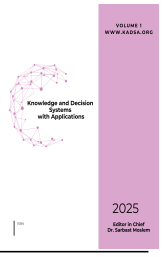
Knowledge and Decision Systems with Applications

www.kadsa.org
eISSN: 3079-6318



Artificial Intelligence-Enabled Reliability and Hazard-Rate Degradation Modeling for Dynamic Student Progression in Outcome-Based Educational Systems

Akbota Nurgaliyeva[1,*], Angisin Seitmuratov[1]

[1] Department of Mathematics, Korkyt Ata Kyzylorda University, Aiteke Bi Street 29A, 12000, Kyzylorda, Kazakhstan

ARTICLE INFO

Article history:
Received 25 September 2026
Received in revised form 20 October 2026
Accepted 29 October 2026
Available online 1 November 2026

Keywords:
Outcome-Based Education (OBE); Reliability Engineering & Hazard Rate; Academic Degradation Modeling; Predictive Educational Analytics; Intelligent Decision Support Systems; Artificial Intelligence

ABSTRACT

In Outcome-Based Education (OBE) environments, student academic progression represents a complex, non-linear dynamic system governed by multi-faceted indicators including performance trajectories, engagement metrics, and learning support mechanisms. Traditional evaluation methods often fail to capture time-dependent performance risks and degradation patterns across successive assessment stages. In this paper, we establish a novel reliability-based degradation modeling framework that adapts engineering reliability principles and hazard-rate survival analysis to educational decision-support systems. Utilizing a standardized, publicly available academic progress dataset, our framework models performance stability via cumulative reliability functions and identifies critical vulnerability thresholds using hazard-rate formulations. Students are classified into distinct progression cohorts (high, moderate, and low achievers) based on normalized performance distributions. Empirical findings demonstrate a strong correlation between early academic stability and long-term progression reliability, showing that multi-year failing trajectories exhibit exponentially higher degradation rates. The proposed framework provides an interpretable, data-driven analytical tool to monitor student risk, support early Artificial Intelligence (AI)-driven interventions, and optimize institutional decision-making in modern outcome-based learning architectures.

[*] *Corresponding author.*
E-mail address: akbota.sci@gmail.com

https://doi.org/10.59543/6m0k3s19

1. Introduction

Outcome Based Education (OBE) is an approach that focuses on the realisation of predetermined learning outcomes, competency building and the systematic evaluation of student performance. This model of education calls for a more nuanced approach to measuring students' academic growth than relying on grades or passing rates at the end of the year. It also involves knowing the change in students' academic achievement at various assessment points and how to identify students who might fall back in performance. This analysis can contribute quantitatively to academic mentoring, remedial teaching, and evidence-based educational planning.

The method of traditional education assessment is mainly descriptive indicators such as exam scores, attendance, study time, and cumulative results of education. These are useful signs of student's achievement but do not necessarily measure the continuity of satisfactory performance or the probability of decline between the assessment stages. Student performance data that is publicly available can be used to examine relationships between student engagement with their education and their behaviour in schools in quantitative ways [20].

The reliability theory provides a mathematical tool to analyze the consistency of good system performance over time. The reliability functions measure the likelihood of continued satisfactory functioning, while the hazard-rate functions measure the instantaneous risk of an adverse event under certain modelling assumptions. These ideas were developed in reliability engineering, but the mathematical model can be modified to the educational context, where the satisfactory condition is stated as the state of interest and the transition to the unsatisfactory condition is expressed as a degradation event. Adaptive preventive maintenance research is one example where system performance measures and tracking changes in system behaviour over time are important [8]. However, the meaning of reliability and hazard in an educational context must be sufficiently defined in terms of the performance criteria, assessment phases, and degradation events that constitute satisfactory performance.

Another source of analysis of the performance of students in heterogeneous classes is mathematical modelling under uncertainty. Rasool et al. [15] proposed multi-attribute decision-making approach based on Pythagorean fuzzy numbers and Ejegwa et al. [16] developed partial correlation coefficient technique based on intuitionistic fuzzy information. Shams et al. [17] discussed a semi analytical approach to intuitionistic fuzzy systems of differential equations and Madasi et al. [18] introduced a generalized decision-making framework based on a fuzzy set to select energy sources. Agilan et al. [19] have investigated existence and uniqueness of solutions of fuzzy fractional integro-differential equations with boundary conditions. Recent advances in neutrosophic methods and PINN-based prediction offer promising opportunities for intelligent decision-making under uncertainty [21]. The diversity of the contributions highlights the wide range of mathematical tools for dealing with uncertainty, decision making and modeling of the dynamics of systems. Nevertheless, they do not directly develop a model of academic degradation based on reliability and they handle different applications.

The current study's motivation is to create a mathematically interpretable framework that can be used to investigate academic stability and deterioration throughout the multiple assessment periods. Students can have a performance that is consistently satisfactory, moderate change in performance, or significant drop in performance. Measuring these trends could help teachers determine the phases in which they need to pay more attention and provide more support when needed. A reliability-based degradation model can be used in conjunction with traditional education statistics to depict the likelihood of staying in the satisfactory range and the risk of moving to an unsatisfactory range.

