3), and three failure effects (016, 020 and 021) were denoted by the weighted radial distance (S_4) and unweighted geometric mean (S_1).

6. Conclusion
In this paper, we developed a flexible multi-level risk evaluation framework that is a crucial step in the customization of risk evaluation. The innovative nature of the proposed TREF is its flexible characteristic that helps to overcome the limitations of current risk evaluation methods, which are static in nature. The proposed method overcomes this weakness using flexible RAP and WS and an arbitrary number of risk factors. TREF allows a novel multi-level, such as a hierarchical and horizontal aggregation. In the case of hierarchical aggregation, TRPNs can be calculated at different levels, such as factor, domain, effect, mode, process and organization. In the case of horizontal aggregation, the TRPN for all processes, common causes and common effects can be calculated. Flexible use of RAP only works together with a flexible WS. Warning events and failure effects can be signaled from each

part of the TREF WS, such as $W1$, $W2$ and $W3$. Specific needs can be addressed locally by specifying different thresholds and warning rules for different risk factors. In addition, we conclude that traditional risk factors such as S, O and D are not sufficient for a comprehensive risk evaluation. To overcome the limitation of current methods, TREF makes it possible to evaluate each effect by means of an arbitrary number of $(n \geq 2)$ risk factors, including criticality. This makes TREF unique compared to other risk evaluation methods. TREF flexibly sets the number and importance of risk factors, which can be different by domain. The proposed framework was examined by means of a case study of an electric motor manufacturing company.

At the same time, TREF has a number of important managerial implications. By adapting TREF, managers can use different and adequate factors, weights, aggregation functions and warning rules at different levels that are adapted to the actual needs of the company. In this way, the proposed framework helps users to more comprehensively identify critical and risky elements in a more tailor-made manner than existing risk evaluation methods. Thus, TREF supports plant managers in designing a more appropriate risk management policy for the company, which might result in lower losses or increased profits. However, in the course of practical application, managers need to bear in mind the actual context of the company.

Future work in the following three directions is considered. First, the proposed framework can only address the acyclic process graph. However, in real operations, elements may occur that are reciprocally related to each other. Regardless, it would be interesting for future studies to incorporate cyclic structures into the framework. The second limitation of the present study is that the RAP works together with the WS. Future studies should try to investigate when is it worthwhile to diverge the structure of RAP from WS and consequently determine the cases for which such divergence is appropriate and advantageous. Third, in addition to the electric motor manufacturing company, in the future, more case studies should be conducted to further verify and reinforce the reliability of the proposed framework.

References

Abad, J., Dalmau, I. and Vilajosana, J. (2014), "Taxonomic proposal for integration levels of management systems based on empirical evidence and derived corporate benefits", *Journal of Cleaner Production*, Vol. 78, pp. 164-173.

Asif, M., Searcy, C., Zutshi, A. and Fisscher, O.A. (2013), "An integrated management systems approach to corporate social responsibility", *Journal of Cleaner Production*, Vol. 56, pp. 7-17, Sustainability management beyond corporate boundaries.

Aven, T. (2016), "Risk assessment and risk management: review of recent advances on their foundation", *European Journal of Operational Research*, Vol. 253 No. 1, pp. 1-13.

Aven, T. and Zio, E. (2014), "Foundational issues in risk assessment and risk management", *Risk Analysis: an Official Publication of the Society for Risk Analysis*, Vol. 34 No. 7, pp. 1164-1172.

Azadeh-Fard, N., Schuh, A., Rashedi, E. and Camelio, J.A. (2015), "Risk assessment of occupational injuries using accident severity grade", *Safety Science*, Vol. 76, pp. 160-167.

Beliakov, G., Pradera, A. and Calvo, T. (2008), Aggregation Functions: A Guide for Practitioners, in *Studies in Fuzziness and Soft Computing*, Vol. 221, Springer, Berlin, Heidelberg.

Bernardo, M. (2014), "Integration of management systems as an innovation: a proposal for a new model", *Journal of Cleaner Production*, Vol. 82, pp. 132-142.

Bowles, J.B. (2003), "An assessment of rpn prioritization in a failure modes effects and criticality analysis", in *Annual Reliability and Maintainability Symposium*, IEEE, Tampa, FL, pp. 380-386.

