Article

Journal of International Students
Volume 16, Issue 18 (2026), pp. 279-300
ISSN: 2162-3104 (Print), 2166-3750 (Online)
jistudents.org
https://doi.org/10.32674/aqmtpx39

Forecasting Global International Student Mobility with Artificial Intelligence: A Confirmatory Multi-Horizon Study

Arjun Rajesh Panicker, Shenley Brook End School, Milton Keynes, United Kingdom
https://orcid.org/0009-0003-9169-8175

Nancy V, Easwari Engineering College, Ramapuram, Chennai, India

Swetha K B, M S Ramaiah University of Applied Sciences, Karnataka, India

Surya Kant Sharma, XLRI – Xavier School of Management, Jharkhand, India

Sudheer Choudari, Centurion University of Technology and Management, Andhra Pradesh, India

Abhishek K Singh, School of Media and Mass Communication, Graphic Era Hill University, Dehradun, India

Asesh Kumar Tripathy, Koneru Lakshmaiah Education Foundation, Andhra Pradesh, India

Somaditya Majumdar, Arka Jain University, Jamshedpur, Jharkhand, India.

Sheifali Gupta, Centre for Research Impact & Outcome, Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab, India

Divya Saleela, University of Southampton, Southampton, United Kingdom

* Corresponding author: Arjun Rajesh Panicker, panickera.19@sbe5d.com; Divya Saleela, d.saleela@soton.ac.uk

ABSTRACT: *Forecasting international student mobility supports institutional planning, education policy and equitable internationalization, but evidence is limited on whether artificial intelligence outperforms temporal benchmarks. This study develops GlobalStudent-AI, a horizon-dependent framework using 9,987 observations from 159 countries across one- to six-year horizons. Models were selected on 2021–2023 data and tested on an untouched 2024 confirmation year. Baselines were compared with direct and persistence-residual machine-learning models using error metrics, clustered bootstrap confidence intervals, adjusted significance tests, ablation, robustness checks and conformal uncertainty. Machine learning offered no reliable one-year advantage, but the benefits increased with increasing horizon. At six years, persistence-residual extra trees reduced the mean absolute error from 3.77 to 2.34, a 38.02% improvement. Overall, artificial intelligence was most effective for medium- and long-term forecasting.*

Keywords: international student mobility, artificial intelligence, international higher education, global student flows; multihorizon forecasting.

INTRODUCTION

International student mobility has become a significant component of contemporary higher education, shaping institutional internationalization, national skills strategies, knowledge exchange and the circulation of human capital. With respect to destination countries, international students contribute to sustainability, academic diversity, research capacity and regional economic activity. For sending countries, overseas study can expand access to specialized education and international networks, although it may also intensify inequalities in educational opportunity and raise concerns regarding talent retention (Chalil & Hoque, 2026; Gribble, 2008; Areesophonpichet et al., 2024). The scale and direction of international student flows are therefore relevant not only to universities but also to governments, funding bodies, immigration authorities and organizations responsible for housing, student support and workforce planning.

International student mobility is influenced by interconnected conditions in both origin and destination countries. Established push–pull perspectives suggest that limited domestic educational capacity, economic uncertainty and limited employment opportunities may encourage outward mobility, whereas educational quality, institutional reputation, economic stability, research opportunities and perceived career prospects may attract international students to particular destinations (Doda et al., 2024). Empirical research has further identified the importance of income differences, higher education participation, linguistic and geographic proximity, migration networks and policy conditions in shaping international student flows (Weber & Van Mol, 2023). More recent studies have extended this perspective by examining sustainability, climate readiness, destination attractiveness and the broader social and economic consequences of international mobility (Shields & Lu, 2024). These findings indicate that mobility patterns emerge from dynamic relationships among historical, educational, economic, demographic and policy-related factors rather than from a single determinant.

The capacity to forecast these patterns has practical importance. Universities must make decisions concerning recruitment, accommodation, teaching capacity and student-support provision before final enrollment numbers are known. Governments likewise require forward-looking evidence when developing visa policy, scholarship programs, internationalization strategies and regional skills plans. Inaccurate projections can result in underused capacity, pressure on accommodations and public services, or poorly targeted recruitment expenditure. Nevertheless, much of the international student mobility literature remains explanatory or descriptive, with greater attention given to the determinants, spatial structure and consequences of mobility than to the comparative accuracy of alternative forecasting approaches.

Existing forecasting research is also limited in terms of geographical and methodological scope. For example, research has examined student mobility using origin–destination matrices, network analysis and spatial models, providing valuable evidence about mobility routes and regional dependencies but not necessarily testing out-of-time forecasting performance (Ruiz et al., 2026). Recent European datasets have substantially improved the spatial representation of Erasmus + mobility, yet these resources primarily support descriptive and network-based analysis rather than global, country-level forecasting across multiple future horizons (Väisänen et al., 2025).

Machine learning offers a potential means of modeling the nonlinear and interacting relationships underlying student mobility. Explainable machine-learning methods have already been used to predict regional first-year student movements, with the results suggesting that ensemble models can achieve strong predictive performance, whereas SHAP values provide insight into nonlinear predictor contributions. However, strong model fit does not by itself demonstrate forecasting value. Student mobility indicators often exhibit substantial temporal persistence, meaning that the most recently observed value may already provide a difficult benchmark to outperform. A machine-learning model can achieve a high coefficient of determination while still producing larger forecasting errors than a simple persistence forecast. Consequently, the relevant question is not merely whether artificial intelligence can predict international student mobility but also whether it adds measurable predictive information beyond appropriate temporal baselines.

This distinction may be particularly important across different forecast horizons. Short-term mobility is likely to reflect enrollment continuity, existing institutional partnerships, visa arrangements and recruitment pipelines, all of which can create considerable inertia. At longer horizons, however, accumulated changes in economic conditions, digital access, tertiary education participation, research investment and demographic structure may become increasingly relevant. The predictive value of machine learning may therefore be horizon dependent: simple persistence may remain competitive for near-term forecasts, whereas models incorporating broader structural indicators may offer greater value over medium- and long-term periods. Existing international student mobility research has not adequately tested this proposition through a consistent global comparison across several forecast horizons.

