Proceedings of the 35th European Safety and Reliability & the 33rd Society for Risk Analysis Europe Conference
Edited by Eirik Bjorheim Abrahamsen, Terje Aven, Frederic Bouder, Roger Flage, Marja Ylönen
 Published by Research Publishing, Singapore.
doi: 10.3850/978-981-94-3281-3_ESREL-SRA-E2025-P2373-cd

Development Of A Hybrid Metric For Physical Security Assessment

Thomas Termin

Institute for Security Systems, University of Wuppertal, Germany. E-mail: thomas.termin@gmx.de.

Daniel Lichte

Institute for the Protection of Terrestrial Infrastructures, German Aerospace Center, Germany. E-mail: daniel.lichte@dlr.de.

Kai-Dietrich Wolf

Institute for Security Systems, University of Wuppertal, Germany. E-mail: wolf@iss.uni-wuppertal.de.

In industrial practice, physical security assessments are increasingly performed using scoring methods. However, since scoring methods involve uncertainty, users face challenges in evaluating investment alternatives. While quantitative metrics have the advantage over scorings in that precise calculations can be made, their application is not as simple and intuitive as simple scorings. From a user's perspective, the key question is how to merge the strengths of both metrics to facilitate risk assessment without compromising accuracy. The goal of this paper is to formulate requirements for a hybrid metric for assessing physical vulnerability and to demonstrate the applicability of the outlined solution approach using a specific use case. In a first step, the problems of scoring are explained using the Harnser metric as an example. In a second step, the quantitative metric used to measure the quality of scoring is defined. The third step explains how the scoring under consideration can be extended and modified to replicate the quantitatively calculated results for all calculation results. In a final step, the proposed adjustment approach is demonstrated. Finally, the results are summarized and starting points for further research are identified.

Keywords: Physical Security, Uncertainty, Metrics, Risk Adjustment, Decision-Making, Hybrid Metric.

1. Introduction

In recent years, physical security has become a major societal and political concern, particularly in response to increasing reports of sabotage and climate activism targeting critical infrastructure (Rucht, 2023). Protecting essential systems from physical attack is critical to maintaining the stability of key processes and services that underpin modern society (Stober, 2024). One of the fundamental tools for designing effective security measures is risk assessment, which evaluates threats and vulnerabilities based on attacker profiles.

A widely used approach in industrial security assessments is scoring, as it provides a straightforward way to categorize risks and support decision making (Krisper, 2021). However, scoring methods introduce a significant degree of uncertainty, making it difficult to justify security investments or compare different protection strategies. Alternatively, quantitative risk assessment provides objective, mathematically precise risk assessments (Termin et al., 2021). While more accurate, these methods often rely on complex calculations that are not as intuitive or easily applied as scoring-based approaches.

From a practical perspective, the key question is how to combine the strengths of both scoring and quantitative approaches into a hybrid metric that ensures both usability and analytical robustness. This paper builds on previous work by Termin et al. (2023) and introduces a refined approach that improves scoring-based vulnerability assessments to match the accuracy of purely quantitative methods. The proposed methodology is demonstrated through a concrete use case that illustrates how the hybrid metric can be applied to assess the effectiveness of security measures for a defined protection barrier.

By bridging the gap between intuitive scoring and precise quantitative evaluation, this approach enables a more reliable assessment of security investments while maintaining accessibility for practitioners.

2. Background

Scoring metrics are commonly used in physical security assessments to evaluate vulnerability levels and provide a simplified yet structured approach. However, to improve accuracy and comparability, a scoring-based metric needs to be adapted to a quantitative metric. In this study, we use the Harnser metric (Harnser, 2010) as a scoring-based approach and compare it to the Intervention Capability Metric (ICM) (Lichte et al., 2016), which is a quantitative method.

The Harnser metric assigns scores to three key variables: Protection (P), Observation (O), and Intervention (I), each of which is scored on a scale of 1 to 5 (with 5 being the highest and 1 being the lowest). The sum of these three scores results in a vulnerability score (V) ranging from 3 to 15 (see Fig. 1).