Bowles, J.B. and Peláez, C. (1995), "Fuzzy logic prioritization of failures in a system failure mode, effects and criticality analysis", *Reliability Engineering and System Safety*, Vol. 50 No. 2, pp. 203-213.

Braglia, M. and Bevilacqua, M. (2000), "Fuzzy modelling and analytical hierarchy processing as a means of quantifying risk levels associated with failure modes in production systems", *Technology, Law and Insurance*, Vol. 5 No. 3-4, pp. 125-134.

Calvo, T., Kolesárová, A., Komorníková, M. and Mesiar, R. (2002), "Aggregation operators: properties, classes and construction methods", in *Aggregation Operators*, Physica-Verlag HD, pp. 3-104.

Carmignani, G. (2009), "An integrated structural framework to cost-based FMECA: the priority-cost FMECA", *Reliability Engineering and System Safety*, Vol. 94 No. 4, pp. 861-871.

Certa, A., Enea, M., Galante, G.M. and Fata, C.M.L. (2017), "Electre tri-based approach to the failure modes classification on the basis of risk parameters: an alternative to the risk priority number", *Computers and Industrial Engineering*, Vol. 108, pp. 100-110.

Chang, C., Wei, C. and Lee, Y. (1999), "Failure mode and effects analysis using fuzzy method and grey theory", *Kybernetes*, Vol. 28 No. 9, pp. 1072-1080.

Chang, K.H. and Wen, T.C. (2010), "A novel efficient approach for DFMEA combining 2-tuple and the OWA operator", *Expert Systems with Applications*, Vol. 37 No. 3, pp. 2362-2370.

Chen, Y., Song, G., Yang, F., Zhang, S., Zhang, Y. and Liu, Z. (2012), "Risk assessment and hierarchical risk management of enterprises in chemical industrial parks based on catastrophe theory", *International Journal of Environmental Research and Public Health*, Vol. 9 No. 12, pp. 4386-4402.

de Oliveira, O.J. (2013), "Guidelines for the integration of certifiable management systems in industrial companies", *Journal of Cleaner Production*, Vol. 57, pp. 124-133.

dos Santos, M.A.R., Salomon, V.A.P. and Marins, F.A.S. (2015), "Analytic network process and balanced scorecard applied to the performance evaluation of public health systems", *Pesquisa Operacional*, Vol. 35 No. 2, pp. 353-361.

Dubé, L. and Paré, G. (2003), "Rigor in information systems positivist case research: current practices, trends, and recommendations", *MIS Quarterly*, Vol. 27 No. 4, pp. 597-636.

Etherton, J.R. and Myers, M.L. (1990), "Machine safety research at niosh and the future directions", *International Journal of Industrial Ergonomics* No. 6, pp. 163-174.

Fattahi, R. and Khalilzadeh, M. (2018), "Risk evaluation using a novel hybrid method based on fmea, extended multimoora, and ahp methods under fuzzy environment", *Safety Science*, Vol. 102, pp. 290-300.

Gauthier, F., Chinniah, Y., Burlet-Vienney, D., Aucourt, B. and Larouche, S. (2018), "Risk assessment in safety of machinery: impact of construction flaws in risk estimation parameters", *Safety Science*, Vol. 109, pp. 421-433.

Hansson, S.O. and Aven, T. (2014), "Is risk analysis scientific?", *Risk Analysis: an Official Publication of the Society for Risk Analysis*, Vol. 34 No. 7, pp. 1173-1183.

Ilangkumaran, M., Karthikeyan, M., Ramachandran, T., Boopathiraja, M. and Kirubakaran, B. (2015), "Risk analysis and warning rate of hot environment for foundry industry using hybrid MCDM technique", *Safety Science*, Vol. 72, pp. 133-143.

ISO 12100 (2010), *ISO 12100:2010 Safety of Machinery General Principles for Design Risk Assessment and Risk Reduction*, ISO Copyright Office, Geneva.

ISO 14001 (2015), *ISO 14001:2015 Environmental Management Systems - Requirements with Guidance for Use*, ISO Copyright Office, Geneva.

ISO 45001 (2017), *ISO 45001:2017 Dis - Occupational Health and Safety*, ISO Copyright Office, Geneva.