Recent global research has examined approximately 1.15 million international student flows and demonstrated that national climate readiness and vulnerability are associated with the distribution of internationally mobile students. Such work significantly advances the understanding of global mobility and supports scenario-based projections. However, explanatory associations and future scenarios are methodologically distinct from out-of-time forecasting. A model intended for planning should be evaluated against strong naïve benchmarks, selected without reference to the final test period and accompanied by uncertainty estimates. Sensitivity to microstates, extreme observations, incomplete country histories and differences in predictor availability should also be examined. Without these safeguards, strong performance may reflect temporal leakage, influential cases or inadequate comparison standards.

This study addresses these gaps by developing GlobalStudent-AI, a confirmatory multihorizon framework for forecasting inbound international student shares. The framework integrates international student mobility records with historical, socioeconomic, demographic, technological and higher education indicators. Exact calendar-aligned targets are constructed for forecast horizons ranging from one to six years. Persistence, drift, historical-mean and pooled linear baselines are compared with direct and persistence-residual machine-learning models. Model development and selection are conducted using earlier temporal folds, while the final selected models are evaluated on an untouched 2024 confirmation year. Forecasting performance is assessed using absolute and squared error measures, persistence-relative skill, country-clustered bootstrap confidence intervals and paired tests corrected for multiple comparisons. Feature ablation, robustness analysis, SHAP-based interpretation and split-conformal prediction intervals are incorporated to evaluate model dependence, predictive contributions and forecast uncertainty. The validated framework is subsequently applied to generate uncertainty-aware country-level forecasts for 2027.

The study makes four principal contributions. First, it reframes international student mobility prediction as an incremental forecasting problem by testing whether machine learning improves upon persistence rather than evaluating artificial intelligence in isolation. Second, it introduces a direct, horizon-specific persistence-residual framework that allows the value of machine learning to vary across one- to six-year forecasting periods without recursive error propagation. Third, it combines temporally separated model selection with country-clustered uncertainty estimation, multiplicity-adjusted statistical testing, feature ablation and robustness analysis, thereby providing a more rigorous assessment than reliance on conventional accuracy measures alone does. Fourth, it develops calibrated country-level forecasts for 2027 and distinguishes expected changes from statistically secure mobility signals, supporting a more cautious and policy-relevant interpretation of future international student flows.

The remainder of this paper reviews the literature on international student mobility, its principal determinants and existing forecasting approaches. It then describes the data sources, predictor construction, forecasting models, temporal validation design, evaluation measures and uncertainty estimation procedures. The development, confirmation, ablation, robustness, explainability and 2027 forecasting results are subsequently presented. The findings are discussed in relation to previous research, with particular attention given to their implications for international higher education policy and institutional planning. Finally, the paper summarizes its main conclusions, acknowledges its limitations and identifies directions for future research.

LITERATURE REVIEW

International student mobility has been examined through several complementary perspectives, including studies of mobility determinants, spatial and network analyses, conventional forecasting, and machine-learning-based educational prediction. These strands have generated substantial knowledge about why students move, how mobility routes are structured, and which national or institutional characteristics influence destination choice. However, considerably less attention has been given to the comparative forecasting value of these factors, particularly across different time horizons and within a globally heterogeneous setting.

Determinants of International Student Mobility

Push–pull theory remains a central framework for understanding international student mobility. It distinguishes conditions that encourage students to leave their home country from attributes that attract them to a particular destination (Gbollie & Gong, 2020; Nikou et al., 2025; Mazzarol & Soutar, 2002). Push factors include limited program availability, restricted domestic educational capacity, political or economic uncertainty and weak employment prospects. Pull factors include educational quality, institutional reputation, research opportunities, linguistic accessibility, employment pathways, immigration prospects and established migrant or student networks.

Cross-national studies have shown that international student flows are influenced by economic disparities, geographical and linguistic distance, higher education capacity, institutional quality and established bilateral relationships (Lin and Gu, 2026). These factors do not operate independently. The attractiveness of a high-income destination, for example, may depend on tuition fees, immigration policy, labor market access, research infrastructure and the openness of its higher education system. More recent scholarship has broadened this perspective by considering sustainability, environmental resilience and the capacity of countries and institutions to respond to changing global conditions.

International student mobility is also increasingly affected by disruptions that are not fully captured by conventional push–pull variables. Changes in visa requirements, poststudy work opportunities, immigration policy, geopolitical relations and exchange rates can rapidly alter the relative attractiveness of study destinations. The COVID-19 pandemic has demonstrated how border closures, travel restrictions and uncertainty can interrupt established mobility patterns, while subsequent recovery has varied across countries and regions. Armed conflict, economic instability, climate-related disruption and the expansion of online, hybrid and transnational education may similarly influence whether students move internationally.

Recent international education research further shows that destination choice is sensitive to immigration pathways, affordability, career prospects, perceived safety and changing government policy. Recent studies have highlighted the importance of poststudy opportunities and institutional conditions, while evidence from changing visa and immigration policies demonstrates how policy shifts can disrupt established international enrollment pipelines (Hemati et al., 2026 ; Crumley-Effinger, 2023). These findings reinforce the need to interpret mobility as a policy-responsive process rather than as a stable continuation of historical trends.

Spatial and Network Perspectives

A second strand conceptualizes student mobility as a spatial and relational system. Studies in this tradition use origin–destination matrices, network measures, spatial statistics, and geographic visualization to identify mobility corridors, central destinations, and regional dependencies. Research on Erasmus mobility has shown that international exchanges form structured networks rather than isolated bilateral movements. Geographic proximity, institutional partnerships, historical relationships, and repeated exchange arrangements can reinforce established routes over time.

Genova et al. (2024) examined student mobility in southern Italy and identified patterns consistent with cumulative and chain migration. Geolocated Erasmus+ datasets have likewise enabled analysis of mobility at the country, regional, and institutional levels. Multilayer representations of Erasmus mobility further demonstrate that the student movement is embedded in evolving institutional and geographic networks.