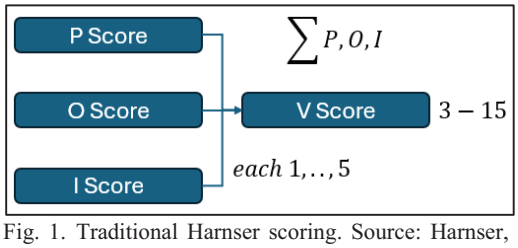

Fig. 1. Traditional Harnser scoring. Source: Harnser, (2010).

In contrast, the ICM applies probabilistic density functions to P, O, and I as inputs and derives vulnerability as a discrete value based on the time-dependent interaction of these variables (see Fig. 2).

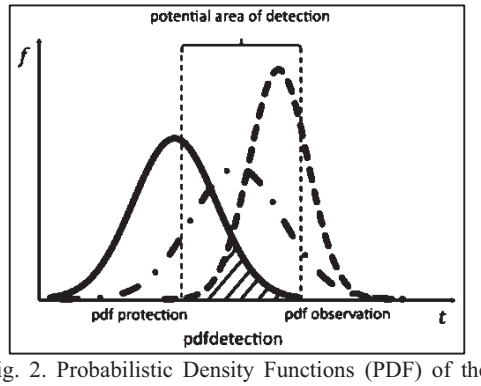

Fig. 2. Probabilistic Density Functions (PDF) of the Intervention Capability Metric (ICM). Source: Lichte et al. (2016).

To enable a meaningful comparison between the scoring and quantitative methods, Termin et al. (2023) introduced a mapping approach that assigns time-based values to each P, O, and I score, making them compatible with the quantitative model. Since the exact probability distribution of scoring-based outcomes is not known in advance, the initial scoring range (3-15) is extended using assumed probability intervals to approximate the expected quantitative outcomes.

A systematic comparison is then performed by evaluating all 125 possible score combinations ($5 \times 5 \times 5$ permutations) in both the scoring and quantitative metrics. The scoring method produces probability intervals, while the quantitative method produces discrete values. By adjusting the assumed probability intervals of the scoring model, the results are aligned so that the quantitative results fall within or close to the corresponding scoring intervals. However, a limitation arises because the scoring method only generates values between 3 and 15, meaning that several different permutations can result in the same sum score (e.g., P = 1, O = 2, I = 1 and P = 2, O = 1, I = 1 both result in a sum score of 4).

If the scoring system had 125 unique score levels, each calculated quantitative value could be directly mapped to a unique scoring result, making the scoring metric as precise as the quantitative method. However, because the Harnser metric groups multiple permutations into the same sum score, it cannot fully replicate the granularity of the quantitative results, i.e. the full result space determined by the number of all possible permutations (combinations). This raises an important methodological question:

How can a scoring metric be designed to generate uniquely distinguishable scores that correspond one-to-one to all possible permutations of the scoring variables?

The key challenge is to determine a generalized formula or rule that defines the relationship between the number of scoring variables, their possible values, and their method of combination (e.g., addition, multiplication, or weighting). Ideally, the scoring system should be structured to produce unambiguous and meaningful results that capture the full range of possible vulnerability levels while maintaining ease of use in security assessments.

3. Approach

The fictitious infrastructure analyzed in this paper consists of one threat (T), one barrier (B) with attributes P, O, and I, and one asset (A) (see Fig. 3). The attacker must overcome the barrier to reach the asset.

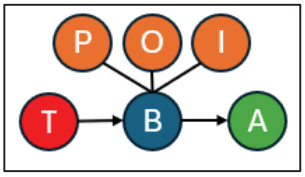

Fig. 3. Fictitious infrastructure under investigation.

To develop a scoring metric that produces uniquely distinguishable scores, ensuring that each permutation of the scoring variables corresponds to a distinct result and accurately replicates quantitative calculations, a general mathematical rule must be defined. Key factors that influence this rule include:

(1) **Number of scoring variables:** This determines the dimensionality of the scoring system and represents the key security factors being evaluated (e.g., protection (P), observation (O), and intervention (I)).

(2) **Range of each scoring variable:** Each variable must have a defined range of values (e.g., P, O, and I range from 1 to 5), which affects the total number of possible score permutations.