For this reason, this research is developed a reliability-based degradation process model for analysis on academic progression in this academic performance data [20] that publicly available. Academic variables available in the selected dataset are considered in the analysis, provided their definitions are verified and they are available, such as previous failures, period-wise grades, final grades, study time

and absences. Students are categorized according to normalized academic achievement measure as fast, moderate, and low achievers. Next, measures of reliability and hazard rates are used to examine the continuity of the satisfactory academic performance and the likelihood of worsening performance across the different assessment points.

This study has made the following contributions to the literature:

1. The academic performance degradation process model is developed based on the reliability.Academic performance degradation process model is formulated to represent academic performance in successive assessment stages in OBE context.

2. Reliability functions and hazard-rate measures are modified to measure the continuity of acceptable academic performance and to describe the likelihood that academic performance will decline.

3. A classification system is created based on performance that is calculated on a normed scale to identify three groups of fast, moderate, and low performance.

4. The proposed framework gives academic monitoring and planning for intervention a quantitative basis and it also supports the use of traditional educational performance measures.

The overall research gap of this study is the creation of an interpretable mathematical model that explores the persistence of satisfactory academic performance, as well as the likelihood of deterioration, across assessment stages. The proposed approach does not look at the level of achievement as an end result, but rather as a process of gradual attainment. The result of this analysis is for systematic academic monitoring and evidence-informed educational planning.

2. Literature Review

An assessment of academic performance is an integral part of the quality assurance in education and implement OBE. Traditional assessment practices typically include the summary of achievement based on examination results, pass rates, school roll and cumulative academic measures. These measures tell you that the performance varies, but not necessarily across successive assessments and do not measure the risk of deterioration.

Student performance data that is publicly available can be used to examine academic outcomes and their correlation with student characteristics [20]. Previous failures, period-wise grades, amount of study time and absences may be included in an empirical model if they are known and defined. The purpose of the present study is to examine academic progress, not to build causal links between student characteristics and achievement.

There may also be institutional and organizational conditions that affect educational performance. Clotfelter et al. [3] looked at teacher mobility, school segregation, and pay-based policies, and found evidence of the larger organizational context in which educational production takes place. While their study does not directly examine the degradation of students' academic performance at the level of the individual student, it shows that empirical results must be understood in the context of the educational performance and the institutional situation.

Reliability theory offers mathematical means of assessing the duration of adequate performance and the frequency of negative events. Reliability functions are used in engineering applications to state the probability that a system will meet certain operating conditions, whereas hazard functions provide a statement of the instantaneous hazard (or risk) of failure for systems that have not failed up to and at a certain point.

Öhman et al. [8] studied the measurement of service outcomes for adaptive preventive maintenance. Their work is an example of measurement across a domain in maintenance. Service maintenance is very different from educational assessment, but the salient feature of monitoring performance over time is the idea of the investigation of comparable measures of academic progression.

Applying the concepts of reliability in education demands a well defined satisfactory performance limit, and a uniform description of the assessment phases. Academic deterioration has to be operationally defined, such as moving from satisfactory to unsatisfactory grade level according to a criterion. These estimates should therefore be interpreted in relation to that definition of reliability and hazard. In particular, hazard rate is a conditional risk measure and not a causal impact of a student characteristic.

In the present study, this overall modelling view is applied to academic achievement, where each stage of assessment is considered as a snapshot of student achievement. It is hoped that the proposed framework will review the performance and determine the phases where deterioration is more likely, according to the defined performance indicators and the assumptions.

Analysing complex systems and making decisions when observations or assessments are imperfect can be done using mathematical approaches to uncertainty. Rasool et al. [15] proposed a multi-attribute decision making approach based on Pythagorean fuzzy numbers and applied it to the hotel evaluation problem. They illustrate their study by comparing alternatives based on multiple criteria with fuzzy information.

Ejegwa et al. [16] proposed a partial correlation coefficient technique based on intuitionistic fuzzy information for pattern recognition. Shams et al. [17] presented a semi-analytical scheme for intuitionistic fuzzy systems of differential equations and Madasi et al. [18] proposed a generalized fuzzy-set-based decision-making process for selection of energy sources. The examples of these studies demonstrate various uses of mathematical tools to deal with uncertainty, to compare alternatives and to model system behaviour.

In [19] Agilan et al. discussed the existence and uniqueness of solutions to fuzzy fractional integro-differential equations with boundary conditions. They are working on the mathematical analysis of uncertain dynamical systems. But the purpose of fuzzy modelling is different from the purpose of reliability-based degradation modelling: the former models and represents uncertainty, while the later quantifies the possibility of a specific model being in that state and the risk of entering an undesirable model state.

In the current research, reliability-based measures are used for analysis of academic progression. Although fuzzy approaches provide possibilities for future extensions under the circumstances that academic classifications or performance thresholds are inherently ambiguous, it is not assumed that they are equal to reliability measures used in this work.

There has been a recent focus on data-driven methods for modelling and analysis of complex systems. To assess structural mathematical reasoning ability in LLM, Zheng et al. [1] introduced a process-aware benchmark; for multi-arm data, Wan et al. [12] explored a causal rule ensemble approach. The studies reflect various methodological issues but show that the computational development of structured evaluation and data analysis is ongoing.

