Journal of Secure Digital Technologies
Volume 1 (№2) September 2026

ISSN (E): 3136-9340
Sertificate № KZ19VPY00128988, date of issue: 25.12.2025

The journal is published 4 times a year
(March, June, September, December)

EDITORIAL TEAM
Editor-in-Chief
Dr. Leila Rzayeva, Association prof., PhD, Digital government support center (Kazakhstan)

Deputy Editor-in-Chief
Dr. Ali Myrzatay, PhD, Digital government support center (Kazakhstan)

Executive Secretary
Dr. Abdenova Zhuldyz, PhD, Digital government support center (Kazakhstan)

EDITORIAL BOARD MEMBERS (KAZAKHSTAN)
Seitkulov E.N. – Candidate of Physical and Mathematical Sciences, Professor Computer Science, Computer Engineering and Control
Abdenova G.A. – PhD in Automation and Control, Candidate of Technical Sciences, Professor L.N. Gumilyov Eurasian National University
Atanov S.K. – Doctor of Technical Sciences, Professor.
Sergazin G.K. – PhD, Associate Professor Computer Science, Computer Engineering and Control
Satybaldina D.Zh. – Candidate of Physical and Mathematical Sciences, Professor Director, Research Institute of Information Security and Cryptology, L.N. Gumilyov Eurasian National University

INTERNATIONAL EDITORIAL BOARD MEMBERS
Ibraheem Shayea – PhD in Electrical and Electronic Engineering Istanbul Technical University (Turkey)
Yadigar Nasiboglu Imamverdiyev – DSc in Technical Sciences, Associate Professor Information Security, Computer Science and Computer Technology (Azerbaijan)
Erdal Irmak – PhD in Electrical and Electronics Engineering, Professor Gazi University (Turkey)
Korhan Kayisli – PhD in Electrical Electronics Engineering, Associate Professor Gazi University (Turkey)
Yonghao Wang – PhD in Electronic Engineering, Associate Professor Queen Mary, University of London (UK)
Amir Mosavi – PhD in Applied Informatics, Professor of AI Obuda University (Hungary)
Gerald Feldman – PhD in General Computing, Associate Professor Birmingham City University (UK)
Mohamed Ragab – PhD in Computer Science, Assistant Professor Birmingham City University (UK)
Seref Sagiroglu – PhD in System Engineering, Professor Gazi University (Turkey)

MANAGING EDITOR
Dr. Ali Myrzatay, PhD, Digital government support center (Kazakhstan)

TECHNICAL EDITOR
Dr. Ali Myrzatay, PhD, Digital government support center (Kazakhstan)

CONTENTS

Computational Assessment of a Reconstructed Fuzzy Model for Categorizing Critical Digital Objects 3

Cyber Incident Damage Assessment for Public Digital Systems and Critical Infrastructure 27

Verifiable but Not Correct: Blockchain Trust Boundaries for Autonomous Agents 61

Calibrated Digital Twins and Geometry-based Perception for Automated Driving in Heavy Rain and Snowfall: Evidence From Research Testing in China and Implications for Kazakhstan. 78

Personal Data Protection in Kazakhstan's AI Enabled Public Services 99

Asking Another Jurisdiction What It Holds Without Telling It What You Hold: Private Artifact Matching with a Ledger-Anchored Transcript 121

On-Chain and Off-Chain Data Placement in Hybrid Blockchain Health Systems 133

An Integrated Project Management Maturity Model for Digital Qazaqstan Initiatives: Embedding Cybersecurity, AI Governance and Forward-Looking Cyber-Risk Requirements into ST RK ISO 21502-2022 148

Privacy Risks and Data Protection Readiness in AI Marketing in Kazakhstan 173

Khassanov Fakhreddin
MSc, Manager, Information Security Center
fakhreddin.khassanov@gmail.com, 0009-0000-5439-8171
National Institute for Cyber Policy and Research, RSE on REM "Digital Government Support Center", Astana, Kazakhstan

Leila Rzayeva
PhD, Associate Professor
e-mail: l.rzayeva@govtec.kz
ORCID: 0000-0002-3382-4685
Digital Government Support Center, Astana, Republic of Kazakhstan

Nursultan Nyssanov
MSc, Information Security Specialist, Information Security Center
n.nysanov@govtec.kz, 0009-0002-8128-0595
National Institute for Cyber Policy and Research, RSE on REM "Digital Government Support Center", Astana, Kazakhstan

COMPUTATIONAL ASSESSMENT OF A RECONSTRUCTED FUZZY MODEL FOR CATEGORIZING CRITICAL DIGITAL OBJECTS

Abstract: Before a categorization model informs protection obligations, its logical consistency and sensitivity need to be understood. We examine a reproducible reconstruction of a three-level Mamdani model for critical digital objects. The assessment covers eighteen scoring indicators, a separate automatic-assignment field, and confidentiality, integrity and availability scenarios. Documented membership functions and rules are retained, while missing normalization values are identified as reconstruction assumptions. Code, random seeds and generated data are described for reproduction. In 30,000 synthetic passports, all tested inference stages had nonzero aggregated membership, and scores ranged from 16.12 to 83.51. We observed 56 score groups at a tolerance of 10^{-8}, including 1,768 passports in [65, 70). Of 28,166 applicable one-level worsening steps, 1,237 (4.39%) lowered the score. None reduced decision status under the original thresholds, but setting the first-category threshold to 68 yielded 24 first-category reversals. Analytical and targeted tests show that a fully critical consequence is preserved at the severity layer; this does not guarantee first-category assignment. Scores were positively correlated with an equal-weight additive comparator ($r = 0.405$). Joint shifts in assumed normalization values changed 6.66% to 31.95% of final outcomes. These findings point to substantial calibration dependence and threshold-sensitive stability. Further specification, testing and external validation are needed before regulatory suitability can be assessed.

Keywords: critical digital objects; fuzzy inference; computational assessment; monotonicity; normalization sensitivity; non-compensation; reproducibility; cybersecurity.

Вычислительная оценка реконструированной нечёткой модели категоризации критически важных цифровых объектов

Аннотация: До использования результатов категоризации для установления требований к защите необходимо проверить логическую согласованность модели и её чувствительность к параметрам. В работе исследуется воспроизводимая реконструкция трёхуровневой модели Мамдани для критически важных цифровых объектов. Восемнадцать расчётных показателей и отдельный признак автоматического присвоения категории рассматриваются для сценариев нарушения конфиденциальности, целостности и доступности. Сохранены документированные функции принадлежности и правила; допущения при нормировании обозначены явно, предоставлены код, начальные состояния генераторов и результаты расчётов. В базовой выборке из 30 000 синтетических паспортов нулевая агрегированная принадлежность не обнаружена; баллы находились в пределах 16,12–83,51. При допуске 10^{-8} выделено 56 групп выходных значений; 1 768 паспортов попали в интервал [65; 70). Из 28 166 применимых шагов ухудшения 1 237 (4,39%) снизили балл. При исходных порогах понижения статуса решения не выявлены, однако снижение порога первой категории до 68 баллов привело к 24 случаям утраты первой категории после ухудшения показателя. Аналитически и вычислительно подтверждено сохранение полностью критического последствия на уровне тяжести, что не означает безусловного присвоения первой категории. Корреляция с равновесовой аддитивной моделью положительна ($r = 0,405$). Совместное изменение восстановленных значений нормирования изменило итоговые решения для 6,66–31,95% паспортов. Результаты показывают зависимость от калибровки и порогов и обосновывают дальнейшую доработку и проверку модели.

Ключевые слова: критически важные цифровые объекты; нечёткий вывод; вычислительная оценка; монотонность; чувствительность к нормированию; некомпенсаторность; воспроизводимость; кибербезопасность.