ISO 9001 (2015), *ISO 9001:2015 Quality Management Systems - Requirements*, ISO Copyright Office, Geneva.

Johanyák, Z. and Kovács, S. (2004), "On the right selection of the fuzzy membership function", *GAMF Journal*, Vol. XIX, pp. 73-84.

Kalantarnia, M., Khan, F. and Hawboldt, K. (2009), "Dynamic risk assessment using failure assessment and bayesian theory", *Journal of Loss Prevention in the Process Industries*, Vol. 22 No. 5, pp. 600-606.

Kanes, R., Marengo, M.C.R., Abdel-Moati, H., Cranefield, J. and Véchot, L. (2017), "Developing a framework for dynamic risk assessment using bayesian networks and reliability data", *Journal of Loss Prevention in the Process Industries*, Vol. 50, pp. 142-153.

Karasan, A., Ilbahar, E., Cebi, S. and Kahraman, C. (2018), "A new risk assessment approach: safety and critical effect analysis (SCEA) and its extension with pythagorean fuzzy sets", *Safety Science*, Vol. 108, pp. 173-187.

Khan, F., Rathnayaka, S. and Ahmed, S. (2015), "Methods and models in process safety and risk management: past, present and future", *Process Safety and Environmental Protection*, Vol. 98, Complete, pp. 116-147.

Kinney, G. and Wiruth, A. (1976), *Practical Risk Analysis for Safety Management*, Naval Weapons Center, China Lake, CA.

Klinke, A. and Renn, O. (2002), "A new approach to risk evaluation and management: risk-based, precaution-based, and discourse-based strategies1", *Risk Analysis: an Official Publication of the Society for Risk Analysis*, Vol. 22 No. 6, pp. 1071-1094.

Kokangl, A., Polat, U. and Dağsuyu, C. (2017), "A new approximation for risk assessment using the ahp and fine kinney methodologies", *Safety Science*, Vol. 91, pp. 24-32.

Kosztyán, Z.T. (2018), "Serviceability of large-scale systems", *Simulation Modelling Practice and Theory*, Vol. 84, pp. 222-231.

Kutlu, A.C. and Ekmekçioğlu, M. (2012), "Fuzzy failure modes and effects analysis by using fuzzy TOPSIS-based fuzzy AHP", *Expert Systems with Applications*, Vol. 39 No. 1, pp. 61-67.

Liu, H.C., Liu, L., Bian, Q.H., Lin, Q.L., Dong, N. and Xu, P.C. (2011), "Failure mode and effects analysis using fuzzy evidential reasoning approach and grey theory", *Expert Systems with Applications*, Vol. 38 No. 4, pp. 4403-4415.

Liu, H.C., Liu, L. and Liu, N. (2013), "Risk evaluation approaches in failure mode and effects analysis: a literature review", *Expert Systems with Applications*, Vol. 40 No. 2, pp. 828-838.

Liu, H.C., Liu, L., Liu, N. and Mao, L.X. (2012), "Risk evaluation in failure mode and effects analysis with extended VIKOR method under fuzzy environment", *Expert Systems with Applications*, Vol. 39 No. 17, pp. 12926-12934.

Liu, H.C., You, J.X., Ding, X.F. and Su, Q. (2015), "Improving risk evaluation in fmea with a hybrid multiple criteria decision making method", *International Journal of Quality and Reliability Management*, Vol. 32 No. 7, pp. 763-782.

Liu, H.C., You, J.X., Fan, X.J. and Lin, Q.L. (2014), "Failure mode and effects analysis using d numbers and grey relational projection method", *Expert Systems with Applications*, Vol. 41 No. 10, pp. 4670-4679.

Liu, H.T. and Tsai, Y.L. (2012), "A fuzzy risk assessment approach for occupational hazards in the construction industry", *Safety Science*, Vol. 50 No. 4, pp. 1067-1078.

Lolli, F., Ishizaka, A., Gamberini, R., Rimini, B. and Messori, M. (2015), "FlowSort-GDSS – a novel group multi-criteria decision support system for sorting problems with application to FMEA", *Expert Systems with Applications*, Vol. 42 No. 17-18, pp. 6342-6349.