A few other works have referred to particular applications in pharmacometrics [2], healthcare analytics [5], risk analysis [6], applied sciences [7], employment [11] and organizational decision-making [10]. The application areas are not directly related to the students' academic trajectory, and therefore do not constitute definitive evidence for the model of education proposed. Only if there is a discussion of their relevant methodological or contextual contributions in the manuscript, should they be kept in the bibliography.

The literature discussed here includes educational data analysis, reliability-based performance assessment, mathematical modelling in the absence of certainty, and computational solutions to struc-

tured assessment. By themselves, however, these references do not provide evidence that such a combination of a reliability function, a hazard-rate measure, and an academic degradation index has already been made into a single model for academic progression based on OBE.

In the present study, we considered this modelling objective by identifying academic performance states during different assessment phases and analyzing how they persist or deteriorate in terms of the reliability measures. The proposed framework aims to be complementary to descriptive educational statistics in order to offer a quantitative quantitative index interpretable in terms of academic stability and degradation. This can be evaluated using the selected data set, clear threshold definitions, appropriate estimation methods and the validation of measurement.

3. Proposed Reliability-Based Academic Progression Framework

Let S_1, S_2, and S_3 denote the first assessment, second assessment, and final assessment stages, corresponding to the grade variables G_1, G_2, and G_3, respectively. Since the grades are measured on a scale from 0 to 20, the normalized performance score at each assessment stage is defined as

$$P_i = \frac{G_i}{20}, \qquad i = 1, 2, 3. \tag{1}$$

A student is considered to maintain satisfactory academic performance if $P_i \geq \theta$, where θ denotes the predefined performance threshold. In this study, $\theta = 0.50$ is adopted as the operational pass-level competency threshold. Since the original grades range from 0 to 20, this threshold corresponds to a grade of 10 and is applied consistently across all three assessment stages.

The reliability function at assessment stage S_i is defined as the probability that a student's normalized performance meets or exceeds the specified threshold:

$$R(S_i) = \Pr(P_i \geq \theta), \qquad i = 1, 2, 3. \tag{2}$$

Accordingly, the reliability measures for the three assessment stages are expressed as

$$R(S_1) = \Pr(P_1 \geq \theta), \tag{3}$$

$$R(S_2) = \Pr(P_2 \geq \theta), \tag{4}$$

and

$$R(S_3) = \Pr(P_3 \geq \theta). \tag{5}$$

These measures represent the proportions of students whose normalized performance meets the specified threshold at the first, second, and final assessment stages, respectively.

To quantify the relative reduction in satisfactory performance between consecutive assessment stages, the stage-wise hazard measure is defined as

$$H(S_i) = \frac{R(S_i) - R(S_{i+1})}{R(S_i)}, \qquad i = 1, 2. \tag{6}$$

This measure quantifies the relative decrease in the proportion of students meeting the performance threshold between consecutive stages. It is defined when $R(S_i) > 0$ and should be interpreted as a discrete academic deterioration measure rather than a conventional time-to-event hazard rate.

To quantify the overall level of academic degradation across the three assessment stages, the Academic Degradation Index (ADI) is defined as

$$ADI = 1 - \frac{1}{3}\sum_{i=1}^{3} P_i. \tag{7}$$

Since each normalized performance score lies in the interval $[0, 1]$, the ADI also lies in $[0, 1]$. A lower ADI indicates a higher average normalized academic performance, whereas a higher ADI indicates a lower average performance. The ADI provides an aggregate measure and does not independently identify the assessment interval in which performance deterioration occurs. Therefore, it is interpreted together with the stage-wise reliability and hazard measures.

For further analysis, students are classified into three groups according to their final normalized performance score P_3. Students with $P_3 \geq 0.75$ are classified as fast learners, those with $0.50 \leq P_3 < 0.75$ as moderate learners, and those with $P_3 < 0.50$ as low achievers. This classification provides a basis for comparing academic progression across different achievement groups.

The proposed framework integrates normalized academic performance, stage-wise reliability, relative deterioration measures, and the Academic Degradation Index to provide a mathematically interpretable approach for analysing academic progression in outcome-based education systems.

4. Data Collection and Educational Implementation

4.1 Dataset Description

The proposed reliability-based academic progression model is evaluated using the publicly available Student Performance Dataset from Kaggle, based on the UCI Machine Learning Repository. The dataset contains academic and behavioural information collected through school reports and student questionnaires.

The file `student-mat.csv` is used in this study. It contains records for 395 students enrolled in mathematics courses, with 33 attributes covering demographic, social, and academic characteristics. The variables selected for the reliability and degradation analysis are presented in Table 1.

Table 1. Selected variables used in the study

Variable	Description
G_1	First assessment grade
G_2	Second assessment grade
G_3	Final assessment grade
`studytime`	Weekly study time category
`failures`	Number of previous course failures
`absences`	Number of school absences
`schoolsup`	Additional educational support
`famsup`	Family educational support
`higher`	Intention to pursue higher education
`internet`	Internet access at home

The grades G_1, G_2, and G_3 are measured on a scale from 0 to 20 and represent the first, second, and final assessment stages, respectively.