Аса маңызды цифрлық объектілерді санаттаудың қайта құрылған бұлыңғыр моделін есептік бағалау

Аңдатпа: Санаттау нәтижелерін қорғау талаптарын белгілеу үшін қолданар алдында модельдің логикалық келісімділігі мен параметрлерге сезімталдығын тексеру қажет. Зерттеуде аса маңызды цифрлық объектілерге арналған Мамданидің үш деңгейлі моделінің қайта өндіруге болатын нұсқасы қарастырылады. Құпиялылықтың, тұтастықтың және қолжетімділіктің бұзылу сценарийлері бойынша он сегіз есептік көрсеткіш және санатты автоматты тағайындауға арналған бөлек белгі бағаланады. Құжатталған тиістілік функциялары мен ережелер сақталып, нормалау кезіндегі жорамалдар нақты көрсетілді; код, кездейсоқ сандар генераторларының бастапқы мәндері және есептеу деректері берілді. 30 000 синтетикалық паспорттан

тұратын негізгі таңдамада агрегатталған тиістіліктің нөлдік мәні анықталмады; баллдар 16,12–83,51 аралығында болды. 10^{-8} төзімділік кезінде шығыс мәндерінің 56 тобы анықталды, 1 768 паспорт [65; 70) аралығына түсті. Көрсеткішті нашарлатудың 28 166 қолданылатын қадамының 1 237-сі (4,39%) баллды төмендетті. Бастапқы шектерде шешім мәртебесінің төмендеуі байқалмады, алайда бірінші санат шегін 68 балға төмендету көрсеткіш нашарлағаннан кейін бірінші санаттан айырылудың 24 жағдайын туғызды. Толық сыни салдардың ауырлық деңгейінде сақталуы аналитикалық және есептік жолмен расталды, бірақ бұл бірінші санаттың шартсыз берілуін білдірмейді. Тең салмақты аддитивті модельмен корреляция оң болды ($r = 0,405$). Қайта құрылған нормалау мәндерін бірге өзгерту паспорттардың 6,66–31,95%-ында қорытынды шешімді өзгертті. Нәтижелер модельді одан әрі жетілдіру және тексеру қажеттігін көрсетеді.

Түйін сөздер: аса маңызды цифрлық объектілер; бұлыңғыр қорытынды; есептік бағалау; монотондылық; нормалауға сезімталдық; компенсацияламау; қайта өндіру мүмкіндігі; киберқауіпсіздік.

1 Introduction

Critical digital objects underpin public services, economic activity and, in some cases, physical processes. A disruption may therefore have consequences far beyond the value of the equipment involved. Categorization should reflect those consequences and provide a clear rationale for the protection requirements that follow. Kazakhstan's Digital Code establishes a legal framework for critical digital objects [1]. The computational approach studied here remains a methodological proposal, not an approved regulatory instrument.

Identifying a critical object is not the same as assigning it a criticality category. ENISA's guidance links essential services to the infrastructure on which they depend [2], while studies of critical-infrastructure interdependencies show how disruptions can spread across technical, physical and organizational boundaries [3–5]. This motivates examining not only the severity of an incident, but also its geographic reach, potential cascading effects and the availability of substitutes.

Fuzzy sets and linguistic variables can represent expert assessments that do not fit neatly into fixed numerical classes [6, 7]. Mamdani inference translates those assessments into a rule-based scoring process [8]. Related methods have been used for failure prioritization [9], cyber-power grids [10], refinery assets [11] and hybrid FMECA [12]. More recent work considers richer representations and more efficient inference [13]. These studies motivate the approach, but do not validate the particular rules or thresholds used here. The aggregation and defuzzification choices still need to be assessed in the complete model [14, 15].

Matching a known example is a useful initial check, but it says little about how a model behaves elsewhere. Even if an individual aggregation operation is monotone, the full

sequence of fuzzification, rule activation and defuzzification may not be. It is also important to distinguish what a test demonstrates from what it merely fails to observe: one score decrease disproves monotonicity for the tested implementation, whereas finding no decreases in a random sample cannot prove monotonicity everywhere. Likewise, an interval with no sampled scores is not necessarily unreachable.

We therefore ask how consistently a fully specified reconstruction behaves under controlled tests, and how much its conclusions depend on normalization values missing from the original description. The analysis addresses four questions: RQ1, whether the tested inputs are covered by the inference rules and what scores they produce; RQ2, whether worsening one indicator can lower a score or decision status; RQ3, which severe-consequence preservation properties can be established independently of an additive comparison; and RQ4, how normalization choices and threshold adjustments affect the results.

This study contributes a reproducible computational assessment of the Variant 4 model [16]. It documents the experimental parameters and the provenance of normalization values, distinguishes fuzzy scores from automatic category assignment, and examines counterexamples under alternative thresholds. We also establish a limited analytical preservation result at the consequence-severity layer. We do not introduce a new inference algorithm, claim exhaustive coverage of the passport domain or validate the model against actual regulatory decisions. The V4-R1 experiments reported here were run anew and replace an earlier numerical series for which the underlying tables, weights, code and random seeds were not available.

2 Materials and methods

2.1 Source model and scope of reconstruction

Our starting point is the unpublished working draft, ‘Methodology for categorizing critical digital objects. Variant 4: a two-level model based on fuzzy inference’ [16]. The source document is identified in Supplementary Material S1 by its SHA-256 digest because it has no formal issue date or version identifier beyond ‘Variant 4’. We call our implementation V4-R1 to distinguish this reconstruction from the original authors' software, which was not available for inspection.

Variant 4 describes the passport fields, seven groups of membership functions, three rule bases, an 800-node defuzzification grid and three automatic-assignment grounds. Its normalization tables, however, are incomplete. We retain the documented functions and rules and make the missing numerical choices explicit in Table 1. Of the 63 mapping entries, 15 are anchored in the source, three cascade values appear in the supplied manuscript, and 45 are reconstruction assumptions. None of these assumptions is presented as an official parameter. The passport contains eighteen scoring fields (I5–I17 and I19–I23) and a separate binary automatic-assignment field (I24). There is no I18. This accounts for the apparent gap in indicator numbering.

2.2 Inference and decision stages

For each breach scenario s, passport p is mapped to six inputs on [0, 1]: societal, state and economic consequences (X1–X3), geographic extent (X4), cascading effects (X5) and irreplaceability (X6). Higher normalized values consistently indicate more severe consequences or less scope for substitution. Appendix B gives the scenario-specific mappings. Figure 1 shows the numerical inference chain separately from the automatic grounds that can override its final category.

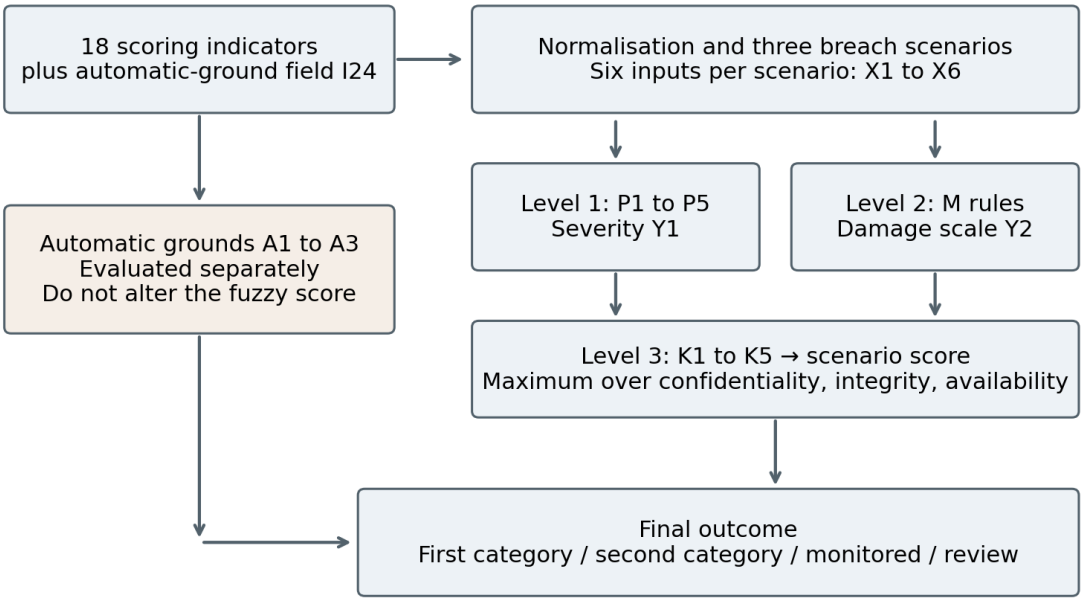


Figure 1. Computational structure of V4-R1. Automatic grounds override the decision outcome without changing the fuzzy score.