MacKenzie, C.A. (2014), "Summarizing risk using risk measures and risk indices", *Risk Analysis: an Official Publication of the Society for Risk Analysis*, Vol. 34 No. 12, pp. 2143-2162.

Maheswaran, K. and Loganathan, T. (2013), "A novel approach for prioritization of failure modes in fmea using mcdm", *IJERA*, Vol. 3, pp. 733-739.

Malekitabar, H., Ardeshir, A., Sebt, M.H., Stous, R. and Teo, E.A.L. (2018), "On the calculus of risk in construction projects: contradictory theories and a rationalized approach", *Safety Science*, Vol. 101, pp. 72-85.

Manuele, F.A. (2005), "Risk assesment & hierarchies of control", *Safety Management* No. 5, pp. 33-39.

Merrick, J.R.W., van Dorp, J.R. and Singh, A. (2005), "Analysis of correlated expert judgments from extended pairwise comparisons", *Decision Analysis*, Vol. 2 No. 1, pp. 17-29.

Ni, H., Chen, A. and Chen, N. (2010), "Some extensions on risk matrix approach", *Safety Science*, Vol. 48 No. 10, pp. 1269-1278.

Øien, K., Utne, I. and Herrera, I. (2011), "Building safety indicators: Part 1 – theoretical foundation", *Safety Science*, Vol. 49 No. 2, pp. 148-161.

O'Keeffe, V.J., Tuckey, M.R. and Naweed, A. (2015), "Whose safety? flexible risk assessment boundaries balance nurse safety with patient care", *Safety Science*, Vol. 76, pp. 111-120.

Ouédraogo, A., Groso, A. and Meyer, T. (2011), "Risk analysis in research environment – part II: weighting lab criticity index using the analytic hierarchy process", *Safety Science*, Vol. 49 No. 6, pp. 785-793.

Panchal, D., Chatterjee, P., Yazdani, M. and Chakraborty, S. (2019a), "A hybrid mcdm approach-based framework for operational sustainability of process industry", in *Advanced Multi-Criteria Decision Making for Addressing Complex Sustainability Issues*, IGI Global, Hershey, PA, pp. 1-13.

Panchal, D., Singh, A.K., Chatterjee, P., Zavadskas, E.K. and Keshavarz-Ghorabaee, M. (2019b), "A new fuzzy methodology-based structured framework for ram and risk analysis", *Applied Soft Computing*, Vol. 74, pp. 242-254.

Panchal, D., Tyagi, M. and Sachdeva, A. (2019c), "A novel framework for evaluation of failure risk in thermal power industry", in Sachdeva, A., Kumar, P. and Yadav, O.P. (Eds) *Operations Management and Systems Engineering*, Springer Singapore, Singapore, pp. 37-50.

Panchal, D., Jamwal, U., Srivastava, P., Kamboj, K. and Sharma, R. (2018a), "Fuzzy methodology application for failure analysis of transmission system", *International Journal of Mathematics in Operational Research*, Vol. 12 No. 2, pp. 220-237.

Panchal, D., Mangla, S.K., Tyagi, M. and Ram, M. (2018b), "Risk analysis for clean and sustainable production in a urea fertilizer industry", *International Journal of Quality and Reliability Management*, Vol. 35 No. 7, pp. 1459-1476.

Panchal, D. and Kumar, D. (2016), "Stochastic behaviour analysis of power generating unit in thermal power plant using fuzzy methodology", *OPSEARCH*, Vol. 53 No. 1, pp. 16-40.

Panchal, D. and Kumar, D. (2017), "Risk analysis of compressor house unit in thermal power plant using integrated fuzzy fmea and gra approach", *international Journal of Industrial and Systems Engineering*, Vol. 25 No. 2, pp. 228-250.

Panchal, D. and Srivastava, P. (2019), "Qualitative analysis of cng dispensing system using fuzzy fmea–gra integrated approach", *International Journal of System Assurance Engineering and Management*, Vol. 10 No. 1, pp. 44-56.

Pasman, H., Reniers, G., Pasman, H. and Reniers, G. (2014), "Past, present and future of quantitative risk assessment (qra) and the incentive it obtained from land-use planning (lup)", *Journal of Loss Prevention in the Process Industries*, Vol. 28, Complete, pp. 2-9.