These findings are directly relevant to forecasting because they imply that previous mobility is likely to be a strong predictor of future mobility (Ma and Zhang, 2022). Established corridors generate information networks, institutional familiarity, and recurring recruitment patterns, all of which contribute to temporal persistence. However, spatial and network studies generally prioritize structural description and explanation rather than out-of-time prediction. They identify where students move and how mobility systems are organized but rarely test whether more complex models outperform a simple continuation of the most recently observed pattern.

Time-Series Forecasting and Scenario-Based Projection

Conventional forecasting studies typically employ autoregressive integrated moving-average models, exponential smoothing, trend extrapolation, and cross-correlation analysis. Forecasting international student enrollment has also been examined using time series methods, which have demonstrated its potential value for institutional planning (Chen et al., 2019). Such methods are transparent and derive predictions primarily from the historical behavior of the outcome. The relationships among economic growth, higher education expansion, and international student mobility in Taiwan are analyzed using cross-correlation and ARIMA-based forecasting (Lo et al., 2022). Their findings indicated that economic and educational indicators evolve jointly with student mobility and that conventional time series models can support long-term planning. Recent research has also applied statistical, deep-learning, and ensemble forecasting models to African international student mobility, demonstrating the growing relevance of forward-looking analytical approaches for institutional planning and international education policy (Shikaa & Igiri, 2026).

The transferability of national forecasting models is nevertheless limited by variations in higher education systems, migration policy, demographic conditions, and economic structure. A model developed within one national context may not be generalizable to a global panel containing countries with markedly different mobility histories and institutional characteristics. Existing studies also tend to concentrate on a single forecast period, offering limited evidence on whether model performance changes across short-, medium-, and long-term horizons.

Scenario-based studies address future uncertainty differently. Rather than estimating one expected outcome, they examine mobility under alternative assumptions concerning policy, economic development, demographic change, or environmental conditions. For example, (Shields, 2019) investigated potential future mobility patterns under different combinations of national climate vulnerability and adaptation readiness. Such work is valuable for strategic analysis, but scenario construction is not equivalent to confirmatory forecasting. Scenarios illustrate plausible futures, whereas forecasting evaluation asks whether a model accurately predicts observations that were not used during its development.

This distinction is becoming increasingly important in the postpandemic international education environment. Border closures, changing immigration rules, uneven economic recovery and shifts in student preferences have shown that mobility patterns may depart abruptly from historical trajectories. The expansion of online, hybrid and transnational education may also alter the relationship between participation in international education and physical movement across borders. Forecasting systems should therefore be evaluated not only under relatively stable conditions but also for their ability to accommodate structural breaks and changing policy environments.

The broader forecasting literature indicates that no model class consistently dominates across all datasets and horizons. Statistical models may perform well when temporal relationships are stable, whereas machine-learning or hybrid approaches may offer advantages when nonlinear relationships and multiple predictors are important. Consequently, model performance should be assessed against credible naïve and conventional benchmarks rather than inferred from goodness of fit alone.

Machine Learning and Explainable Educational Prediction

Machine learning has become increasingly prominent in educational research, particularly in studies of academic performance, dropout, retention, employability, and student success (Hammoudi Halat et al., 2023 ; Boujmiraz et al., 2026). Random forests, gradient boosting, support vector machines, neural networks, and ensemble approaches can capture nonlinear associations and interactions that may not be represented adequately by conventional linear models.

Litmeyer and Hennemann (2024) applied explainable machine learning to first-year student mobility and reported that ensemble models could achieve competitive predictive performance, whereas SHAP values supported the interpretation of nonlinear predictor contributions. Their study demonstrated the value of combining prediction with interpretability and provided evidence that mobility decisions may be influenced by complex interactions among contextual variables.

Nevertheless, much educational machine-learning research remains focused on individual students, institutions, or domestic regional mobility rather than global international student flows. In addition, validation procedures do not always preserve temporal order, creating a risk that future information may indirectly influence model development. Many studies also compare several algorithms without determining whether any model improves upon a strong persistence benchmark. This is particularly important in international education, where visa regimes, recruitment networks, institutional partnerships, and enrollment pipelines may create considerable short-term stability.

Similarly, the wider forecasting literature cautions that algorithmic complexity does not guarantee improved out-of-sample performance. Tree-based ensembles can represent nonlinear relationships effectively, but their forecasting value must be demonstrated through temporally ordered validation, comparison with simple benchmarks, statistical testing, and explicit assessment of uncertainty. Explainability measures such as SHAP are also most appropriately interpreted as indicators of predictive contribution rather than evidence of causal influence.

Deep-learning forecasting approaches can capture complex sequential and nonlinear patterns, but their effectiveness depends on the data volume, temporal density, architecture, and validation design. Because the present country-year panel is relatively short, irregular, heterogeneous, and affected by missing data, extra trees provide a suitable balance between nonlinear modeling capacity, robustness, and interpretability. The methodological novelty of GlobalStudent-AI lies primarily in its confirmatory temporal design, horizon-specific evaluation, persistence-relative modeling, robustness analysis, multiplicity-adjusted inference, and calibrated uncertainty (Kong et al., 2025).

Synthesis and Research Gap

Taken together, the existing research provides substantial insight into the determinants, spatial organization, and potential future development of international student mobility. Determinant-based studies identify relevant economic, educational, demographic, policy, and environmental conditions.

Spatial and network analyses reveal persistent mobility corridors and institutional relationships. Time series and scenario studies demonstrate the value of historical trends and alternative future assumptions, whereas machine learning studies show the potential of nonlinear and explainable prediction.

However, these studies remain only partially connected. Explanatory studies do not establish whether identified determinants improve future forecasts. Spatial and network studies rarely evaluate the predictive performance of multiple horizons. National time series studies offer limited evidence of global generalizability, and scenario-based projections are not direct tests of out-of-sample accuracy. Existing machine learning studies seldom combine strict temporal ordering, strong naïve benchmarks, horizon-specific evaluation, statistical inference, robustness analysis, and calibrated uncertainty.

The principal unresolved issue is therefore whether machine learning adds reliable predictive value beyond temporal persistence in global international student mobility forecasting and whether that value changes as the forecast horizon lengthens. Short-term mobility may be dominated by existing enrollment pipelines, visa arrangements, institutional partnerships, and recruitment practices. Over longer periods, however, changes in economic capacity, digital connectivity, higher education participation, research investment, and demographic structure may become increasingly informative.