Membership functions are trapezoidal, with parameters (a, b, c, d). Membership equals one on the closed core [b, c], increases linearly from a to b, decreases linearly from c to d and is zero elsewhere. At coincident shoulder endpoints a = b or c = d, the core condition takes precedence. This convention prevents an endpoint from being assigned both zero and one. All parameter tuples appear in Appendix A. Conjunction, disjunction and negation are defined as follows:

$$u \wedge v = min(u,v), \quad u \vee v = max(u,v), \quad \neg u = 1 - u \quad (1)$$

Two branches are evaluated in parallel. Five Level 1 rules combine X1–X3 into consequence severity Y1, while ten Level 2 rules combine X4–X6 into damage scale Y2. Five Level 3 rules then produce the scenario score C_s from Y1 and Y2. Appendix B lists the antecedents, consequents and suppression factors. For each output term τ, capτ is the largest

activation among rules leading to that term. The aggregated membership function is defuzzified using a discrete centroid:

$$\mu*(x_i) = max\tau\, min(cap\tau, \mu\tau(x_i)), \quad Y = \Sigma_i\, x_i\, \mu*(x_i) \,/\, \Sigma_i\, \mu*(x_i) \quad (2)$$

The grid contains 800 equally spaced nodes including both endpoints, on [0, 1] for Y1 and Y2 and on [0, 100] for C_s. A zero denominator invokes the source's lower-endpoint fallback. Such a case is recorded separately as uncovered inference; receiving a numerical fallback is not counted as successful rule coverage. The maximum scenario score is

$$C(p) = max\,\{Cconf(p), Cint(p), Cavail(p)\} \quad (3)$$

Without an automatic ground, $C \geq 70$ yields the first category, $65 \leq C < 70$ yields referral for expert review, $30 \leq C < 65$ yields the second category, and $C < 30$ yields monitored status. The review interval is a procedural status, not a third criticality category. Automatic first-category assignment applies if I24 confirms a defence, emergency or mobilization role (A1); if the object is the sole provider of a national-level service (A2); or if at least two sectors are affected, dependent CDOs are present and substitution requires weeks or is impossible (A3). All comparisons use unrounded values. Membership-based K2 and K3 flags are retained in the output data, but they are not silently substituted for the score-based review interval: the source does not specify an operational firing threshold for converting these graded flags into referrals.

2.3 Complete experimental normalization

Table 1 defines the ordered levels used to generate the synthetic passports and their numerical mappings. S denotes a numerical anchor in Appendix 2 of the source methodology, M an anchor appearing only in Table 1 of the supplied manuscript, and R an assumption introduced for reconstruction. Values between fixed anchors are interpolated by ordinal level. Where no numerical anchor exists, the listed ordered values are explicit experimental choices. Such unanchored scales cannot be recovered by interpolation alone.

Table 1. Scoring indicators and normalization used in V4-R1

Indicator	Ordered levels	Normalized values and provenance
I5 Service significance	ordinary / sectoral / national	0.1 R / 0.6 R / 0.95 R
I6 Public service reach	none / individual users / mass public service	0.05 R / 0.45 R / 0.9 R
I7 Provider substitutability	one of many / main provider / sole provider	0.1 R / 0.65 R / 0.95 R
I8 Operating regime	working hours / extended hours / continuous	0.3 R / 0.6 R / 0.9 R

Indicator	Ordered levels	Normalized values and provenance
I9 Number of users	up to 10000 / 10001 to 124999 / 125000 to 500000 / 500001 to 2000000 / over 2000000	0.1 S / 0.325 R / 0.55 S / 0.75 R / 0.95 S
I10 Territorial coverage	one region / two to four regions / five or more regions / whole country	0.3 S / 0.55 R / 0.8 S / 0.95 S
I11 Sectoral share	up to 10 percent / over 10 to 30 percent / over 30 percent	0.1 R / 0.5 R / 0.9 R
I12 Dependent objects	fewer than 10 / 10 to 100 / over 100 / dependent CDO present	0.1 R / 0.35 R / 0.65 R / 0.95 R
I13 Data category	open / personal / special personal / restricted	0.05 S / 0.4 R / 0.75 S / 0.9 S
I14 Personal records	up to 10000 / 10001 to 50000 / 50001 to 100000 / 100001 to 300000 / over 300000	0.1 R / 0.3125 R / 0.525 R / 0.7375 R / 0.95 R
I15 Legal significance of data	none / supporting / rights establishing	0.05 R / 0.55 R / 0.95 R
I16 Integrity recoverability	full / partial / impossible	0.05 R / 0.55 R / 0.95 R
I17 Cascade character	none / within organisation / within sector / other sectors or CDOs	0.05 M / 0.35 R / 0.65 M / 0.95 M
I19 Affected sectors	one / two / three	0.15 R / 0.55 R / 0.95 R
I20 Substitution time	hours / days / weeks / impossible	0.15 S / 0.45 R / 0.75 S / 0.95 S
I21 Time to consequences	at least one day / hours / minutes	0.2 R / 0.55 R / 0.9 R
I22 Redundancy	geographical / partial / none	0.05 R / 0.55 R / 0.95 R
I23 Physical process link	none / indirect / direct	0.05 S / 0.5 S / 0.9 S

The original draft does not define every intermediate count range for I9 and I14. The boundaries shown in Table 1 are therefore experimental choices, not regulatory cutoffs. Where both I12 descriptions apply, the dependent-CDO state takes precedence over the count state. Including I24, the encoded domain contains 9,069,926,400 possible level

combinations. This is a Cartesian product of the declared fields; many combinations may not describe plausible real objects.

To test sensitivity to those choices, we construct two additional variants by shifting all reconstruction-only (R) values together by −0.10 or +0.10. We clip the values to [0, 1] and to the nearest fixed S or M anchors, preserving indicator order and the source values. All three variants use the same passports and worsening pairs. These coordinated changes are sensitivity tests, not a probability model for parameter uncertainty or bounds on every possible reconstruction.

2.4 Samples and comparison design

We generate a baseline of 30,000 synthetic passports. For each of the eighteen scoring indicators, an ordinal level is sampled independently with equal probability; I24 is independently set to 'yes' with probability 0.05. Generation uses NumPy's PCG64 generator with seed 20261021. Two further samples of 30,000 use seeds 20261024 and 20261025. This design deliberately explores combinations of declared levels. It does not reproduce the distribution of real critical digital objects in Kazakhstan and should not be used to estimate their prevalence.

The monotonicity experiment uses seed 20261022 to draw 40,000 passport-row and scoring-indicator pairs, both with replacement. In each pair, we worsen the selected indicator by one ordinal level and leave every other field, including I24, unchanged. A step is skipped when the indicator is already at its highest level. We retain row identifiers so duplicate tests can be distinguished from unique pairs. Automatic grounds are recalculated after every applicable change.

For comparison, we use a simple additive score, $A(p) = (100/18)\Sigma_j \varphi_j(p)$, with equal weight 1/18 for every scoring indicator. These are deliberately chosen weights, not a reconstruction of the undocumented expert weights used in an earlier comparison. We select the same number of first-category passports as the fuzzy score does before automatic overrides. Additive scores are ranked without rounding, and sample identifiers break ties. Equalizing the number selected makes the sets comparable in size, not necessarily in quality.