Petrović, D.V., Tanasijević, M., Milić, V., Lilić, N., Stojadinović, S. and Svrkota, I. (2014), "Risk assessment model of mining equipment failure based on fuzzy logic", *Expert Systems with Applications*, Vol. 41 No. 18, pp. 8157-8164.

Ragin, C.C. (1999), "The distinctiveness of case-oriented research", *Health Services Research*, Vol. 34 No. 5 Pt 2, pp. 1137-1151.

Rebelo, M.F., Santos, G. and Silva, R. (2016), "Integration of management systems: towards a sustained success and development of organizations", *Journal of Cleaner Production*, Vol. 127, pp. 96-111.

Reiman, T., Rollenhagen, C., Pietikinen, E. and Heikkil, J. (2015), "Principles of adaptive management in complex safety–critical organizations", *Safety Science*, Vol. 71, pp. 80-92.

Riahi, R., Bonsall, S., Jenkinson, I. and Wang, J. (2012), "A seafarer's reliability assessment incorporating subjective judgements", *Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment*, Vol. 226 No. 4, pp. 313-334.

Saaty, R. (1987), "The analytic hierarchy processwhat it is and how it is used", *Mathematical Modelling*, Vol. 9 No. 3, pp. 161-176.

Saaty, T.L. (2004), "Decision making the analytic hierarchy and network processes (ahp/anp)", *Journal of Systems Science and Systems Engineering*, Vol. 13 No. 1, pp. 1-35.

Shaker, F., Shahin, A. and Jahanyan, S. (2019), "Developing a two-phase qfd for improving fmea: an integrative approach", *International Journal of Quality and Reliability Management*.

Silva, M.M., de Gusmão, A.P.H., Poleto, T., e Silva, L.C. and Costa, A.P.C.S. (2014), "A multidimensional approach to information security risk management using FMEA and fuzzy theory", *International Journal of Information Management*, Vol. 34 No. 6, pp. 733-740.

SRA (2015), *Foundations of Risk Analysis, Discussion Document*, www.sra.com.resources (accessed 15 august 2017).

Stamatis, D.H. (2003), *Failure Mode and Effect Analysis: FMEA from Theory to Execution*, ASQ Quality Press, Milwaukee, WI.

Tay, K.M. and Lim, C.P. (2006), "Fuzzy fmea with a guided rules reduction system for prioritization of failures", *International Journal of Quality and Reliability Management*, Vol. 23 No. 8, pp. 1047-1066.

Torabi, S., Soufi, H.R. and Sahebjamnia, N. (2014), "A new framework for business impact analysis in business continuity management (with a case study)", *Safety Science*, Vol. 68, pp. 309-323.

Wang, W., Liu, X., Qin, Y. and Fu, Y. (2018), "A risk evaluation and prioritization method for fmea with prospect theory and choquet integral", *Safety Science*, Vol. 110, pp. 152-163.

Wang, Y.M., Chin, K.S., Poon, G.K.K. and Yang, J.B. (2009), "Risk evaluation in failure mode and effects analysis using fuzzy weighted geometric mean", *Expert Systems with Applications*, Vol. 36 No. 2, pp. 1195-1207.

Xu, K., Tang, L., Xie, M., Ho, S. and Zhu, M. (2002), "Fuzzy assessment of fmea for engine systems", *Reliability Engineering and System Safety*, Vol. 75 No. 1, pp. 17-29.

Yin, R.K. (2013), *Case Study Research: Design and Methods*, Sage publications, Los Angeles, CA.

Yousefi, S., Alizadeh, A., Hayati, J. and Baghery, M. (2018), "HSE risk prioritization using robust DEA-FMEA approach with undesirable outputs: a study of automotive parts industry in Iran", *Safety Science*, Vol. 102, pp. 144-158.

Zammori, F. and Gabbrielli, R. (2012), "ANP/RPN: a multi criteria evaluation of the risk priority number", *Quality and Reliability Engineering International*, Vol. 28 No. 1, pp. 85-104.

Zheng, G., Zhu, N., Tian, Z., Chen, Y. and Sun, B. (2012), "Application of a trapezoidal fuzzy AHP method for work safety evaluation and early warning rating of hot and humid environments", *Safety Science*, Vol. 50 No. 2, pp. 228-239.