4.2 Data Preprocessing

Before conducting the reliability analysis, the dataset is preprocessed using the following steps:

1. Identification and removal of duplicate records.
2. Verification of missing values and invalid observations.

3. Extraction of the variables required for the analysis.

4. Normalization of the assessment grades using

$$P_i = \frac{G_i}{20}, \qquad i = 1, 2, 3. \tag{8}$$

5. Classification of students into learner categories based on the final normalized score P_3.

Normalization transforms the grade values to the interval $[0, 1]$, facilitating consistent computation of the reliability measures. Duplicate records were checked to ensure that each student contributed only one record to the analysis. The selected variables were also examined for missing and invalid observations. No missing values were identified in the variables used for the reliability calculations; therefore, no imputation was required.

4.3 Experimental Framework

The overall workflow of the proposed reliability-based academic progression model is illustrated in Fig. 1.

The analysis begins with data acquisition, followed by preprocessing and normalization, as illustrated in Fig. 1. Reliability values are calculated at each assessment stage, and the relative decline measures are estimated between consecutive stages. The resulting degradation indicators are then interpreted to support educational monitoring and intervention planning.

The framework provides a systematic process for converting educational records into mathematically interpretable reliability measures. Unlike conventional descriptive approaches based primarily on aggregate grades or pass percentages, the proposed method examines changes in the proportion of students meeting a specified performance threshold across successive assessment stages.

4.4 Computation of Reliability Measures

After preprocessing, 395 student records were retained for the reliability analysis. Let $N = 395$ denote the total number of eligible student records. For each assessment stage S_i, let n_i denote the number of students whose normalized performance satisfies $P_i \geq 0.50$. The stage-wise reliability values are calculated as

$$R(S_i) = \frac{n_i}{N}, \qquad i = 1, 2, 3. \tag{9}$$

Here, n_i represents the number of students meeting the specified performance threshold at assessment stage S_i, while N denotes the total number of eligible student records. The threshold $P_i \geq 0.50$ corresponds to a grade of at least 10 on the original 0–20 scale.

The resulting reliability values quantify the proportion of students achieving satisfactory performance at each assessment stage. Comparing these values across S_1, S_2, and S_3 provides a basis for examining changes in satisfactory academic performance throughout the assessment sequence.

4.5 Hazard Rate Estimation

The discrete hazard measures are calculated from the reliability values at consecutive assessment stages using

$$H(S_i) = \frac{R(S_i) - R(S_{i+1})}{R(S_i)}, \qquad i = 1, 2. \tag{10}$$

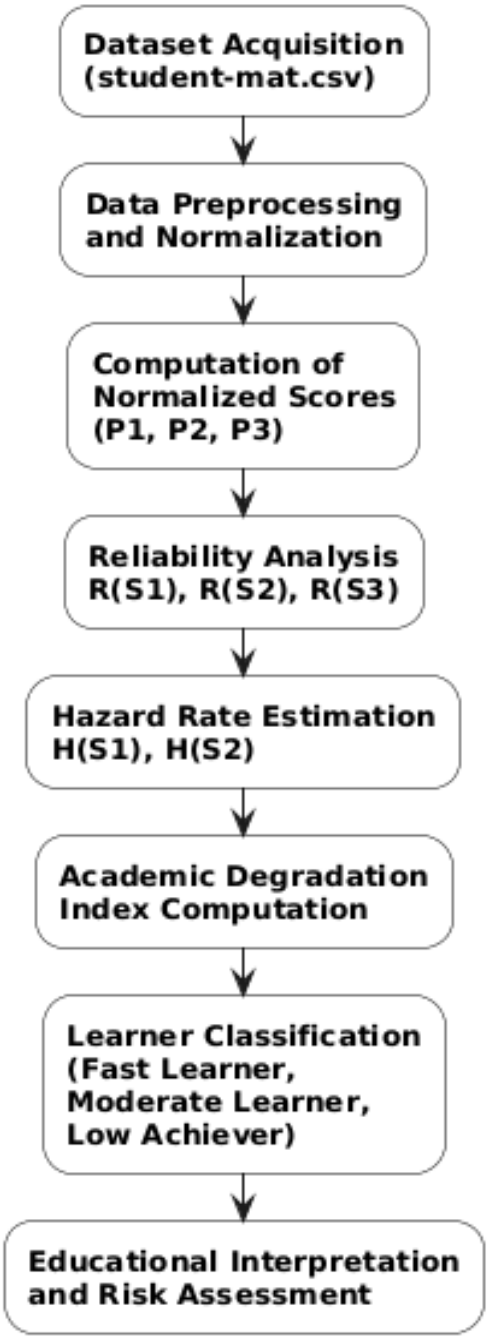


Fig. 1. Experimental framework of the proposed reliability-based academic progression model.

For the transition from S_1 to S_2, the hazard measure is given by

$$H(S_1) = \frac{R(S_1) - R(S_2)}{R(S_1)}. \tag{11}$$

Similarly, for the transition from S_2 to S_3, it is given by

$$H(S_2) = \frac{R(S_2) - R(S_3)}{R(S_2)}. \tag{12}$$

These measures quantify the relative reduction in the proportion of students meeting the specified performance threshold between consecutive assessment stages. The values reported in Table 7 are

calculated directly from the corresponding reliability values in Table 6. The measure is defined when the reliability value at the preceding stage is greater than zero.