We also test the inference core directly using 30,000 independent vectors drawn uniformly from $[0, 1]^6$ (seed 20261023). This bypasses the mapping from passport fields to model inputs. Some vectors may not correspond to any valid passport, so the experiment does not establish passport-level reachability. Instead, it helps determine whether a restricted-looking score distribution comes from the inference rules themselves or from the chosen input mappings.

2.5 Properties and test criteria

Coverage and empirical output. For every evaluated scenario, the test records finite output, the permitted output interval and a nonzero centroid denominator at all three levels. Sorted output values are grouped by a greedy rule: a group extends only while its largest

value differs from its first value by at most $\varepsilon = 10^{-8}$ score points. This bounds each group's diameter and avoids chaining pairwise closeness. Counts are also computed at other tolerances. A finite list of observed groups is a property of the evaluated sample and numerical convention, not a proof of the full attainable set.

Monotonicity. Let p' differ from p only by a one-level worsening of one scoring indicator. Score monotonicity requires $C(p') \geq C(p)$; a computational violation is recorded when

$$C(p') < C(p) - \varepsilon, \quad \varepsilon = 10^{-8} \quad (4)$$

We separately compare first-category membership and decision status. For diagnostic ordering only, statuses are coded monitored = 0, second category = 1, review referral = 2 and first category = 3. This convention treats removal of a required referral as a decrease in procedural scrutiny; it does not turn referral into a legal category. Category and referral comparisons use the documented thresholds and recomputed automatic grounds. Rates are descriptive of the sampling procedure. Because pairs can reuse passports and include duplicates, no independence-based confidence interval is attached to the violation rate.

Severe-consequence preservation. We define a restricted non-compensation property at the consequence layer: once one of X1–X3 is fully critical, changing the other two consequence inputs cannot reduce Y1. Full critical membership occurs for $X_j \in [0.90, 1]$. The property is checked analytically from P1–P5 and numerically at $X_j = 0.90$, for each of the three choices of j, while the other inputs span $\{0, 0.05, \ldots, 1\}^2$. This gives 1,323 cases. Separately, the scale inputs span $\{0, 0.05, \ldots, 1\}^3$, giving 9,261 combinations. We test the resulting scores and the conditional first-category property when the scale has positive high membership and zero low membership. This separates preserved consequence severity from the further scale requirements of the decision rule.

Threshold sensitivity. For any first-category threshold t, the binary first-category decision is $F_t(p) = 1$ when $C(p) \geq t$ or an automatic ground holds. We calculate the score-only and final first-category shares separately. For a worsening pair with no automatic ground after the change, a first-category reversal occurs exactly when

$$C(p') < t \leq C(p) \quad (5)$$

The source's automatic predicates are nondecreasing under the declared ordinal changes, so a pair that has no automatic ground after the change also had none before it. Equation (5) therefore identifies the threshold interval exposed by each such score decrease. We evaluate integer thresholds from 35 to 85 and all relevant observed score endpoints. For full decision-status sensitivity, the review band retains width five, $[t - 5, t)$, and the monitored threshold remains 30. These alternatives are diagnostic recalibrations, not proposed regulatory thresholds.

2.6 Numerical implementation and reproducibility

The implementation uses Python 3.12.14 and NumPy 2.3.5 with binary64 arithmetic. The archive contains model.py, specification.json, run_experiments.py, figure-generation code, exact seeds, the source methodology, passport and score CSV files, paired monotonicity records, output groups and machine-readable summaries. The code stores all scenario scores, intermediate values, automatic-ground flags and K2/K3 activations. A reference-example check and shoulder-endpoint assertions are executed before the main experiments. Numerical sensitivity is assessed by repeating the entire baseline calculation at 400 and 1,600 centroid nodes and comparing it with 800 nodes. These checks establish specific implementation agreements; they do not substitute for independent code review or full formal verification.

3 Results

3.1 Agreement with the reported reference inputs

Table 2 reproduces the six scenario inputs reported for the regional heat-supply example in the supplied manuscript. V4-R1 returns 49.322933 for all three scenarios, matching the published rounded value of 49.3. The example also meets automatic ground A3: dependent critical digital objects are present, at least two sectors are affected, and substitution is impossible. Its fuzzy score therefore implies the second category, while the automatic rule assigns the first. This agreement checks the stated intermediate inputs and override logic; it does not reconstruct every field of the original passport.

Table 2. Reference check using the reported scenario inputs

Scenario	X1	X2	X3	X4	X5	X6	Y1	Y2	C
Confidentiality	.35	.45	.40	.30	.95	.95	.525000	.557154	49.322933
Integrity	.55	.45	.55	.30	.95	.95	.452633	.557154	49.322933
Availability	.55	.60	.55	.30	.95	.95	.452633	.557154	49.322933

This single-point agreement is reassuring but limited. It tests only the rules activated by that example and cannot recover omitted normalization entries, establish coverage of every rule combination or validate the model for other types of object.

3.2 Coverage and empirical output distribution

All 90,000 scenario calculations for the baseline passports produced nonzero aggregated membership at every inference level, and all maximum scores stayed within [0, 100]. The sample minimum was 16.117982 and the maximum 83.511428; the mean was 69.109277 and the median 82.661771. These figures describe the sampled passports only. They do not rule out other scores elsewhere in the full domain.

Table 3. Baseline outcomes before and after automatic assignment

Outcome	Score count	Score share	Final count	Final share
First category	16056	53.52%	18883	62.94%
Review referral	1768	5.89%	1394	4.65%
Second category	11931	39.77%	9515	31.72%
Monitored	245	0.82%	208	0.69%

At ε = 10^{-8}, 56 output groups were observed. The same count was obtained for tolerances from 10^{-10} to 10^{-4}; grouping with tolerance 0.005 merged one additional pair and gave 55 groups. The largest observed gap was 12.137624 points, between 69.749153 and 81.886777. The review interval [65, 70) contained 1,768 passports (5.89%) before overrides and 1,394 (4.65%) after overrides. Figure 2 therefore contradicts a universal assertion that this interval is unreachable under the documented rule architecture. Its occupancy depends on the declared input mapping.

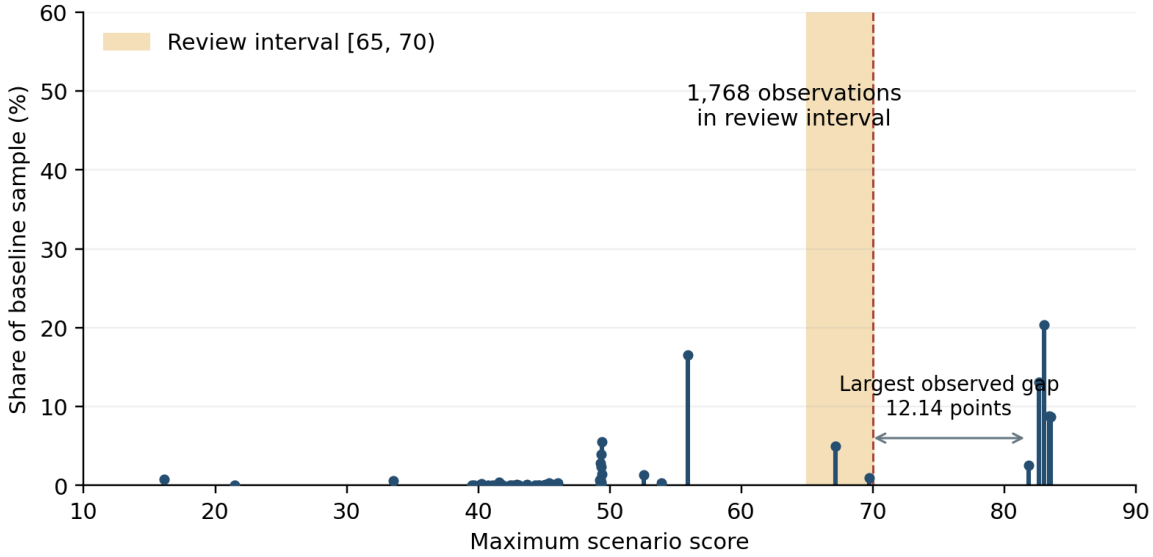


Figure 2. Empirical output distribution for the baseline 30,000 passports. Stems show observed score groups, not an enumeration of all attainable outputs.