(The Appendix follows overleaf)

Corresponding author
Zsolt Tibor Kosztyán and can be contacted at: kzst@gtk.uni-pannon.hu

Appendix A. Evaluation scales

Value	Quality Occurrence	Severity	Detection	Control	Information	Range
1	It is almost impossible	There is no implication or minor financial costs	Always, 100%	It does not require any regulation, spontaneously regularized	No information is required	Nobody is involved
2	Annually	A slight error, or no greater than 2000 CHF-cost associated	Very easy to detect, 98–100%	Open loop (nonfeedback) control is enough	All the information is available to know the failure modes and effects	Max. 1 person is involved in the factory, max. 1 day is the correction
3	Quarterly	Serious error, or 2vd5 000 CHF costs	Easy to detect, 95–98%	Only local control, easily controllable using feedback	The failure modes and effects are known, but not all the information about effects given in advance, but can be acquired after the occurrence	2 or 3 people are involved in the factory, or 4–5 days is the correction
4	Once a month	Customer annoyance (informal notice) or minor damage to component or system, or 5–10 000 CHF costs	Easier than medium difficulty to detect, 90–95%	Requires local forecast, feed-forward control	The failure modes and effects are known, but not all the information about the effects is given in advance and also cannot be obtained after the occurrence	4 or 5 people are involved in the company, or 4–7 days is the correction
5	Several times a month	Customer complaints or 10–100 000 CHF costs-	Slightly harder than medium difficulty to detect, 80–90%	Higher-level control, easily controllable using feedback	The failure modes are known, but not all the effects, or not all information about the effects is given in advance, but can be obtained after the occurrence	Min. 6 people are involved in the company, 7–10 days is the correction

(*continued*)

Table AI.
Risk factors in quality aspect

Value	Quality Occurrence	Severity	Detection	Control	Information	Range
6	Once a week	High degree of customer dissatisfaction or 100–250 000 CHF costs	Medium difficulty to detect, 70–80%	Requires forecast from higher level, feed-forward control	The failure modes are known, but not all the effects, or not all information about the effects is given in advance, and can be only estimated after the occurrence	Impact on the total internal operation of the company but can be corrected internally
7	Several times a week	High degree of customer dissatisfaction or 250–500 000 CHF costs-	Harder than medium difficulty to detect, 50–70%	Requires forecast from higher level, feed-forward control, but only partially effective	Failure modes are nor known fully and effects can be estimated only partly	Widespread involvement within the company, one business partner is involved
8	Once a day	Customer loss, or 500–750 000 CHF costs	Difficult to detect, 25–50%	It requires the use of multiple control tools for multiple levels, but only partially effective	There is no information about failure modes and effects in advance, but can be obtained after occurrence	Widespread involvement within the company, more business partners are involved
9	Several times a day	Violation of law, or 150–1,000,000 CHF costs-	Very difficult to detect, less than 25%	There is no influence, not controllable	There is no information about failure modes and effects in advance, and they can be obtained after occurrence only in part.	Widespread involvement among the business partners, social environment and inside the company
10	Per shift	Human injury or the cost of more than CHF 1 million	Almost impossible to detect, close to 0%	There is no influence, going to unfavorable direction	There is no and cannot be obtained information about failure modes and effects	Widespread involvement among the business partners, social environment and inside the company, large publicity

Table AI.

Appendix B

Value	Environment Occurrence	Severity	Detection	Control	Information	Range
1	It is almost impossible	No effect: the effect is unverifiable, change is not sensible	Sense perception is directly and immediately possible by anyone, at any time of the day	It does not require any regulation, spontaneously regularized	No information is required	Nobody is involved
2	Annually	Neutral: the effect is verifiable, but the caused change is not sensible	Sense perception is directly and immediately possible by anyone, but only in smooth atmospheric/optical/noise conditions (e.g. only in daytime)	Open-loop (nonfeedback) control is enough	All the information is available to know the failure modes and effects	Max. 1 person is involved in the factory, max. 1 day is the correction
3	Quarterly	Tolerable: unwanted changes can be detected, but they do not affect any essential features of the given unit	Especially difficult to detect sensory, even in daytime, under optimal conditions	Only local control, easily controllable using feedback	The failure modes and effects are known, but not all the information about effects given in advance, but can be acquired after the occurrence	2 or 3 people are involved in the factory, or 4–5 days is the correction
4	Once a month	Acceptable as tolerated status: unwanted changes can be detected, they do affect more essential features of the given unit	Sense perception is possible only using tool (e.g. light, the use of contrast agents, the use of indicators, etc.)	requires local forecast, feed-forward control	The failure modes and effects are known, but not all the information about the effects is given in advance and also cannot be obtained after the occurrence	4 or 5 people are involved in the company, or 4–7 days is the correction
5	Several times a month	Annoying: functions are sustainable, but conditions are deteriorating	Sense perception is not possible; however, the change is demonstrable mathematically (calculation, inventory, mass balance, etc.)	Higher-level control, easily controllable using feedback	The failure modes are known, but not all the effects, or not all information about the effects is given in advance, but can be obtained after the occurrence	Min. 6 people are involved in the company, 7-10 days is the correction