(3) **Combination Method:** The scoring variables must be mathematically linked in such a way that each combination produces a unique and distinguishable result. This can be achieved through additive, multiplicative, or weighted functions, depending on the desired level of granularity and the need for a one-to-one mapping to the quantitative model.

3.1. General Formula for the Scoring Metric

Let n represent the number of scoring variables (e.g., n = 3 for P, O and I). Each scoring variable X_i $(where\ i \in \{1, 2, \dots, \mathrm{n}\}$ takes values from a set $X_i = \{1, 2, \dots, m_i\}$, where m_i represents the number of possible values for each variable, e. g. $m_1 = m_2 = m_3 = 5$ for P, O and I. The total number of unique permutations N_{perm} of the scoring variables is given by Eq. (1):

$$N_{perm} = m_1 \times m_2 \times \dots \times m_n = \prod_{i=1}^{n} m_i \quad (1)$$

For the example where each variable has five values (1-5), the number of possible score permutations is $N_{perm} = (5 \times 5 \times 5 =)\ 125$. To ensure that each unique combination of X_i values leads to a distinct score, a weighted sum function is introduced as follows (see Eq. (2)):

$$VS = \sum_{i=1}^{n} w_i \cdot X_i \quad (2)$$

where:

- w_i are the weighting factors assigned to each scoring variable X_i.
- X_i is the assigned value of the scoring variable.

The key requirement is that the weights w_i must be chosen so that each possible combination of scoring variables $X_1, X_2, \dots, X_n$ results in a unique score.

3.2. Ensuring Unique Scores Using an Exponential Weighting Scheme

A practical wat to achieve the uniqueness is by assigning exponentially decreasing weights, ensuring that no two different combinations produce the same score. For n = 3 scoring variables, each with $(m_1 = m_2 = m_3 =)$ five possible values, the weightings can be defined as follows (see Eq. (3)):

$$w_1 = m_2 \times m_3,\ w_2 = m_3,\ w_3 = 1 \quad (3)$$

For **P, O, I with values from 1 to 5**, this results in the following equation (compare Eq. (4)):

$$VS = w_1 \cdot P + w_2 \cdot O + w_3 \cdot I = 25 \cdot P + 5 \cdot O + 1 \cdot I \quad (4)$$

These weights ensure that each combination of P, O, and I generate a unique vulnerability

score by giving a higher weight to variables positioned earlier in the hierarchy. Importantly, this weighting does not imply relative importance of the variables—it purely serves to differentiate outcomes numerically, e.g., the total number of possible scores is equal to the number of permutations, in this case 125.

3.3. Example Calculations for Unique Scoring

Using this scheme presented, each unique combination of P, O, and I result in a distinct vulnerability score (VS). Example calculations (see Eq. (5)):

For P = 3, O = 4, I = 2: (5)
$$VS = 25 \cdot 3 + 5 \cdot 4 + 2 = 97$$

For P = 5, O = 5, I = 5:
$$VS = 25 \cdot 5 + 5 \cdot 5 + 5 = 155$$

By applying this approach, the scoring system now generates 125 uniquely distinguishable vulnerability scores instead of the previous range of "3" to "15". The range of possible scores can be seen in Eq. (6):

Lowest score (P = 1, O = 1, I = 1) (6)
Combination 1:

P = 1, O = 1, I = 1:
$$VS = 25 \cdot 1 + 5 \cdot 1 + 1 = 31$$

Highest score (P = 5, O = 5, I = 5)
Combination 125:

P = 5, O = 5, I = 5:
$$VS = 25 \cdot 5 + 5 \cdot 5 + 5 = 155$$

This transformation ensures a one-to-one mapping between the scoring system and the number of unique permutations.

3.4. Establishing a General Rule for Unique Scoring Metrics

The general rule for constructing a Harnser-based scoring metric that produces distinct vulnerability scores is VS (Eq. (2)), where the weights w_i are assigned exponentially decreasing values, ensuring that the number of unique scores matches the number of possible variable permutations. For n = 3 and 5 values per variable, the rule in Eq. (4) ensures each of the 125 possible permutations maps to a unique score.