Automatic grounds applied to 6,782 passports (22.61%). A1 applied to 1,516, A2 to 3,336 and A3 to 2,507; these groups overlap, so their counts cannot be summed. For 2,827 passports, automatic assignment alone changed the outcome to the first category. The final first-category count was therefore 16,056 + 2,827 = 18,883 (62.94%). Reporting the counts alongside the percentages avoids rounding discrepancies.

In 14,907 passports (49.69%), two or more scenarios tied for the maximum within 10^{-8} points. A first-maximum convention in the order confidentiality, integrity, availability selected these scenarios 10,343, 12,255 and 7,402 times. These counts are bookkeeping outcomes under the tie rule; they should not be interpreted as evidence that one breach type intrinsically governs more often.

3.3 Score monotonicity and decision status

Of 40,000 proposed one-level worsening steps, 28,166 were applicable. In 1,237 of these (4.39%), the fuzzy score fell by more than 10^{-8}. After removing repeated row–indicator pairs, 1,185 violations remained among 27,184 unique tests (4.36%). The effect therefore cannot be explained solely by duplicate sampling. Each recorded decrease is a counterexample to score monotonicity in V4-R1; the overall frequency across the full domain remains unknown.

Table 4. Magnitudes and consequences of observed score decreases

Quantity	Result
Applicable steps / distinct pairs	28,166 / 27,184
Score decreases / distinct-pair decreases	1,237 / 1,185
Median / mean decrease	0.365598 / 0.547403 points
95th percentile / maximum decrease	1.624651 / 9.208157 points
First-category reversals at t = 70	0
Decision-status decreases at original thresholds	0
I10 applicable steps / violations	1,596 / 760
I10 within-indicator violation rate	47.62%

Indicator I10 (territorial coverage) accounted for 760 decreases, or 61.44% of all observed violations. Among the 1,596 applicable I10 steps, 47.62% reduced the score. These percentages describe different denominators and should not be conflated. I19 (affected sectors) and I17 (cascade character) contributed 127 and 99 decreases, respectively. These results identify useful starting points for tracing the rule activations, but do not show that any one parameter change would resolve every violation.

The maximum worsening-related score reduction was greater than nine points, so the observed decreases are not all negligible numerical effects. Nevertheless, none of the tested pairs moved to a lower decision status at the original cutoffs of 30, 65 and 70. As Section 3.7 demonstrates, this apparent decision stability depends on the thresholds.

3.4 Severe consequences and the scope of non-compensation

A restricted preservation property follows directly from the specified consequence rules. If any Xj lies in [0.90, 1], its critical membership is one and its low membership is zero. Therefore P1 activates at one, P2, P3 and P5 are suppressed to zero, and P4 is zero. The aggregated Y1 output contains only the fully activated critical term. Its 800-node centroid is 0.844768, independent of the other two consequence inputs. Since this value lies inside the critical core of Y1, the downstream critical membership is also one. This proves preservation at the severity layer for a fully critical input under the stated numerical convention.

The 1,323 targeted severity cases all returned Y1 = 0.844768. Across the 9,261 scale combinations, final scenario scores ranged from 49.237220 to 83.511428, and 4,679 were below 70. Thus the strong claim that any fully critical consequence necessarily produces first-category assignment is false for this model. The rule base deliberately also requires sufficient scale. For the 4,120 tested combinations with high-scale membership greater than 10^{-8} and low-scale membership at most 10^{-8}, no score was below 70; the minimum was 81.567688.

The results support a precise statement: other consequence inputs cannot compensate away a fully critical consequence at the Y1 layer, and the final score preserves this influence at fixed scale. They do not establish non-compensation for partially activated critical terms, for arbitrary changes in territorial or cascade conditions, or for every individual passport indicator before X1–X3 are formed. A severity guarantee should therefore be specified at the layer where it is required and tested separately from the rule defining the first category.

3.5 Comparison with the additive model

The Pearson correlation between maximum fuzzy scores and equal-weight additive scores was 0.404610, indicating a positive association. The equal-volume comparison selected 16,056 first-category passports in each model. The additive cutoff was 52.916667; 46 passports had exactly the cutoff score in the implementation, of which 26 were selected by the declared sample-identifier rule. This tie handling is part of the experimental comparison, not a recommended administrative decision rule.

The fuzzy and additive first-category sets differed for 10,658 passports (35.53%). Of these, 5,329 were selected only by the fuzzy score and 5,329 only by the additive score. Figure 3 illustrates how passports with similar additive scores can receive quite different fuzzy scores. The comparison establishes that the models aggregate information differently; it does not tell us which selections are preferable, or prove non-compensation. We examine the latter property directly in Section 3.4. Other additive weights or calibration targets could produce different levels of agreement.

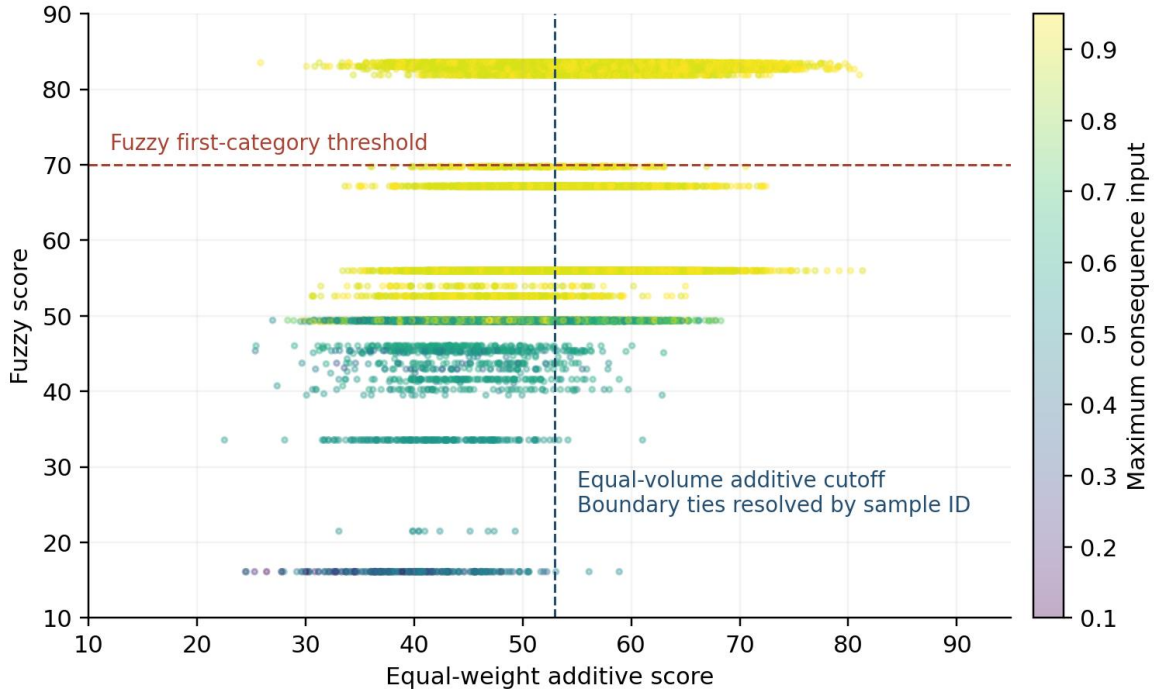


Figure 3. Fuzzy and equal-weight additive scores for the same baseline passports. Automatic grounds are excluded from both first-category sets in this comparison.

3.6 Sensitivity to reconstructed normalization

Table 5 and Figure 4 show that normalization uncertainty materially affects the findings. Changing only R entries, while retaining every fixed anchor and the entire rule architecture, changed the final outcome for 9,585 passports under the lower variant and 1,997 under the upper variant. These correspond to 31.95% and 6.66% of the same baseline sample. The largest absolute score changes were 34.106021 and 30.901588 points. The shifts were joint and coordinated, so these figures characterize sensitivity scenarios rather than confidence limits.