4.6 Academic Risk Categories

To facilitate educational interpretation, students are classified into three academic risk categories according to their final assessment grade, as presented in Table 2.

Table 2. Academic risk categorization

Category	**Final Score Range**	**Risk Level**
Fast Learner	$G_3 \geq 15$	Low Risk
Moderate Learner	$10 \leq G_3 < 15$	Moderate Risk
Low Achiever	$G_3 < 10$	High Risk

This classification supports comparisons of academic reliability and relative decline measures among students with different final achievement levels.

4.7 Evaluation Strategy

The proposed framework is evaluated using stage-wise reliability values, relative decline measures, the Academic Degradation Index, and group-wise performance comparisons. These components provide quantitative indicators for examining satisfactory academic performance and its variation across the three assessment stages.

4.8 Research Workflow

The research workflow consists of the following steps:

1. Load the `student-mat.csv` dataset.
2. Select the relevant academic and behavioural attributes.
3. Normalize the grade variables.
4. Compute reliability values for each assessment stage.
5. Estimate relative decline measures and the Academic Degradation Index.
6. Classify students into learner categories.
7. Perform comparative academic progression analysis.
8. Interpret the results in the context of outcome-based education.

4.9 Educational Implementation and Data Collection

The proposed framework is intended to support the monitoring of student progression across three assessment stages within an outcome-based education (OBE) environment. The illustrative sixteen-week implementation schedule is presented in Table 3.

The proposed responsibilities of the participants are summarized in Table 4.

Table 3. Implementation schedule

Weeks	Activity	Duration
1–2	OBE orientation and learning outcome briefing	2 weeks
3–6	Lectures, tutorials, and assignments	4 weeks
7	First assessment (G_1)	1 week
8–11	Mentoring and remedial support	4 weeks
12	Second assessment (G_2)	1 week
13–15	Revision, project guidance, and performance monitoring	3 weeks
16	Final assessment (G_3)	1 week

Table 4. Roles and responsibilities

Participant	Responsibilities
Students	Attend classes, complete assignments, participate in mentoring sessions, and respond to questionnaires
Instructors	Conduct teaching activities, administer assessments, provide feedback, and monitor learning outcomes
Mentors	Identify academically at-risk students, conduct counselling sessions, and recommend remedial actions

Table 5. Data collection instruments

Instrument	Developer	Items	Administration Time
OBE Learning Outcome Questionnaire	Researcher-developed	15	Week 2
Mentoring Effectiveness Questionnaire	Researcher-developed	10	Week 11
Student Satisfaction Questionnaire	Researcher-developed	12	Week 16

The proposed educational implementation also considers three questionnaire instruments, summarized in Table 5.

The questionnaire instruments describe a possible educational implementation of the framework and may support future institutional applications. However, the quantitative reliability and relative decline analyses presented in this study are based on the publicly available student performance dataset described above. The questionnaire results are not treated as empirical findings unless the instruments were actually administered and the responses were collected and analysed. If administered, the questionnaires use a five-point Likert scale ranging from Strongly Disagree (1) to Strongly Agree (5). Participation should be voluntary and anonymous, with only aggregated responses used for analysis.

5. Results and Discussion

This section presents the reliability-based analysis of student academic progression within the proposed framework at three assessment stages, namely G_1, G_2, and G_3. Stage-wise reliability and relative decline measures are used to examine the proportion of students maintaining satisfactory academic performance and to identify assessment transitions associated with a reduction in satisfactory performance.

5.1 Reliability and Hazard-Rate Analysis

The proportion of students achieving satisfactory performance, defined by $P_i \geq 0.50$, at each assessment stage is presented in Table 6. For each stage, the number of students satisfying this criterion is denoted by n_i. The stage-wise reliability is calculated as

$$R(S_i) = \frac{n_i}{N}, \qquad i = 1, 2, 3, \tag{13}$$

where $N = 395$ is the total number of eligible student records.

Table 6. Reliability values across assessment stages

Stage	Assessment	Reliability
S_1	G_1	0.6405
S_2	G_2	0.5696
S_3	G_3	0.5595

Approximately 64.05% of students met the specified performance threshold at the first assessment stage. This proportion decreased to 56.96% at the second stage and 55.95% at the final stage. These results indicate a decline in the proportion of students meeting the performance criterion across the three assessments.

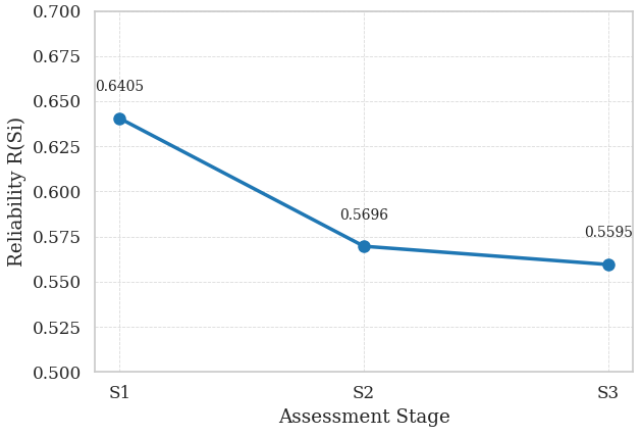


Fig. 2. Reliability progression across assessment stages.