Table AII.
Risk factors in
environment aspect

Value	Environment Occurrence	Severity	Detection	Control	Information	Range
6	Once a week	Little problem: the irreversibility described above exists; however, the change is not over limit or qualification barrier, or the change is over limit, but the effect is reversible without any corrective action	Immediately can be detected by instruments in own ownership and use	Requires forecast from higher level, feed-forward control	The failure modes are known, but not all the effects, or not all information about the effects is given in advance, and can be only estimated after the occurrence	Impact on the total internal operation of the company but can be corrected internally
7	Several times a week	Problem: irreversibility exists, the change is over one limit or qualification barrier, the effect is reversible without any corrective action	Immediately can be detected by instruments from external source	Requires forecast from higher level, feed-forward control, but only partially effective	Failure modes are nor known folly and effects can be estimated only partly	Widespread involvement within the company, one business partner is involved
8	Once a day	Restrictive: the change is over each limit, regulation and so on, as a consequence the unit gets into lower quality class, but the change is reversible	Not all parameters can be detected using in-situ procedure	It requires the use of multiple control tools for multiple levels, but only partially effective	There is no information about failure modes and effects in advance, but can be obtained after occurrence	Widespread involvement within the company, more business partners are involved, environmental damage in max. 50 m radius
9	Several times a day	Harmful: the change is over each limit, regulation and so on, as a consequence the unit gets into lower quality class, the change is irreversible	Can be detected only based on sampling, in laboratory conditions	There is no influence, not controllable	There is no information about failure modes and effects in advance, and they can be obtained after occurrence only in part	Widespread involvement among the business partners, social environment and inside the company, environmental damage in 50200 m radius
10	Per shift	Terminating: the unit or the whole system ceases to exist or loses determinating features	Instrumental analytical tests cannot detect (or close to error limit), at most generation changes can be detected	There is no influence, going to unfavorable direction	There is no and cannot be obtained information about failure modes and effects	Widespread involvement among the business partners, social environment and inside the company, large publicity, environmental damage is in over 200 m radius

Table AII.

Value	Health and safety Occurrence	Severity	Detection	Control	Information	Range
1	It is almost impossible	There is no harm	Always, 100%	It does not require any regulation, spontaneously regularized	No information is required	Nobody is involved
2	Annually	Not serious work accident, the duration of incapacity to work is not more than 3 days	Very easy to detect, 98–100%	Open-loop (nonfeedback) control is enough	All the information is available to know the failure modes and effects	Max. 1 person is involved in the factory, max. 1 day is the correction
3	Quarterly	Not serious work accident, the period of incapacity more than 3 days	Easy to detect, 95–98%	Only local control, easily controllable using feedback	The failure modes and effects are known, but not all the information about effects given in advance, but can be acquired after the occurrence	2 or 3 people are involved in the factory, or 4–5 days is the correction
4	Once a month	Work accident causing no serious truncation	Easier than medium difficulty to detect, 90–95%	Requires local forecast, feed-forward control	The failure modes and effects are known, but not all the information about the effects is given in advance and also cannot be obtained after the occurrence	4 or 5 people are involved in the company, or 4–7 days is the correction
5	Several times a month	Serious work accident with truncation	Slightly harder than medium difficulty to detect, 80–90%	Higher-level control, easily controllable using feedback	The failure modes are known, but not all the effects, or not all information about the effects is given in advance, but can be obtained after the occurrence	Min. 6 people are involved in the company, 7–10 days is the correction