Table 5. Normalization sensitivity on identical passports and worsening steps

Quantity	Lower	Baseline	Upper
Observed output groups	48	56	38
First category by score	19.65%	53.52%	61.02%
Final first category	37.42%	62.94%	69.14%
Scores in [65, 70)	0	1,768	0
Final review referrals	0	1,394	0
Changed final outcome versus baseline	9,585	—	1,997
Score decreases in 28,166 steps	1,446	1,237	1,067
Decision-status decreases at original thresholds	0	0	0

The review band contained scores in the baseline sample, but was empty in both alternative normalization variants. Its occupancy therefore depends on values that the source methodology does not fully specify. The number of observed output groups also varied. Neither feature can be explained as an unavoidable consequence of the three-level rule architecture or centroid defuzzification alone.

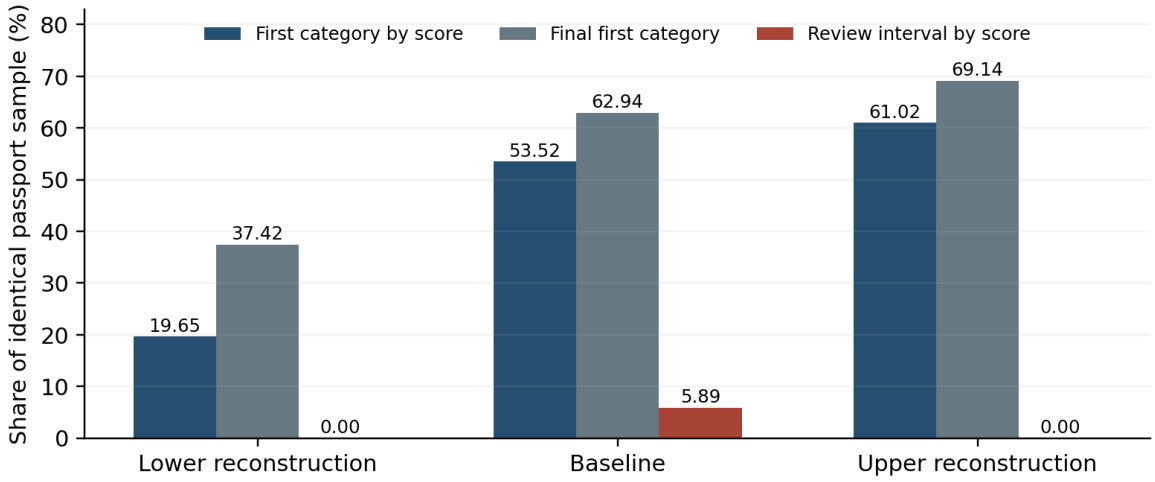


Figure 4. Effect of jointly shifting reconstructed normalization entries while preserving fixed anchors. Percentages refer to the same 30,000 passports.

Two independently generated samples, using seeds 20261024 and 20261025, produced 58 and 53 output groups and included 1,696 and 1,735 scores in the review interval. Their score-based first-category shares were 53.63% and 53.29%, while final first-category

shares were 63.03% and 62.52%. The broad distributional pattern was similar across these runs. The exact number of observed score groups, however, is a sample result, not the cardinality of the full attainable output set.

3.7 Threshold recalibration exposes decision reversals

First-category prevalence and monotonicity answer different questions. A threshold can produce an apparently stable aggregate share while concealing decreases for individual objects. At t = 70, no first-category reversal occurred in the tested worsening pairs. At t = 68, 24 reversals remained after automatic grounds were applied. At t = 83, there were 386 final first-category reversals. Table 6 uses exact tested thresholds; it does not describe a plateau as a closed interval when one endpoint already changes the result.

Table 6. Diagnostic threshold changes using unchanged scores and paired inputs

Threshold t	Score first count	Final first count	Final first reversals	Status decreases
55	22776	24004	0	0
68	16338	19097	24	24
70	16056	18883	0	0
73	16056	18883	0	24
83	11364	15430	386	386

For example, baseline passport 21667 produced 69.749153 points. Advancing indicator I19 by one ordinal level reduced the score to 67.160956, with no automatic ground before or after the change. Both scores lie in the original review interval, explaining the absence of a status change at t = 70. At t = 68, however, the first category is lost after worsening. The complete passport and paired record are included in S1. This observed counterexample shows why category stability must be retested after calibration.

For the baseline sample, the score-only first-category set is constant for thresholds t in (69.749153, 81.886777], where the endpoints shown are rounded for display and exact values are supplied in the data. The left endpoint is excluded because observations equal to it still qualify under $C \geq t$; the right endpoint is included because observations at that score remain selected. This is an empirical plateau of one sample, not a domain-wide insensitivity guarantee. It also does not guarantee unchanged referral status, because moving the lower edge of the review interval can change the treatment of lower scores.

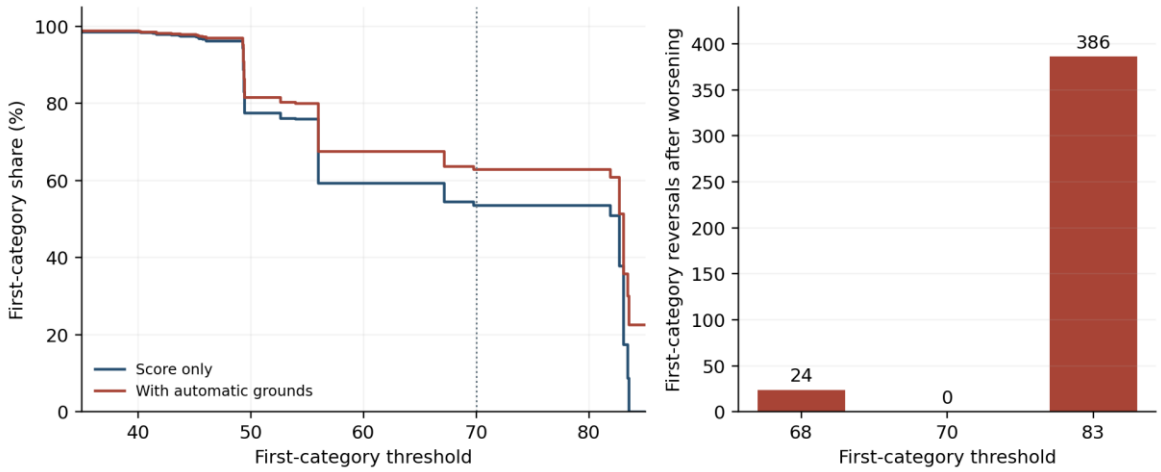


Figure 5. Threshold dependence of score-only and final first-category shares, and first-category reversals in the worsening tests. Alternative thresholds are diagnostic choices.

3.8 Numerical sensitivity and relaxed-domain diagnostics

Relative to the 800-node baseline, using 400 nodes changed a score by at most 0.120347 points and using 1,600 nodes by at most 0.060376 points. Neither alternative changed a final decision status at the original thresholds in the baseline sample. These results quantify sensitivity to the tested quadrature resolutions. They do not prove that every finer grid would yield identical categories, especially after thresholds are moved.

The continuous-input diagnostic produced 14,597 output groups and 277 scores in [65, 70), without invoking the zero-area fallback. This much richer distribution indicates that the inference architecture itself does not limit outputs to a few dozen discrete values. Because these input vectors do not necessarily correspond to valid passports, the result concerns the inference core rather than the full passport domain.

4 Discussion

4.1 What the computational evidence establishes

RQ1 receives a sample-limited positive answer: inference was covered at all stages for the tested passports and direct-input vectors. The source's fallback rule means that numerical finiteness alone would have been insufficient evidence; counting zero-area cases is necessary. The observed score distribution and its gaps characterize the implemented reconstruction and sampling design. They cannot be promoted to an exhaustive statement about the full passport domain.