The present study addresses this issue through a direct multihorizon forecasting framework that compares machine learning models with persistence, drift, historical-mean, and pooled linear benchmarks across one- to six-year horizons. Model selection is conducted using earlier temporal observations, while final performance is assessed on an untouched confirmation year. The analysis further incorporates country-clustered bootstrap confidence intervals, multiplicity-adjusted statistical tests, feature ablation, robustness analysis, SHAP-based interpretation, and conformally calibrated prediction intervals. In this way, the study evaluates not only whether artificial intelligence can forecast international student mobility but also when it contributes predictive information beyond established temporal patterns.

METHOD

This study adopted a global longitudinal forecasting design to test whether machine learning models add predictive value beyond conventional temporal benchmarks in forecasting inbound international student mobility. The analytical design was direct, multihorizons and confirmatory: each one- to six-year target was constructed independently using exact calendar alignment, candidate models were selected only from temporally earlier development years, and the final model for each horizon was assessed on an untouched 2024 confirmation set. The complete workflow is summarized in Figure 1.

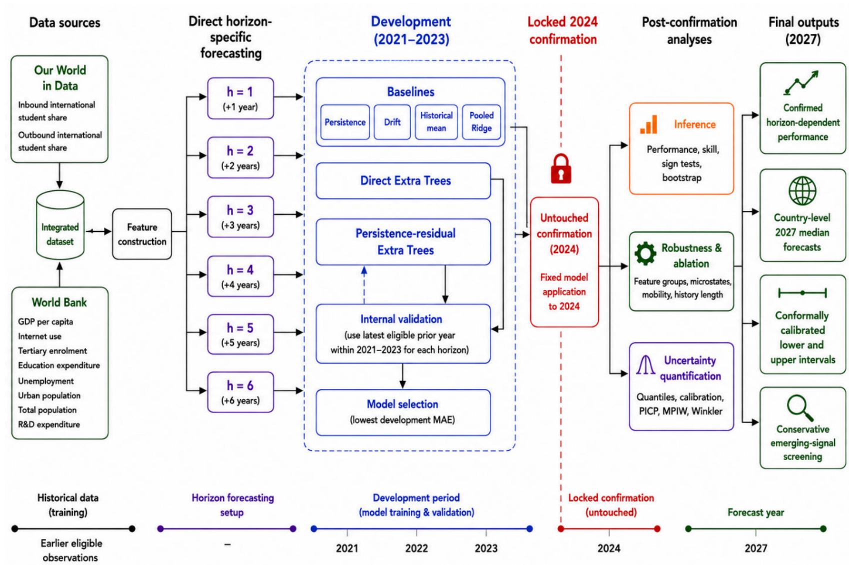

Figure 1: Analytical workflow of the global student-AI forecasting framework.

Data and Variables

International student mobility indicators were obtained from Our World in Data (UNESCO Institute for Statistics, 2025). The principal outcome was the inbound international student share, defined as the percentage of tertiary students in a country who originated from abroad. The corresponding outbound share was included as a predictor. The available inbound series covered 1998–2025, while the outbound series covered 1998–2023.

Country-level socioeconomic, demographic, technological and higher education indicators were retrieved through the World Bank application programming interface (World Bank, 2026). The variables included gross domestic product per capita, internet use, tertiary enrollment, government expenditure on education, unemployment, urban population, total population and research and development expenditure. Records were merged by ISO-3 country code and calendar year. After target construction and eligibility screening, the analytical panel contained 9,987 exact calendar-aligned forecasting observations representing 159 countries.

The predictors were organized into historical mobility, socioeconomic-demographic and higher education domains. Historical features included the current inbound and outbound shares, one- to three-year inbound lags, a one-year outbound lag, expanding historical mean and standard deviation, history length, recent change and acceleration. Gross domestic product per capita and population were transformed using $\log(1 + x)$. Values outside the valid 0–100% range were treated as missing. The median imputation and missingness indicators were estimated within each training fold so that information from the confirmation period could not influence preprocessing.

Table 1: Data domains, variables and analytical roles

Domain	Variables	Role in analysis	Source
Mobility outcome	Inbound international student share	Forecast target at horizons of one to six years	Our World in Data
Historical mobility	Current and lagged inbound and outbound shares; historical mean; variability; change; acceleration; history count	Captures pathway persistence and recent mobility dynamics	Our World in Data
Socioeconomic and demographic	GDP per capita; internet use; unemployment; urban population; total population	Represents economic capacity, connectivity and demographic structure	World Bank
Higher education	Tertiary enrollment; education expenditure; research and development expenditure	Represents education-system and research capacity	World Bank

Forecasting Framework

For country i, base year t and horizon h, the target was the observed inbound share in year $t+h$. Separate targets were generated for $h = 1, \ldots, 6$through exact country–year matching. This ensured that, for example, a three-year target always represented a genuine three-calendar-year interval rather than the third subsequently available observation. The change target was defined as follows:

$$\Delta y_{i,t+h} = y_{i,t+h} - y_{i,t}. \quad (1)$$

Three simple benchmarks were included. Persistence carried the most recent inbound share forward unchanged. Drift extrapolated the latest annual change across the relevant horizon, whereas the historical-mean forecast used the expanding country-specific mean. A pooled ridge model provided a regularized linear benchmark. These comparators were necessary because slowly changing education indicators can yield high model fit even when a learned model fails to outperform a simple temporal forecast.

Two extra tree formulations were evaluated. The direct model predicted the future level itself. The persistence-residual model predicted change relative to the current inbound share and then added a shrinkage-adjusted change estimate to persistence:

$$\hat{y}_{i,t+h} = y_{i,t} + \alpha_h f_h(\mathbf{x}_{i,t}). \quad (2)$$

The residual shrinkage coefficient was selected from 0.0 to 1.0 in increments of 0.1 using an internal validation year. The extra trees used 700 estimators, a minimum of three observations per terminal leaf and 90% of the available features at each split. All predictions were restricted to the valid 0–100% range.