As illustrated in Fig. 2, the reported reliability values decrease across the assessment stages. The larger reduction between S_1 and S_2 indicates a more pronounced decrease in the proportion of students meeting the threshold during this transition. The smaller reduction between S_2 and S_3 indicates a comparatively smaller change in the proportion meeting the criterion during the final transition.

The relative decline measures between consecutive assessment stages are calculated using

$$H(S_i) = \frac{R(S_i) - R(S_{i+1})}{R(S_i)}, \qquad i = 1, 2. \tag{14}$$

The resulting values are presented in Table 7.

The relative decline measure for the transition from S_1 to S_2 is greater than that for the transition from S_2 to S_3. This indicates that the relative reduction in stage-wise reliability is more pronounced during the first transition.

Table 7. Relative decline measures between assessment stages	
Transition	**Relative Decline**
$S_1 \to S_2$	0.1107
$S_2 \to S_3$	0.0178

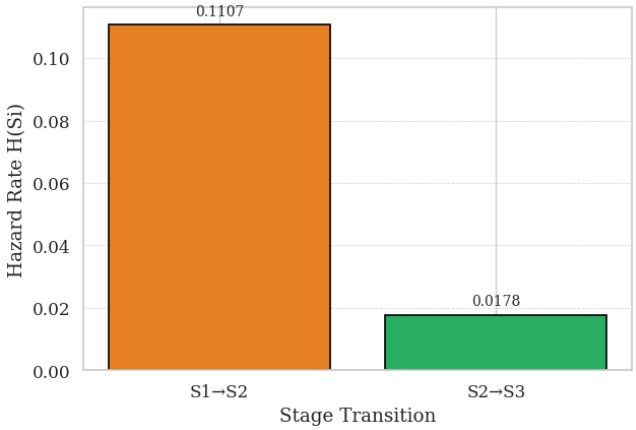


Fig. 3. Relative decline measures between consecutive assessment stages.

As shown in Fig. 3, the relative decline value of 0.1107 corresponds to an 11.07% reduction relative to the reliability at S_1. The value of 0.0178 corresponds to a 1.78% reduction relative to the reliability at S_2. These values describe changes in the proportions meeting the threshold; they should not be interpreted as directly observed individual transition probabilities because the stage-wise reliability values alone do not identify which individual students moved from satisfactory to unsatisfactory performance.

5.2 Learner Group Distribution and Academic Degradation

Students are classified into three learner groups according to their final assessment grade G_3. Students with $G_3 \geq 15$ are classified as fast learners, those with $10 \leq G_3 < 15$ as moderate learners, and those with $G_3 < 10$ as low achievers. These operational categories correspond to normalized final scores of at least 0.75, from 0.50 to below 0.75, and below 0.50, respectively. The classification is summarized in Table 8.

Table 8. Learner-group classification		
Group	**Condition**	**Risk Interpretation**
Fast Learners	$G_3 \geq 15$	Low Risk
Moderate Learners	$10 \leq G_3 < 15$	Moderate Risk
Low Achievers	$G_3 < 10$	High Risk

The distribution of students across the three learner categories is illustrated in Fig. 4.

The learner-group distribution provides a descriptive overview of final academic achievement. Students in the fast learner group meet the higher final-grade criterion, whereas students in the low-achiever group fall below the specified pass-level threshold. The latter group may warrant further investigation to identify potential needs for mentoring, remedial teaching, or other academic support. The actual proportion of students in each group should be interpreted from the observed distribution shown in Fig. 4.

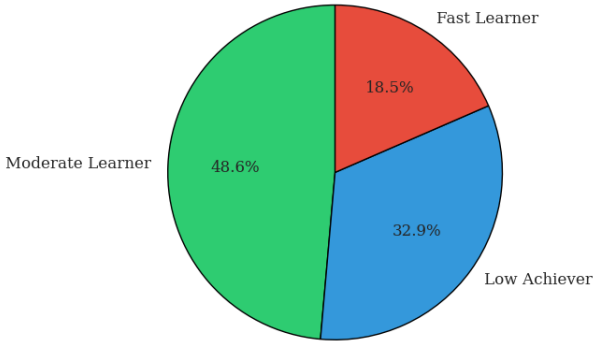


Fig. 4. Distribution of learner categories.

Academic performance across the three assessment stages is further examined using the mean normalized scores illustrated in Fig. 5.

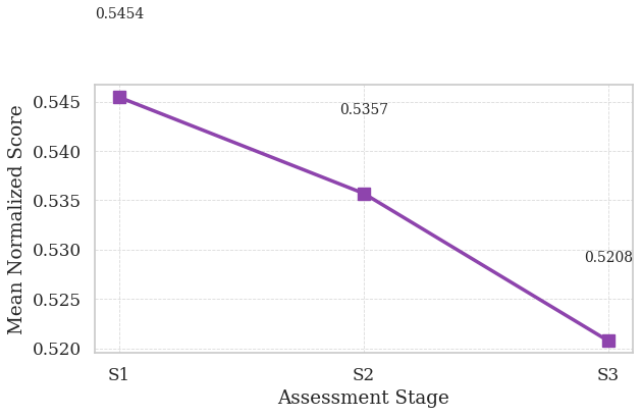


Fig. 5. Mean normalized academic performance across assessment stages.