3.5. Practical Implications for Physical Security Assessment

This revised scoring system allows for direct alignment with quantitative security metrics by ensuring that

(1) Each quantitatively calculated vulnerability score (derived as a discrete probability) corresponds to a unique score in the revised scoring system.
(2) Each permutation of the scoring variables (P, O, I) produces a unique, identifiable vulnerability score.
(3) The scoring method can now be directly mapped to quantitative scores, allowing for better comparability and risk prioritization in security assessments.

By structuring the scoring metric in this way, the approach bridges the gap between simple, intuitive scoring models and precise, quantitative assessments, addressing a key limitation of traditional security scoring frameworks.

4. Demonstration

To illustrate the proposed approach for adapting the Harnser scoring metric, we consider a reference security architecture consisting of a barrier protecting a critical asset. The barrier is characterized by three key security properties: Protection (P), Observation (O), and Intervention (I). An attacker must successfully bypass the barrier to reach the asset, while a defender aims to detect and stop the attack before the asset is compromised.

4.1. Attack and Defence Dynamics

The system is considered vulnerable if the attacker reaches the asset before the defender can intervene. The analysis of an attack scenario is complete when

(1) The attacker successfully reaches the asset (breaches the security barrier), or
(2) The defender intervenes in time to neutralize the threat.

To model this interaction, we compare Harnser's scoring-based evaluation to the Intervention Capability Metric (ICM), a quantitative evaluation method. The ICM uses probability distributions to model the interaction between attack and defense times. Each Harnser score (1 to 5) is mapped to a corresponding time step, which is then used as input to the ICM.

4.2. Time-based probabilistic mapping in the ICM

In the ICM, the time variables associated with protection (P), observation (O), and intervention (I) are treated as probabilistic density functions (PDFs), each characterized by a mean and a standard deviation. The behavior of these functions follows the expected security principles:

(1) Protection (P): A higher P value corresponds to increased resistance, i.e., a longer time required for an attacker to breach the barrier.
(2) Observation (O): A higher O-score corresponds to increased detection capability, i.e. the time required for observation decreases.
(3) Intervention (I): A higher I-score represents a faster defensive response, i.e., the intervention time decreases.

For demonstration purposes, ICM Configuration 1 (ICM 1) is defined by assigning specific time levels to each Harnser score to ensure a structured comparison (see Table 1).

Table 1. Mapping of Harnser scores to time stages in the ICM (Termin et al., 2023).

P	ICM 1	O	ICM 1	I	ICM 1
1	$\mu = 15$ $\sigma = 30$	1	$\mu = 135$ $\sigma = 30$	1	$\mu = 135$ $\sigma = 30$
2	$\mu = 45$ $\sigma = 30$	2	$\mu = 105$ $\sigma = 30$	2	$\mu = 105$ $\sigma = 30$
3	$\mu = 75$ $\sigma = 30$	3	$\mu = 75$ $\sigma = 30$	3	$\mu = 75$ $\sigma = 30$
4	$\mu = 105$ $\sigma = 30$	4	$\mu = 45$ $\sigma = 30$	4	$\mu = 45$ $\sigma = 30$
5	$\mu = 135$ $\sigma = 30$	5	$\mu = 15$ $\sigma = 30$	5	$\mu = 15$ $\sigma = 30$

In the traditional Harnser metric, the scores of the three assessment variables (Protection (P), Observation (O), and Intervention (I)) are simply summed to determine the vulnerability score. In the proposed approach, however, the weighted formula from equation (4) is applied instead, ensuring that each combination of P, O, and I result in a unique vulnerability score.

As described by Termin et al. (2023), the 125 possible vulnerability scores are initially distributed with equal probability (100% total), since the exact match between the scoring-based results and the quantitative ICM values is not known a priori. A ranking approach is used to determine an initial probability distribution:

(1) The lowest vulnerability score is assigned the highest probability, reflecting a more secure system.
(2) Conversely, the highest vulnerability score is assigned the lowest probability, indicating a more vulnerable system.

This initial distribution serves as a baseline for comparison before adjusting the scoring intervals to better match the quantitative results of the ICM (see Table 2).

Table 2. Mapping of Harnser scores to probability intervals.