Temporal Validation and Evaluation

For each horizon, the four most recent target years with at least 35 eligible countries were retained. The first three target years—2021, 2022 and 2023—formed development folds, while 2024 was reserved as the untouched confirmation year. Within each fold, all observations with earlier target years were used for training, and the latest year within that training period was reserved for internal validation. This nested temporal structure prevented future observations from influencing model selection or shrinkage tuning. The untouched 2024 confirmation set included 67, 65, 63, 64, 64 and 59 countries at the one- to six-year forecast horizons, respectively.

Models were ranked separately for each horizon using the mean development-period mean absolute error, with the root mean squared error used as a secondary criterion. The selected model was then frozen and evaluated once on 2024. The performance is reported as the mean absolute error, root mean squared error, median absolute error, coefficient of determination, symmetric mean absolute percentage error, weighted absolute percentage error and bias.

Incremental predictive value was measured using the MAE skill relative to persistence:

$$\mathrm{Skill}_h = 1 - \frac{\mathrm{MAE}_{\mathrm{model},h}}{\mathrm{MAE}_{\mathrm{persistence},h}}. \qquad (3)$$

Country-level absolute errors were compared using paired Wilcoxon signed-rank tests. The resulting p values were adjusted across the six horizons using the Holm procedure. Country-clustered bootstrap confidence intervals were generated from 1,500 resamples, with countries sampled with replacement. The analysis reported median skill, the 95% bootstrap interval and the estimated probability that skill was positive.

Explainability, Robustness and Forecast Uncertainty

Ablation analysis assessed whether forecasting performance depended on particular predictor domains. The selected extratrees formulation was re-estimated using the full feature set, historical features only, socioeconomic features only, higher education features only and a version excluding explicit missingness indicators. Each reduced model was assessed for the same confirmation year.

Robustness was examined by excluding countries with populations below 100,000, removing observations in the upper 1% of absolute mobility changes and restricting the sample to countries with at least eight historical inbound observations. These analyses tested whether the central results were driven by microstates, extreme changes or limited country histories.

SHAP values were computed for the one-year persistence–residual extra trees model to identify variables with the greatest predictive contribution. The values were interpreted as model-based contributions rather than causal effects.

Forecast uncertainty was assessed using gradient-boosting quantile regressors for the 5th, 50th and 95th percentiles of future change. Raw intervals were subsequently widened through split-conformal calibration using the latest preconfirmation year. Interval performance was evaluated using the prediction interval coverage probability, mean prediction interval width and Winkler score.

2027 Forecasting Application

After confirmation, the horizon-specific quantile models were re-estimated using all available labeled observations. Each country's forecast horizon was determined by the difference between 2027 and its latest available inbound-mobility year. Countries with implied horizons between one and six years were retained, and the analysis produced a median forecast with conformally calibrated lower and upper bounds.

Potentially emerging mobility signals were screened conservatively. A country was required to have a below-median current inbound share, an expected increase in the upper quartile, a positive conformal lower-bound change, an interval width below the 75th percentile, at least eight historical observations and a population of at least 100,000. These criteria distinguished positive point forecasts from increases supported across the complete calibrated interval.

All the stochastic procedures used a fixed seed of 42, and data acquisition, preprocessing, modeling and evaluation were implemented through reproducible Python pipelines. The study used aggregated, publicly available country-level data and involved no individual-level records, personal identifiers or direct participation by human subjects. Participant consent was therefore not applicable.

RESULTS

The confirmation results revealed a clear horizon-dependent pattern. Machine learning did not improve upon persistence at the one-year horizon, but its relative advantage increased progressively as the forecast interval lengthened. The greatest gain occurred at six years, when the persistence-residual extratrees model reduced the mean absolute error by 38.02% relative to that of persistence. Across all six horizons, the horizon-balanced mean improvement was 14.79%.

Horizon-Dependent Forecasting Performance

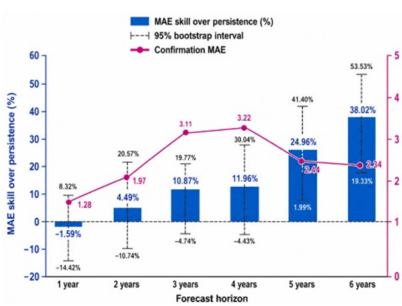

Figure 2 (a): Confirmation-period MAE skill over persistence, 95% bootstrap confidence intervals and model MAE across forecast horizons.

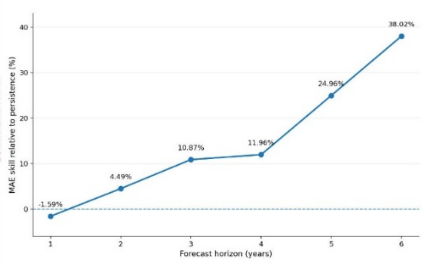

Figure 2 (b): Persistence-relative MAE skill across one- to six-year forecast horizons.

As shown in Figure 2(a), the forecasting performance improved progressively as the prediction horizon increased. The model performed slightly worse than persistence at the one-year horizon, with negative MAE skill, but achieved positive improvements from two years onward. The greatest gains occurred at the five- and six-year horizons, where the MAE skill reached 24.96% and 38.02%, respectively. The bootstrap confidence intervals indicate greater uncertainty at shorter horizons, whereas the longer-horizon improvements were more clearly separated from zero. The line plot also shows that the confirmation MAEs increase with the forecast horizon, reflecting the greater absolute difficulty of predicting outcomes further into the future.

The persistence-residual extra trees model was selected during development at every horizon. Its performance on the untouched 2024 confirmation set is summarized in Figure 2 (b). At one year, the model performed slightly worse than persistence did. Small positive gains appeared at two to four years, followed by substantially larger improvements at five and six years. This pattern indicates that recent mobility levels were difficult to outperform in the short term, whereas socioeconomic, demographic and historical information became more useful over longer intervals.

The statistical analysis supported the horizon-dependent pattern. The bootstrap confidence intervals included zero at horizons one to four, indicating uncertainty regarding improvement over persistence. At five years, the estimated skill was positive, and the bootstrap interval excluded zero; however, the paired Wilcoxon test was not statistically significant after the Holm correction for multiple comparisons. At six years, the bootstrap interval was entirely positive, and the Holm-adjusted Wilcoxon test confirmed a statistically significant improvement over persistence.