For RQ2, the recorded one-level worsening pairs establish that V4-R1 is not score-monotone. The original thresholds masked these score decreases at the decision-status level in the tested pairs, but Equation (5) explains when a decrease becomes a category reversal.

The 24 reversals observed at a threshold of 68 show that this can occur near the original first-category threshold of 70.

RQ3 calls for a distinction between model layers. When one consequence input is fully critical, the consequence rule base preserves that severity regardless of the other two consequence inputs. Whether the object reaches the first category still depends on damage scale and any automatic grounds. The additive comparison shows a positive association but substantial disagreement between the scoring methods. It cannot establish which method better reflects the intended legal or substantive criteria.

For RQ4, the uncertainty surrounding normalization is substantial. Changing the reconstructed mappings altered review-band occupancy and final categories far more than changing the centroid grid from 800 to 400 or 1,600 nodes. Specifying and calibrating the passport-to-input mapping is therefore more urgent than interpreting the baseline histogram as an inherent model feature. Our coordinated lower and upper shifts are only two possibilities; varying individual mappings, bin limits or scenario assumptions could reveal other sensitivities.

4.2 Implications for model revision

Three changes should precede consideration of regulatory use. First, the methodology needs a complete versioned specification covering indicator domains, normalization, scenario equations, endpoint conventions, inference fallbacks and decision precedence. Parameters should be justified against the intended consequence definitions. Calibration should not merely target a preferred proportion of first-category objects in an artificial sample.

Second, the stored counterexamples should be traced through the intermediate memberships and activated rules. Territorial coverage and cascading effects are reasonable places to begin because they account for many observed decreases. Possible remedies include revising rule consequents or overlaps, or changing a hierarchical inference stage. None has been tested here. Any revision should be evaluated on the existing counterexamples, a broader targeted search and the threshold tests—not just a few selected cases.

Third, the purpose of expert review must be stated operationally. A numerical band, disagreement between adjacent memberships, K2/K3 flags and low confidence in passport evidence represent different referral criteria. Membership-based referral could complement a score band, but its activation rule, threshold and workload would require validation. An empty band in one calibration does not establish that all expert-review mechanisms are dormant, and moving a threshold cannot be assumed to repair the underlying score behavior.

4.3 Limitations and external validation

V4-R1 is an explicit reconstruction of an incompletely specified draft. The documented rule and membership parameters are retained, but 45 normalization entries and

some count-bin boundaries are assumptions. The baseline statistics therefore apply to this reconstruction, not automatically to the original unpublished implementation. Although the archive makes the new experiments reproducible, recovering the earlier code and parameter files remains necessary to explain the exact numerical differences from the previous version of the manuscript.

The synthetic sampling design does not reproduce the relationships seen in practice between service importance, user counts, territorial reach and technical dependencies. Some combinations may be implausible. Such cases can still be valuable for stress testing, but their frequencies say nothing about actual exposure or regulatory workload. We also lack independently assessed real objects and observed incident outcomes. Accordingly, false-positive and false-negative categorization rates cannot be estimated from this study.

The monotonicity tests cover sampled one-level changes, not every possible transition or simultaneous change. The normalization analysis considers only two coordinated alternatives, and the analytical preservation result concerns fully critical inputs at one particular layer. Before deployment, the model would need additional tests for missing or uncertain passport evidence, exhaustive or formal checks on manageable subdomains, expert-reviewed reference cases and validation against real consequences. Until then, claims about regulatory suitability must remain conditional.

5 Conclusion

We reconstructed and tested the Variant 4 fuzzy categorization model using explicitly documented parameters, random seeds and synthetic data. All tested inference paths were covered, but 1,237 of 28,166 applicable worsening steps lowered the score. Although none lowered decision status at the original thresholds, recalibrating the first-category cutoff to 68 produced 24 first-category reversals after automatic grounds were considered.

The output distribution also depended on reconstruction choices. The baseline included 56 observed score groups and 1,768 passports in the [65, 70) review interval, whereas the two alternative normalizations left that interval empty. A small set of sampled outputs—or an empty review band in one run—therefore cannot be treated as a universal property of the inference model. We established preservation of fully critical consequences at the severity layer, but not unconditional first-category assignment.

The main practical contribution is a reproducible set of counterexamples and a way to assess how normalization and threshold choices affect decisions. Before the model can be judged suitable for regulatory use, it needs an agreed and complete specification, an explanation or correction of its non-monotone responses, and validation on independently assessed real cases. Computational tests can inform that process, but cannot replace it.

Supplementary material

Supplementary Material S1 is supplied as Fuzzy_Categorization_Reproducibility.zip. It contains the source methodology, the V4-R1

specification, executable model and experiment scripts, figure-generation code, requirements, generated passport and score files, paired worsening records, diagnostic experiments and a results summary. The README identifies reconstruction assumptions and provides reproduction commands. The archive concerns the new V4-R1 experiments and does not claim to contain the unavailable historical implementation.

Author contributions

Conceptualization: L.R.; Methodology: F.K. and L.R.; Software: F.K. and N.N.; Validation: F.K., L.R. and N.N.; Formal analysis: F.K.; Investigation: N.N.; Data curation: N.N.; Writing — original draft preparation: F.K.; Writing — review and editing: L.R. and N.N.; Visualization: N.N.; Supervision: L.R.; Project administration: L.R.

Conflicts of interest

The authors declare no conflict of interest.

Data and code availability

The code and generated data supporting the V4-R1 results are included in Supplementary Material S1 supplied with this revision. No public repository DOI is claimed. The synthetic records do not represent identified real objects. The study makes no claim to provide the original implementation, random seeds or expert weights underlying the superseded numerical series.

References

[1] Digital Code of the Republic of Kazakhstan No. 255-VIII. (2026, January 9). Adilet Legal Information System. https://adilet.zan.kz/rus/docs/K2600000255

[2] Mattioli, R., & Levy-Bencheton, C. (2014). Methodologies for the identification of critical information infrastructure assets and services. European Union Agency for Network and Information Security. https://doi.org/10.2824/38100

[3] Rinaldi, S. M., Peerenboom, J. P., & Kelly, T. K. (2001). Identifying, understanding, and analyzing critical infrastructure interdependencies. IEEE Control Systems Magazine, 21(6), 11–25. https://doi.org/10.1109/37.969131

[4] Kelic, A., Warren, D. E., & Phillips, L. R. (2008). Cyber and physical infrastructure interdependencies (SAND2008-6192). Sandia National Laboratories.

[5] Malinichev, D. M., Kuchmezov, Kh. Kh., & Mochalov, V. V. (2022). Kategorirovanie vzaimosvyazannykh ob'ektov kriticheskoi informatsionnoi infrastruktury [Categorization of interconnected critical information infrastructure objects]. Applied Informatics, 17(3), 105–116. https://doi.org/10.37791/2687-0649-2022-17-3-105-116

[6] Zadeh, L. A. (1965). Fuzzy sets. Information and Control, 8(3), 338–353. https://doi.org/10.1016/S0019-9958(65)90241-X

[7] Zadeh, L. A. (1975). The concept of a linguistic variable and its application to approximate reasoning—I. Information Sciences, 8(3), 199–249. https://doi.org/10.1016/0020-0255(75)90036-5

[8] Mamdani, E. H., & Assilian, S. (1975). An experiment in linguistic synthesis with a fuzzy logic controller. International Journal of Man-Machine Studies, 7(1), 1–13. https://doi.org/10.1016/S0020-7373(75)80002-2

[9] Bowles, J. B., & Peláez, C. E. (1995). Fuzzy logic prioritization of failures in a system failure mode, effects and criticality analysis. Reliability Engineering & System Safety, 50(2), 203–213. https://doi.org/10.1016/0951-8320(95)00068-D

[10] Zúñiga, A. A., Fernandes, J. F. P., & Branco, P. J. C. (2023). Fuzzy-based failure modes, effects, and criticality analysis applied to cyber-power grids. Energies, 16(8), 3346. https://doi.org/10.3390/en16083346

[11] Bahadir, C., & Kahraman, C. (2026). Criticality analysis of refinery assets using picture fuzzy inference system in reliability centered maintenance. International Journal of Fuzzy Systems, 28(5), 1767–1781. https://doi.org/10.1007/s40815-025-02069-y

[12] Chakhrit, A., Djelamda, I., Bougofa, M., Guetarni, I. H. M., Bouafia, A., & Chennoufi, M. (2024). Integrating fuzzy logic and multi-criteria decision-making in a hybrid FMECA for robust risk prioritization. Quality and Reliability Engineering International, 40(6), 3555–3580. https://doi.org/10.1002/qre.3601

[13] Ontiveros, E., Melin, P., & Castillo, O. (2024). Towards an efficient approach for Mamdani interval type-3 fuzzy inference systems. International Journal of Fuzzy Systems, 26, 2172–2190. https://doi.org/10.1007/s40815-024-01722-2

[14] Grabisch, M., Marichal, J.-L., Mesiar, R., & Pap, E. (2009). Aggregation functions. Cambridge University Press.