Harnser Score	Lower Interval Limit (LIL)	Upper Interval Limit (UIL)	Mean Value
31	0.992	1	0.996
32	0.984	0.992	0.988
33	0.976	0.984	0.98
34	0.968	0.976	0.972
35	0.96	0.968	0.964
…	…	…	…
153	0.016	0.024	0.02
154	0.008	0.016	0.012
155	0	0.008	0.004

As shown in Table 3, the next step is to compute all 125 possible permutations using both the scoring-based metric and the quantitative ICM for direct comparison.

Table 3. Calculation of vulnerability using both metrics.

P	O	I	Harnser Vuln. Score	LIL	UIL	Mean Value	ICM 1 Vuln. Value
1	1	1	31	0.992	1	0.996	0.9999999
1	1	2	32	0.984	0.992	0.988	0.9999999
1	1	3	33	0.976	0.984	0.98	0.9999997
…	…	…	…	…	…	…	…

Significant discrepancies between the results of the two metrics are evident in Figure 4, as illustrated by the plotted calculation results.

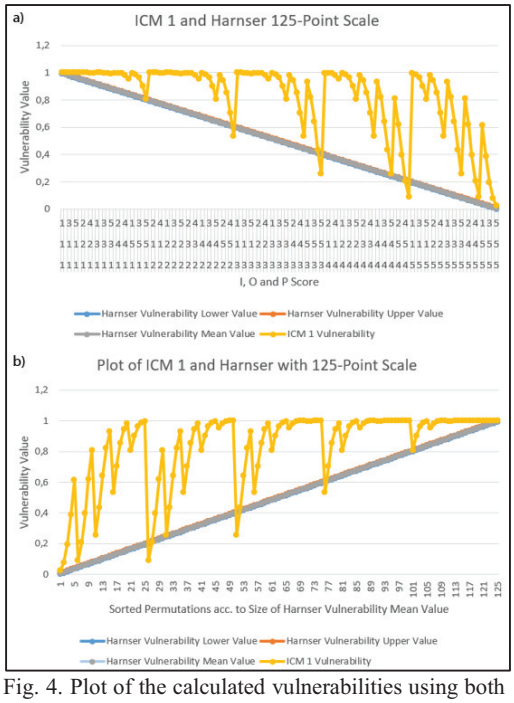

Fig. 4. Plot of the calculated vulnerabilities using both metrics. a) Sorting by permutation, b) Sorting by the size of the Harnser vulnerability mean values. Yellow: ICM 1 vulnerability values. Else: Harnser vulnerability values.

Since each vulnerability score is uniquely distinguishable and corresponds to the total number of calculated permutations, each score can now be mapped directly to the discrete vulnerability probability value determined by ICM 1 for the corresponding permutation. This ensures a one-to-one correspondence between Harnser vulnerability scores and ICM 1 vulnerability values (see Table 4).

Table 4. Adjustment of the probabilities behind the Harnser Vulnerability Scores.

P	O	I	Sco -re	LIL	UIL	Mean Value	ICM 1 Vuln. Value	Harnser Vuln. Adjusted
5	5	5	155	0	.008	.004	.0239	.0239
5	5	4	154	.008	.016	.012	.0766	.0766
5	5	3	153	.016	.024	.02	.1951	.1951
…	…	…	…	…	…	…	…	…

As a result of the adjustments outlined in Table 4, Figure 1b is modified accordingly, resulting in the updated representation shown in Figure 5.

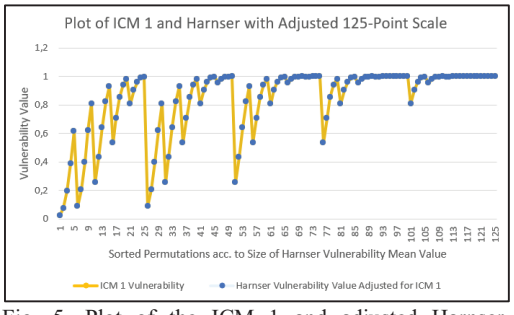

Fig. 5. Plot of the ICM 1 and adjusted Harnser Vulnerability values. The Harnser vulnerability values (blue dots) are unique match one-to-one to the ICM 1 vulnerability values (yellow dots).