Table 2: Principal confirmation findings

Forecast horizon	Interpretation
1 year	No improvement over persistence
2–4 years	Modest but statistically uncertain gains
5 years	Substantial numerical improvement, but not significant after multiplicity correction
6 years	38.02% lower MAE than persistence; statistically significant and bootstrap-supported

The secondary error measures were consistent with this conclusion. The strongest long-horizon model maintained high explanatory accuracy and near-zero bias, indicating that its improvement was not produced by systematic overprediction or underprediction. Full MAE, RMSE, MedianAE, (R^2), sMAPE, WAPE and bias values are provided in the supplementary results.

Feature Contribution and Robustness

The feature-group analysis revealed that socioeconomic and demographic variables contained much of the information responsible for the improvement in the medium- and long-horizon forecasts. As shown in Figure 3, the socioeconomic-only model matched or exceeded the full model from the one- to five-year horizons and performed almost as strongly as the full specification at six years. Historical mobility remained useful, particularly at longer horizons, whereas the higher-education-only specification produced smaller and less consistent gains. Removing explicit missingness indicators had a negligible influence on performance, suggesting that the observed improvements were driven primarily by substantive predictor values rather than patterns of data absence.

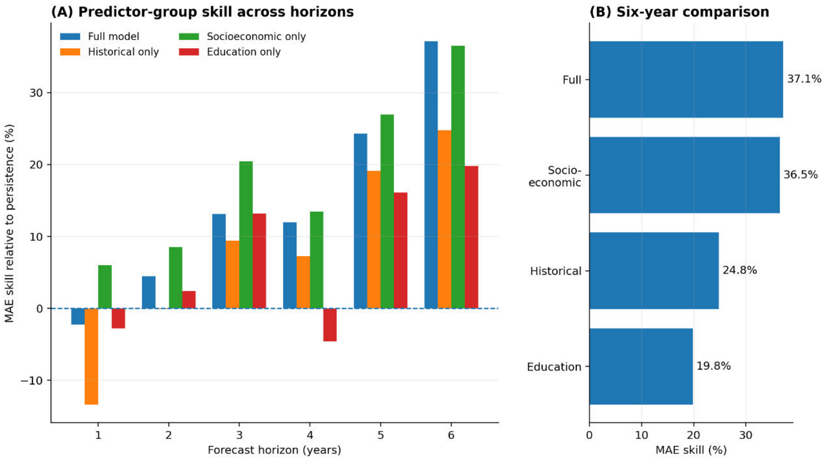

Figure 3: Forecasting skill by predictor group across the six horizons.

Because the ablation models were re-estimated separately for each predictor specification, the six-year full-model skill shown in Figure 3 differs slightly from the 38.02% skill obtained by the final frozen confirmation model.

The six-year result remained stable across all sample restrictions. After excluding microstates, the persistence-relative skill was 41.55%, 38.28% after the upper 1% of absolute mobility changes were removed and 37.12% when the analysis was restricted to countries with at least eight historical observations. The similarity of these estimates to the full-sample result of 38.02% indicates that the principal finding was not driven by very small countries, influential mobility changes or limited reporting histories.

SHAP values were examined for the one-year persistence–residual extratrees model to complement the predictor-group ablation analysis. The current inbound international student share made the greatest predictive contribution to the estimated change in mobility. Internet use, historical variability, GDP per capita, population and recent mobility change also influenced the model output, whereas the remaining variables had relatively small effects. The direction and magnitude of the SHAP values indicate how each feature contributed to individual predictions within the fitted model and should not be interpreted as evidence of causality.

Forecast Uncertainty

Raw quantile intervals undercovered the observed outcomes at all horizons. As shown in Figure 4(a), conformal calibration improved coverage substantially, particularly at the two longest horizons, where empirical coverage reached approximately the nominal 90% level or higher. The calibrated intervals were wider than the unadjusted intervals, reflecting the expected trade-off between interval precision and reliability.

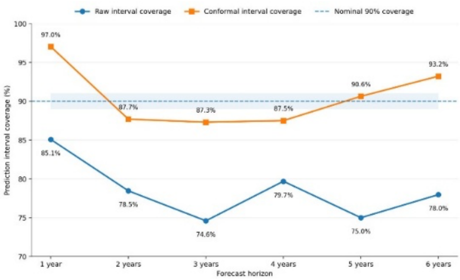

Figure 4 (a) : Raw and conformally calibrated prediction-interval coverage across forecast horizons.

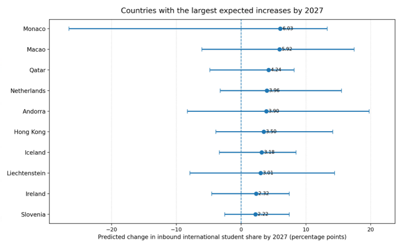

Figure 4 (b) : Countries with the largest expected increases in inbound international

At five and six years, the conformal coverage reached 90.63% and 93.22%, respectively. The coverage at horizons two to four remained slightly below the nominal target, while the one-year interval became conservative. The full interval width, pinball loss and Winkler score results are provided in the supplementary material.

Country-Level Forecasts for 2027

The final models generated median forecasts and conformally calibrated uncertainty intervals for 2027. The greatest expected median increases were observed for Monaco, Macao, Qatar, the Netherlands, Andorra, Hong Kong and Iceland, as shown in Figure 4 (b).

However, the conformal lower-bound change remained negative for every country, indicating that none of the expected increases remained positive across the complete uncertainty interval. Country-level forecasts should therefore be interpreted as emerging signals rather than as secure predictions of future growth. The complete set of country forecasts and calibrated intervals is provided in the supplementary material.

Overall, the results demonstrate that the forecasting value of machine learning strongly depends on the prediction horizon. Persistence remained difficult to outperform over one year, while gains at intermediate horizons were modest or statistically uncertain. In contrast, persistence-residual extra trees produced a substantial and statistically significant improvement at six years. This long-horizon advantage remained stable across robustness checks, was largely associated with socioeconomic and demographic information, and was accompanied by improved uncertainty calibration. Although the 2027 forecasts identified several potential growth destinations, no predicted increase remained positive across the full conformal interval.