[15] van Leekwijck, W., & Kerre, E. E. (1999). Defuzzification: Criteria and classification. Fuzzy Sets and Systems, 108(2), 159–178. https://doi.org/10.1016/S0165-0114(97)00337-0

[16] Methodology for categorizing critical digital objects. Variant 4: a two-level model based on fuzzy inference. Undated unpublished working draft. Source file “Вариант 4(3).docx”, archived in Supplementary Material S1; SHA-256 recorded in specification.json.

Appendix A Membership function parameters

The tuples below are transcribed from Appendix 4 of the source methodology. Every tuple gives (a, b, c, d). All variable domains are [0, 1], except the category-score variable C, whose domain is [0, 100]. Shoulder endpoints follow the core-first convention stated in Section 2.2.

Table A1. Complete membership-function specification

Variable	Term	Parameters
Severity X1–X3	low	(0, 0, 0.2, 0.4)
Severity X1–X3	medium	(0.2, 0.4, 0.5, 0.7)
Severity X1–X3	high	(0.5, 0.7, 0.78, 0.9)
Severity X1–X3	critical	(0.78, 0.9, 1, 1)
Territory X4	organisational	(0, 0, 0.2, 0.4)
Territory X4	regional	(0.2, 0.4, 0.5, 0.62)
Territory X4	interregional	(0.5, 0.62, 0.72, 0.88)
Territory X4	national	(0.72, 0.88, 1, 1)
Cascade X5	isolated	(0, 0, 0.25, 0.45)
Cascade X5	intrasectoral	(0.25, 0.45, 0.6, 0.8)
Cascade X5	crosssectoral	(0.6, 0.8, 1, 1)
Irreplaceability X6	high substitutability	(0, 0, 0.25, 0.45)
Irreplaceability X6	medium substitutability	(0.25, 0.45, 0.55, 0.72)
Irreplaceability X6	low substitutability	(0.55, 0.72, 0.82, 0.93)
Irreplaceability X6	irreplaceable	(0.82, 0.93, 1, 1)

Variable	Term	Parameters
Consequence criticality Y1	insignificant	(0, 0, 0.25, 0.45)
Consequence criticality Y1	significant	(0.25, 0.45, 0.6, 0.8)
Consequence criticality Y1	critical	(0.6, 0.8, 1, 1)
Damage scale Y2	organisational	(0, 0, 0.2, 0.4)
Damage scale Y2	sectoral regional	(0.2, 0.4, 0.52, 0.66)
Damage scale Y2	interregional	(0.52, 0.66, 0.76, 0.9)
Damage scale Y2	national	(0.76, 0.9, 1, 1)
Category score C	monitored	(0, 0, 20, 42)
Category score C	second	(20, 42, 58, 78)
Category score C	first	(58, 78, 100, 100)

Appendix B Scenario inputs and complete rule bases

Let φ_i be the normalized value of indicator Ii; $\wedge$ and $\vee$ denote minimum and maximum. The following equations reproduce the specified scenario mapping. For a confidentiality breach, the special X2 branch applies only to the restricted-data category.

Table B1. Consequence inputs by breach scenario

Scenario	X1 societal	X2 state	X3 economic
Confidentiality	$\varphi_{13} \wedge \varphi_{14}$	φ_{13} if restricted; otherwise $\varphi_{13} \wedge \varphi_6$	$\varphi_{11} \wedge 0.40$
Integrity	$[0.9(\varphi_{13} \wedge \varphi_{14})] \vee (\varphi_{23} \wedge \varphi_{16})$	$\varphi_{15} \vee (\varphi_6 \wedge \varphi_{16})$	$\varphi_{11} \wedge \varphi_{16}$
Availability	$(\varphi_{23} \wedge \varphi_9) \vee (\varphi_{21} \wedge \varphi_9)$	$\varphi_6 \vee (\varphi_5 \wedge \varphi_8)$	$(\varphi_{11} \wedge \varphi_{21}) \vee (\varphi_{12} \wedge \varphi_{21})$

$X4 = \varphi_{10}$. $X5 = \varphi_{17} \vee \varphi_{19} \vee w$, where $w = \varphi_{12} \wedge 0.80$ when dependent CDOs are present, and $w = 0$ otherwise. $X6 = \varphi_7 \vee \varphi_{20} \vee \varphi_{22}$, additionally combined by maximum with φ_{16} for the integrity scenario. Constants 0.9, 0.40 and 0.80 are retained from the source. They are modeling choices, not empirical estimates validated in this study; in particular, the confidentiality cap does not establish a general upper bound on real economic loss from a disclosure.

For level 1, l_j, m_j, h_j and c_j denote low, medium, high and critical severity memberships of Xj for j = 1, 2, 3. Let m(1) and m(2) be the largest and second-largest medium memberships. The notation αr denotes activation of rule r.

Table B2. Level 1 consequence rules

Rule	Activation	Y1 consequent
P1	$\max(c_1, c_2, c_3)$	Critical
P2	$\min[\max(h_1, h_2, h_3), 1 - \alpha P1]$	Significant
P3	$\min[m(2), 1 - \alpha P1]$	Significant
P4	$\min(l_1, l_2, l_3)$	Insignificant
P5	$\min[0.6m(1), 1 - \alpha P1, 1 - \alpha P2, 1 - \alpha P3]$	Insignificant

For level 2, o, r, i and n are the organizational, regional, interregional and national memberships of X4. The cascade memberships of X5 are z (isolated), v (within sector) and q (cross-sector). The irreplaceable membership of X6 is u.

Table B3. Level 2 scale rules

Rule	Activation	Y2 consequent
M1	n	National
M2a	i	Interregional
M2b	$q \wedge (r \vee i \vee n)$	Interregional
M2c	$q \wedge o$	Sectoral-regional
M3	$r \wedge (z \vee v)$	Sectoral-regional
M4	$o \wedge z$	Organizational
M4b	$o \wedge v$	Sectoral-regional
M5a	$u \wedge o$	Sectoral-regional
M5b	$u \wedge r$	Interregional
M5c	$u \wedge i$	National

For level 3, let $H = \mu_{national}(Y2) \vee \mu_{interregional}(Y2)$ and $L = \mu_{sectoral\text{-}regional}(Y2) \vee \mu_{organizational}(Y2)$. Let c, s and n denote the critical, significant and insignificant memberships of Y1, respectively. These symbols are local to Table B4.

Table B4. Level 3 category-score rules

Rule	Activation	C consequent and flag
K1	c ∧ H	First
K2	c ∧ L	Second; completeness-control flag
K3	s ∧ H	Second; review-related flag
K4	s ∧ L	Second
K5	n	Monitored

At every level, the cap for an output term is the maximum activation of its rules; the output is then defuzzified using Equation (2). Thus the category-score caps are αK1 for first, max(αK2, αK3, αK4) for second, and αK5 for monitored. The graded flags are exported for inspection. The source's score thresholds and automatic grounds determine the reported decision outcome.