As shown in Figure 2, the scoring scale was successfully adjusted to match the quantitatively calculated vulnerability values for the respective permutations. Within the given time step assumptions, the proposed scoring method demonstrates an accuracy comparable to that of the quantitative approach.

Furthermore, Figure 6, which presents the results sorted by ICM 1 vulnerability values, clearly confirms that each quantitatively derived value can be effectively mapped using the adapted scoring system. This underscores the robustness of the proposed approach in bridging the gap between scoring-based and quantitative assessments.

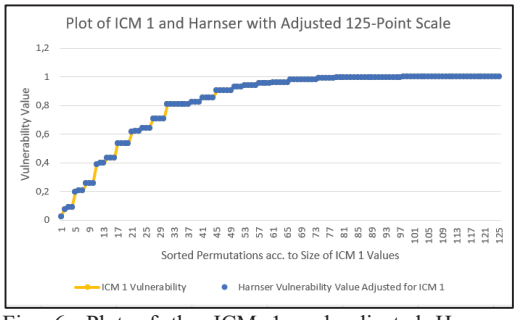

Fig. 6. Plot of the ICM 1 and adjusted Harnser Vulnerability values sorted by the size of the ICM 1 values. Blue dots: Adjusted Harnser vulnerability values. Yellow dots: ICM 1 vulnerability values.

5. Conclusion

This paper has demonstrated a systematic approach to extending the Harnser scoring metric to produce quantifiable, directly comparable results to a quantitative vulnerability assessment. By extending and modifying the traditional Harnser method, we have successfully ensured that each of the (5 × 5 × 5 =) 125 possible score combinations corresponds to a unique, quantitatively verifiable vulnerability score.

5.1. Key findings and practical implications

The core principle behind this adaptation is that scoring-based security assessments can be aligned with quantitative calculations if

(1) The scoring scale is expanded to accommodate the number of possible permutations of scores and time levels.
(2) Time-dependent levels are assigned within the scoring metric, ensuring a case-specific mapping between scoring results and real-world vulnerability estimates based on a quantitative metric.

To achieve this, the traditional Harnser vulnerability scale (ranging from 3 to 15) was extended to 125 unique categories, exactly matching the number of possible score permutations (5 × 5 × 5). This was accomplished by introducing three weighting factors, each applied to the protection (P), observation (O), and intervention (I) scores. The weighting factors were scaled exponentially to ensure that each final score remained uniquely distinguishable and could be mapped to a quantitative vulnerability score derived from the ICM. Unlike the previous approach in Termin et al. (2023), which relied on interval-based probabilities, the adapted method allows for precise, discrete vulnerability values for each score (see Fig. (7)).