DISCUSSION

This study examined whether machine learning improves upon temporal persistence when international student mobility is forecast across one- to six-year horizons. Its value depended strongly on the forecast length. Persistence remained difficult to outperform at one year, whereas the persistence-residual extra trees model became increasingly effective over longer horizons. At six years, it achieved a substantial and statistically significant reduction in forecast error, suggesting that recent mobility levels dominate short-term forecasts, while socioeconomic, demographic, and educational conditions become more informative over time.

The weak one-year performance is consistent with the institutional inertia of international higher education. Recruitment cycles, visa arrangements, bilateral agreements, program capacity, and institutional partnerships generally change gradually, meaning that the latest inbound share already provides a strong near-term forecast. Over five or six years, however, changes in economic capacity, digital connectivity, population structure, tertiary participation, and research investment can accumulate sufficiently to alter mobility patterns. The persistence-residual approach was therefore effective because it retained the latest value as an anchor while longer-term adjustments were estimated. The five-year improvement was substantial but did not remain significant after multiplicity correction, whereas the six-year gain was supported by both bootstrap inference and the Holm-adjusted Wilcoxon test.

Socioeconomic and demographic indicators accounted for much of the medium- and long-horizon performance. The socioeconomic-only model matched or approached the full model, indicating that economic capacity, internet access, population structure, urbanization, and labor market conditions provide useful forecasting information. Historical mobility remained relevant, but the strongest forecasts combined persistence with changing national conditions rather than simply reproducing recent trends. The higher education variables contributed less independently, possibly because they overlapped with broader measures of national capacity. The negligible effect of removing missingness indicators also suggests that performance was driven by substantive predictor values rather than patterns of data availability.

The six-year results remained robust after the exclusion of microstates, extreme changes, and countries with limited histories. However, generalizability remains limited because the confirmation was based on one untouched year. Omitted factors such as tuition fees, scholarships, visa rules, poststudy work rights, exchange rates, geopolitical events, and bilateral relationships may create structural breaks before their effects appear in historical data. The forecasts should therefore be interpreted as conditional planning estimates rather than policy-invariant predictions.

Conformal calibration improved interval coverage, although wider intervals reflected the trade-off between precision and reliability. Several countries showed positive median changes for 2027, but all the lower bounds remained negative. These forecasts should therefore be treated as emerging signals rather than as secure predictions of growth.

The findings extend international student mobility research by distinguishing explanatory relevance from forecasting usefulness. Economic, educational, and demographic variables may be associated with mobility without necessarily improving prediction. They added clear value mainly at medium and long horizons. The results also reinforce the persistence of established mobility corridors, institutional relationships, and recruitment networks, showing why high predictive accuracy alone is insufficient unless a model outperforms an appropriate temporal benchmark. The ablation and SHAP analyses aid interpretation, but they indicate a predictive contribution rather than causality.

For short-term planning, persistence may offer a sufficiently accurate and transparent basis for recruitment targets, admissions capacity, and student-service provision, supplemented by current application, visa, and recruitment information. For longer-term planning, machine-learning forecasts may support strategic enrollment management, program development, scholarship allocation, staffing, accommodation, student support, immigration-system capacity, and institutional partnerships. Universities may use horizon-specific forecasts to assess the geographical balance of recruitment portfolios and identify countries requiring closer monitoring. Governments and international education organisations may use them to inform internationalization strategies, scholarships, regional skills planning, and poststudy employment capacity. However, forecasts should complement professional judgment, local market knowledge, partner engagement, and policy analysis.

Forecasts should be presented as ranges rather than single values. Because all 2027 intervals cross zero, the identified countries should be monitored rather than treated as guaranteed growth markets. The remaining limitations are the use of inbound shares rather than absolute or bilateral flows and the absence of explicit spatial dependence, bilateral networks, and abrupt structural-break modeling.

CONCLUSION

This study developed GlobalStudent-AI to forecast inbound international student mobility across 159 countries and horizons of one to six years. The findings show that machine learning provides little advantage over persistence in the short term but becomes increasingly valuable over longer horizons. At six years, the persistence-residual extra trees model reduced the mean absolute error by 38.02% and remained statistically significant after multiplicity correction. Socioeconomic and demographic variables contributed strongly to this improvement, whereas conformal calibration produced more reliable uncertainty intervals. Although several countries showed positive expected changes for 2027, none remained positive across the complete calibrated interval.

For practical application, universities should use transparent persistence-based forecasts for short-term enrollment and capacity planning while combining longer-horizon machine-learning forecasts with institutional knowledge, policy monitoring and scenario analysis. Governments and international education organisations may use such forecasts to support strategic enrollment management, scholarship allocation, visa-processing capacity, accommodation, student services and long-term internationalization planning. However, because all 2027 conformal intervals cross zero, the country-level projections should be treated as signals for monitoring rather than firm evidence for recruitment or resource-allocation decisions.

Future research should extend confirmation across multiple future years, incorporate bilateral origin–destination flows and absolute student numbers, and examine additional predictors such as tuition fees, visa policy, poststudy work opportunities, institutional rankings and scholarships. Further work could also compare graph-based, hierarchical and dynamic forecasting models and investigate adaptive conformal methods to improve interval efficiency under temporal and regional distribution shifts.

Acknowledgment

In the preparation of this manuscript, we utilized artificial intelligence (AI) tools for content creation with the following capacity:

▢ *Some sections, with minimal or no editing*

This article uses artificial intelligence (AI) tools solely to support language editing, grammatical correction, and improvements in clarity and readability. AI tools were not used to generate the study data, conduct the analysis, interpret the results, or make scholarly conclusions. All AI-assisted content was thoroughly reviewed, verified, and edited by the authors, who take full responsibility for the accuracy, integrity, and final content of the manuscript.