Figure 5 presents the average normalized performance of students at each assessment stage. Any observed downward trend indicates a reduction in mean performance across the assessments, while an upward trend indicates improvement. This analysis complements the reliability measures because the mean score summarizes overall achievement, whereas stage-wise reliability measures the proportion of students meeting the specified performance threshold.

Taken together, the reliability results in Table 6 and Fig. 2, the relative decline measures in Table 7 and Fig. 3, the learner-group classification in Table 8 and Fig. 4, and the mean performance analysis in Fig. 5 provide complementary perspectives on academic progression. These measures can support the identification of assessment stages requiring further investigation and provide quantitative information for educational monitoring within outcome-based education systems.

6. Conclusions

This study introduced a reliability- and hazard-rate-based degradation model for analysing academic advancement in OBE systems. The cumulative reliability analysis showed a decreasing share of students maintaining satisfactory performance across assessment stages, and the hazard-rate analysis identified the transition between the first and second assessments as the most vulnerable period.

The proposed framework's practical applicability is strengthened by the educational implementation plan, mentoring support, and structured data collection procedure. The findings demonstrate the potential of reliability-oriented modelling to provide a mathematically interpretable and educationally meaningful approach for identifying academically at-risk students, tracking progression stability, and informing evidence-based interventions in Outcome-Based Education.

In addition, future research may explore the application of recent artificial intelligence models to academic performance prediction and compare their results with those obtained using the proposed framework [22-25].

Acknowledgement

This research was not funded by any grant.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Zheng, X., Zhai, W., Wang, W., Yang, B., Li, W., Luo, R., Sun, H., et al. (2026). *Unmasking reasoning processes: A process-aware benchmark for evaluating structural mathematical reasoning in LLMs*. arXiv preprint arXiv:2602.00564. https://doi.org/10.48550/arXiv.2602.00564.

[2] Cheng, X., Ji, S., Lee, Y., and Dong, H. (2026). Applications of Pharmacometrics in Antibody–Drug Conjugate Development. *Pharmaceutics*, 18(3), 354. https://doi.org/10.3390/pharmaceutics18030354.

[3] Clotfelter, C. T., Ladd, H. F., and Vigdor, J. L. (2011). Teacher mobility, school segregation, and pay-based policies to level the playing field. *Education Finance and Policy*, 6(3), 399–438. https://doi.org/10.1162/EDFP_a_00040.

[4] Meng, X., and Bojinov, I. (2025). *Time-Varying Causal Survival Learning*. arXiv preprint arXiv:2503.00730. https://doi.org/10.48550/arXiv.2503.00730.

[5] Acharya, S., and Das, S. (2024). Statistical Modeling for Predictive Healthcare Analytics. In *Revolutionizing Healthcare Treatment With Sensor Technology*, 265–293. IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3693-2762-3.ch017.

[6] Sorochynskyi, O., Guibert, Q., Planchet, F., and Schwarzinger, M. (2024). Estimating Disease-Free Life Expectancy Based on Clinical Data from the French Hospital Discharge Database. *Risks*, 12(6), 92. https://doi.org/10.3390/risks12060092.

[7] Charizanis, G., Mavridou, E., Vrochidou, E., Kalampokas, T., and Papakostas, G. A. (2025). Data-Driven Decision Support in SaaS Cloud-Based Service Models. *Applied Sciences*, 15(12), 6508. https://doi.org/10.3390/app15126508.

[8] Öhman, M., Finne, M., and Holmström, J. (2015). Measuring service outcomes for adaptive preventive maintenance. *International Journal of Production Economics*, 170, 457–467. https://doi.org/10.1016/j.ijpe.2015.06.020.

[9] Takenoshita, H., and Tagami, K. (2022). Gender Inequality in Access to Managerial Positions in Japan from a Cross-National Comparative Perspective: The Role of Labor Markets and Welfare States. *Japan Labor Issues*. Available at: https://www.jil.go.jp/english/jli/documents/2022/036-04.pdf.

[10] Zhang, Q., and Yi, H. (2024). How do university–industry alliances respond to the trust crisis in green technology innovation activities? *Nankai Business Review International*, 15(2), 269–299. https://doi.org/10.1108/NBRI-08-2022-0079.

[11] Neumark, D., Burn, I., and Button, P. (2019). Is it harder for older workers to find jobs? New and improved evidence from a field experiment. *Journal of Political Economy*, 127(2), 922–970. https://doi.org/10.1086/701029.

[12] Wan, K., Tanioka, K., and Shimokawa, T. (2025). *Causal rule ensemble approach for multi-arm data*. arXiv preprint arXiv:2504.17166. https://doi.org/10.48550/arXiv.2504.17166.