Table AIII.
Risk factors in health and safety aspect

(*continued*)

Value	Health and safety Occurrence	Severity	Detection	Control	Information	Range
6	Once a week	Permanent damage, causing the loss independent living ability	Medium difficulty to detect, 70–80%	Requires forecast from higher level, feed-forward control	The failure modes are known, but not all the effects, or not all information about the effects is given in advance, and can be only estimated after the occurrence	Impact on the total internal operation of the company but can be corrected internally
7	Several times a week	The loss or significant damage of sensory organ (or sense), reproductive capacity	Harder than medium difficulty to detect, 50–70%	Requires forecast from higher level, feed-forward control, but only partially effective	Failure modes are nor known fully and effects can be estimated only partly	Widespread involvement within the company, one business partner is involved
8	Once a day	Work accident causing the loss of speaking skill, striking distortion, paralysis, mental disorder	Difficult to detect, 25–50%	It requires the use of multiple control tools for multiple levels, but only partially effective	There is no information about failure modes and effects in advance, but can be obtained after occurrence	Widespread involvement within the company, more business partners are involved
9	Several times a day	Life-threatening injury, damage based on medical opinion	Very difficult to detect, less than 25%	There is no influence, not controllable	There is no information about failure modes and effects in advance, and they can be obtained after occurrence only in part.	Widespread involvement among the business partners, social environment and inside the company
10	Per shift	Fatal work accidents (injury, fetus, newborn)	Almost impossible to detect, close to 0%	There is no influence, going to unfavorable direction	There is no and cannot be obtained information about failure modes and effects.	Widespread involvement among the business partners, social environment and inside the company, large publicity

Quality

	Occurrence	Severity	Detection	Control	Information	Range	Criticality	Geometrical (f,w,S1)	Un-weighted Geom. (f,1,S1)	Weighted Median (f,w,S3)	Radial (f,w,S4)	Max Value (f,1,S2)	Failure effect
Maximum:	3	3	2	4	3	6	0	2.66	5				
Geom. avg:	3	3	2	4	3	6	0	2.66	3.3				
Average:	3	3	2	4	3	6	0	2.66	3.3				
Warnings:	0	0	0	0	0	1	0	0	1				
Domain	3	3	2	4	3	6	0	2.66	3.3				
	Occurrance	Severity	Detection	Control	Information	Range	Criticality	Geometrical	Un-weighted Geom.	Weighted Median	Radial	Max Value	
weight:	0.161	0.246	0.426	0.094	0.036	0.037	1.000	(f,w,S1)	(f,1,S1)	(f,w,S3)	(f,w,S4)	(f,1,S2)	
Critical value	5	5	8	4	4	4	0	4	4	4	4	4	
Effect 014 Time loss, extra work time requirement	3	3	2	4	3	6	0	2.66	3.30	2.00	2.80	6.00	Yes

Environment

	Occurrence	Severity	Detection	Control	Information	Range	Criticality	Geometrical (f,w,S1)	Un-weighted Geom. (f,1,S1)	Weighted Median (f,w,S3)	Radial (f,w,S4)	Max Value (f,1,S2)	Failure effect
Maximum:	2	3	2	3	2	2	0	2.48	2.29				
Geom. avg:	2	3	2	3	2	2	0	2.48	2.29				
Average:	2	3	2	3	2	2	0	2.48	2.29				
Warnings:	0	0	0	0	0	0	0	0	0				
Domain	2	3	2	3	2	2	0	2.48	2.29				
	Occurrance	Severity	Detection	Control	Information	Range	Criticality	Geometrical	Un-weighted Geom.	Weighted Median	Radial	Max Value	
weight:	0.136	0.446	0.044	0.080	0.040	0.254	1.000	(f,w,S1)	(f,1,S1)	(f,w,S3)	(f,w,S4)	(f,1,S2)	
Critical value	5	5	5	4	4	4	0	4	4	4	4	4	
Effect 050 Pollutants into the environment	2	3	2	3	2	2	0	2.48	2.29	3.00	2.24	3.00	No

Table BI.
Calculation of TRPNs for effects 014E(Q) and 050E(E)