REFERENCES

Areesophonpichet, S., Bhula-Or, R., Malaiwong, W., Phadung, S., & Thanitbenjasith, P. (2024). Thai Higher Education Institutions: Roles and Challenges in Attracting International Talent to Accelerate Thai Competitiveness in the Main Economy and Industry. *International Journal of Learning, Teaching and Educational Research*, *23*(2), 145–164. https://doi.org/10.26803/ijlter.23.2.7

Boujmiraz, S., Darhmaoui, H., & Drissi El Maliani, A. (2026). Predicting student performance: A comprehensive review of machine learning, deep learning, and explainable AI approaches. *Computers and Education: Artificial Intelligence*, *10*, 100548. https://doi.org/10.1016/j.caeai.2026.100548

Chalil, K., & Hoque, M. A. (2026). Internationalization of Higher Education: Concerns of Financing, Affordability and Equity. In P. K. Misra & N. Moitra (Eds.), *India's Higher Education in the Global Context* (pp. 251–265). Springer Nature Singapore. https://doi.org/10.1007/978-981-95-2367-2_14

Chen, Y., Li, R., & Hagedorn, L. S. (2019). Undergraduate International Student Enrollment Forecasting Model: An Application of Time Series Analysis. *Journal of International Students*, *9*(1), 242–261. https://doi.org/10.32674/jis.v9i1.266

Crumley-Effinger, M. (2023). ISM Policy Pervasion: Visas, Study Permits, and the International Student Experience. *Journal of International Students*, *14*(1). https://doi.org/10.32674/jis.v14i1.5347

Doda, S., Hysa, A., & Liça, D. (2024). International Student Mobility: Pushing-Pulling Factors for International Students. *Journal of Educational and Social Research*, *14*(6), 402. https://doi.org/10.36941/jesr-2024-0182

Gbollie, C., & Gong, S. (2020). Emerging destination mobility: Exploring African and Asian international students' push-pull factors and motivations to study in China. *International Journal of Educational Management*, *34*(1), 18–34. https://doi.org/10.1108/IJEM-02-2019-0041

Genova, V. G., Ruiu, G., Attanasio, M., Ermacora, M., & Breschi, M. (2024). Student mobility in Southern Italy: An empirical analysis of preferential patterns. *Genus*, *80*(1), 17. https://doi.org/10.1186/s41118-024-00225-0

Gribble, C. (2008). Policy options for managing international student migration: The sending country's perspective. *Journal of Higher Education Policy and Management*, *30*(1), 25–39. https://doi.org/10.1080/13600800701457830

Hammoudi Halat, D., Abdel-Salam, A.-S. G., Bensaid, A., Soltani, A., Alsarraj, L., Dalli, R., & Malki, A. (2023). Use of machine learning to assess factors affecting progression, retention, and graduation in first-year health professions students in Qatar: A longitudinal study. *BMC Medical Education*, *23*(1), 909. https://doi.org/10.1186/s12909-023-04887-w

Hemati, M., Pourasghari, H., Meshkani, Z., & Motamed-Jahromi, M. (2026). Factors influencing international students' choice of educational destination: A systematic review through the socioecological model lens. *Journal of International Students*, *16*(4), 33–54. https://doi.org/10.32674/ccv5sm82

Kong, X., Chen, Z., Liu, W., Ning, K., Zhang, L., Muhammad Marier, S., Liu, Y., Chen, Y., & Xia, F. (2025). Deep learning for time series forecasting: A survey. *International Journal of Machine Learning and Cybernetics*, *16*(7–8), 5079–5112. https://doi.org/10.1007/s13042-025-02560-w

Lin, Y., & Gu, H. (2026). Knowledge pursuit across distance: Revealing the multipolar trend of international student attraction with the A-index. *Geography and Sustainability*, *7*(2), 100426. https://doi.org/10.1016/j.geosus.2026.100426

Litmeyer, M.-L., & Hennemann, S. (2024). Forecasting first-year student mobility using explainable machine learning techniques. *Review of Regional Research*, *44*(1), 119–140. https://doi.org/10.1007/s10037-024-00207-x

Author bios

ARJUN RAJESH PANICKER is a Year 13 student at Shenley Brook End School, UK. His interests include artificial intelligence, machine learning, data science, educational analytics, and predictive modelling.
Email: panickera.19@sbe5d.com

NANCY V is an Assistant Professor at Easwari Engineering College, India. Her research interests include artificial intelligence, machine learning, deep learning, computer vision, natural language processing, cybersecurity, and intelligent healthcare.
Email: nancy.victorsagayee@gmail.com

SWETHA K B is an Associate Professor at M S Ramaiah University of Applied Sciences, India. Her research interests include artificial intelligence, machine learning, explainable AI, data analytics, and ubiquitous computing.
Email: drswethakb@gmail.com

SURYA KANT SHARMA is a Doctoral Scholar at XLRI – Xavier School of Management, India. His research interests include higher education management, public policy, leadership, and organisational structures.
Email: suryakantsharma@yahoo.com

SUDHEER CHOUDARI is an Assistant Professor at Centurion University of Technology and Management, India. His research interests include structural engineering, education, management, and artificial intelligence.
Email: sudheer.13031@gmail.com

ABHISHEK K. SINGH is a Professor at Graphic Era Hill University, India. His research interests include digital media, artificial intelligence, emerging technologies, and their applications in education.
Email: asingh8319@gmail.com

ASESH KUMAR TRIPATHY is an Associate Professor and Research Group Lead at K L University, India. His research interests include energy-aware protocols, AI-driven network modelling, and quantum-safe communication.
Email: asesh.tripathy@gmail.com

SOMADITYA MAJUMDAR is an Assistant Professor at Arka Jain University, India. His research interests include higher education, career psychology, labour laws, human dignity, and related interdisciplinary studies.
Email: somaditya.m@arkajainuniversity.ac.in

SHEIFALI GUPTA is a Professor at Chitkara University, India. Her research interests include image processing, pattern recognition, machine intelligence, biomedical imaging, agricultural image analysis, and deep learning.
Email: sheifali.gupta@chitkara.edu.in

DIVYA SALEELA is a Postdoctoral Research Fellow at the University of Southampton, UK. Her research interests include artificial intelligence, machine learning, medical imaging, predictive modelling, and healthcare applications.
Email: d.saleela@soton.ac.uk