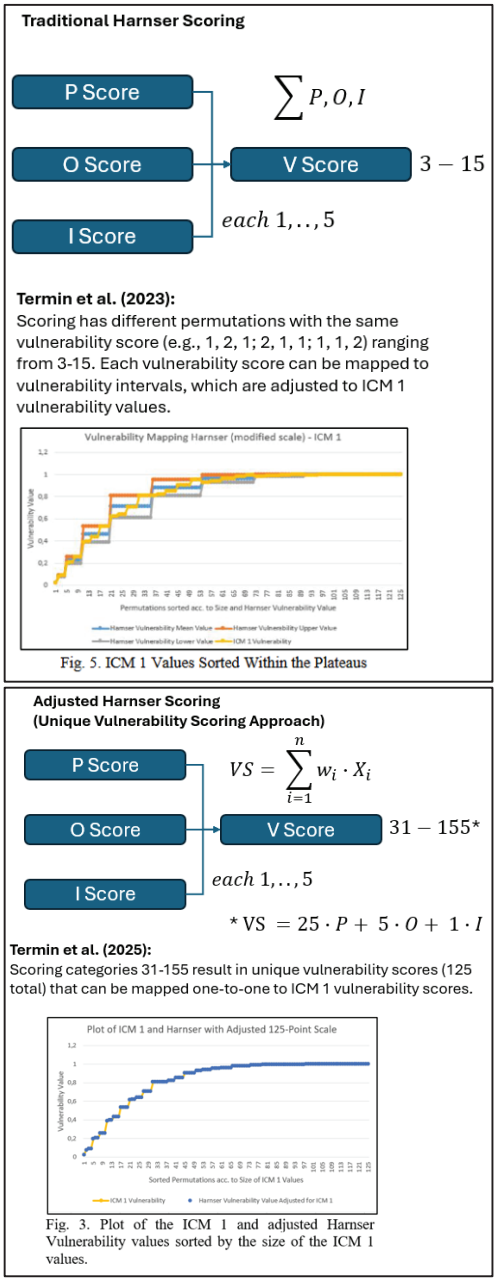

Fig. 7. Comparison of the approach of Termin et al. (2023) with the approach of Termin et al. (2025) (this document).

5.2. Benefits and Practical Applications

While this customization does not address fundamental limitations of scoring systems, such as the ordinal nature of scores, it does provide a viable solution for aligning semi-quantitative ratings with quantitative security assessments. This approach allows practitioners to use scoring while maintaining quantitative accuracy, making it highly relevant to practical security assessments. Key benefits include

(1) Improved decision making: Security professionals can now compare and validate scoring results with quantitative models, increasing confidence in investment decisions.
(2) Minimal added complexity: The introduction of weighting factors slightly increases computational complexity but remains intuitive and easy to use.
(3) Use case adaptability: The method is flexible and can be tailored to different time step assumptions based on the specific characteristics of a given security scenario. For example, different barrier types or attack scenarios can have customized probability distributions, allowing for more context-sensitive risk assessments.
(4) Potential for standardization: The adapted Harnser scale, including its probability mappings, could serve as a reference tool for vulnerability assessments in different security environments, ensuring a standardized risk assessment process in industrial applications.

5.3. Future research and industrial implementation

The feasibility of this approach has been demonstrated using a simple barrier-asset model, but further research is needed to validate its applicability to more complex infrastructure systems. Future studies should focus on:

(1) Cost-benefit analysis: Applying this method to large-scale infrastructure projects to assess its impact on security investment decisions.
(2) Real-world industrial implementation: Assessing how the adapted scoring model can be integrated into existing security assessment frameworks used in industry, government, and critical infrastructure protection.
(3) Alternative configurations: Investigating how different time step assumptions affect scoring accuracy and whether dynamic time adjustments could further improve risk assessments.

This research serves as a foundation for future work that bridges the gap between scoring metrics and rigorous quantitative analysis. By improving the practical applicability of scoring-based assessments, this approach provides a scalable and adaptable tool for security professionals, policy makers, and risk analysts alike.

References

Harnser Group (2010). A Reference Security Management Plan for Energy Infrastructure. European Commission.

Krisper, M. (2021). Problems with risk matrices using ordinal scales. arXiv preprint arXiv:2103.05440.

Lichte, D., S. Marchlewitz and K.-D. Wolf (2016). A Quantitative Approach to Vulnerability Assessment of Critical Infrastructures With Respect to Multiple Physical Attack Scenarios. In: Future Security 2016, Proc. intern. conf., Berlin, Germany.

Rucht, D. (2023). Die Letzte Generation. Beschreibung und Kritik. Berlin: ipb working papers.

Stober, R. (2024, March). Der Entwurf des KRITIS-Dachgesetzes: Ein Rechtsrahmen mit offenen Flanken. In FORSI-Jahresband 2023 Der Schutz Kritischer Infrastrukturen (KRITIS) (pp. 25-28). Richard Boorberg Verlag GmbH & Co KG.

Termin, T., Lichte, D. and K.-D. Wolf (2021): Physical Security Risk Analysis for Mobile Access Systems Including Uncertainty Impact. In: Proceedings of the 31st European Security and Reliability Conference (Angers, France, 19.-23. Sept. 2021). Hrsg. von B. Castanier; M. Cepin; D. Bigaud; C. Berenguer; ISBN: 978-981-18-2016-8; doi:10.3850/978-981-18-2016-8_175-cd.

Termin, T., Lichte, D. and K.-D. Wolf (2023). Risk Adjusting of Scoring-based Metrics in Physical Security Assessment.