[13] Swiss Society of Neurosurgery (SSNS), and Swiss Society of Neuroradiology (SSNR). (2024). *Abstracts of the Joint Annual Meeting 2024 of the Swiss Society of Neurosurgery (SSNS) and the Swiss Society of Neuroradiology (SSNR) Together with the Association of Neurosurgical Nursing Staff Switzerland*. *Clinical and Translational Neuroscience*, 8(2), 24. https://doi.org/10.3390/ctn8020024.

[14] Berkok, U. G., Penney, J., and Kivinen, S. (2013). *Factor and Organizational Substitutions to Minimize Costs in the Navy*. Defence Research and Development Canada, Centre for Operational Research and Analysis, Contract Report DRDC-RDDC-2013-C10. Available at: https://publications.gc.ca/collections/collection_2016/rddc-drdc/D68-3-C10-2013-eng.pdf.

[15] Rasool, E., Kausar, N., Syed Ahmad, S. S., Aydin, N., and Olanrewaju, O. A. (2024). Multi-attribute decision-making based on Pythagorean fuzzy numbers and its application in hotel evaluations. *Decision Making: Applications in Management and Engineering*, 7(2), 559–571. https://doi.org/10.31181/dmame7220241204.

[16] Ejegwa, P. A., Onyeke, I. C., Kausar, N., and Kattel, P. (2023). A new partial correlation coefficient technique based on intuitionistic fuzzy information and its pattern recognition application. *International Journal of Intelligent Systems*, 2023, Article 5540085. https://doi.org/10.1155/2023/5540085.

[17] Shams, M., Kausar, N., Alayyash, K., Al-Shamiri, M. M., Arif, N., and Ismail, R. (2023). Semi-analytical scheme for solving intuitionistic fuzzy system of differential equations. *IEEE Access*, 11, 33205–33223. https://doi.org/10.1109/ACCESS.2023.3241482.

[18] Madasi, J. D., Khan, S., Kausar, N., Pamucar, D., Addis, G. M., Gulistan, M., and Kumar, R. (2022). A novel decision-making process in the environment of generalized version of fuzzy sets for the selection of energy source. *Advances in Mathematical Physics*, 2022, Article 7057639. https://doi.org/10.1155/2022/7057639.

[19] Agilan, K., Parthiban, V., Kausar, N., and Salman, M. A. (2026). Existence and uniqueness of solutions for fuzzy fractional integro-differential equations with boundary conditions. *Scientific Reports*, 16, Article 2082. https://doi.org/10.1038/s41598-025-31808-2.

[20] Cortez, P., & Silva, A. M. G. (2008). Student performance dataset. UCI Machine Learning Repository. https://www.kaggle.com/datasets/ishandutta/student-performance-data-set.

[21] Parthiban, Y. and Raghavendran, P. (2026). Application of Nanomaterial Selection and PINN-Based Prediction Using Neutrosophic Over Soft Complex Locally Closed Sets and Locally Continuous Functions. *Next Materials*, 13, Article 103274. https://doi.org/10.1016/j.nxmate.2026.103274.

[22] Demir, A. T. (2026). A Physics-Grounded Synthetic Data and Machine Learning Surrogate Framework for the Inverse Design of Functionally Graded Scaffolds in Osteochondral Tissue Engineering. *Knowledge and Decision Systems With Applications*, 2, 642–657. https://doi.org/10.59543/jccq8g73.

[23] Ma, Y., Zhang, X., and Meszaros, F. (2026). Event-Triggered Preference Adaptation with Large Language Model-Guided Conditional Reinforcement Learning for Pedestrian-Aware Traffic Signal Control. *Journal of Urban Intelligence and Smart Systems*, 1, 208–236. https://doi.org/10.59543/110jfp29.

[24] Wagdy Farag Ghaly, A. (2026). Governing the Intelligent University: An Empirical Analysis of Artificial Intelligence (AI) Integration Challenges in Smart City Knowledge Systems. *Journal of Urban Intelligence and Smart Systems*, 1, 180–198. https://doi.org/10.59543/43wm4v09.

[25] Farag, M. I. H. (2026). Artificial Intelligence and Strategic Decision-Making for Sustainable Governance. *Knowledge and Decision Systems With Applications*, 2, 447–468. https://doi.org/10.59543/aspp4925.

A Outcome-Based Learning Outcome Questionnaire

Students responded using the following scale:

1 = Strongly Disagree 2 = Disagree 3 = Neutral 4 = Agree 5 = Strongly Agree

1. I clearly understand the course learning outcomes.
2. The course activities are aligned with the expected program outcomes.
3. Continuous assessments help me monitor my learning progress.
4. The assessment criteria are clearly communicated.
5. Classroom activities support the development of problem-solving skills.
6. I can relate theoretical concepts to practical applications.
7. Feedback from instructors helps me improve my performance.
8. The course encourages independent learning.
9. Collaborative activities improve my understanding of the subject.
10. The learning resources provided are adequate for achieving the course outcomes.

B Mentoring Effectiveness Questionnaire

1. Mentoring sessions helped me identify my academic weaknesses.
2. The mentor provided useful guidance for improving my performance.
3. Remedial activities were helpful for understanding difficult topics.
4. Regular interaction with the mentor increased my confidence.
5. Mentoring support motivated me to participate actively in learning activities.