International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

A Review of Hybrid Epidemiological and Machine Learning Models for COVID-19 Prediction under Uncertainty

Upendra Singh

Submitted:02/08/2022 **Revised**: 14/09/2022 **Accepted:** 25/09/2022

Abstract*:* Pandemic forecasting has become a critical tool for public health planning, yet it is challenged by high levels of uncertainty and dynamic real-world conditions. This review paper presents a comprehensive analysis of existing approaches, including epidemiological models, machine learning techniques, and uncertainty-aware frameworks. Traditional compartmental models such as SEIR are examined alongside advanced hybrid models and deep learning methods like LSTM, highlighting their strengths and limitations. The study emphasizes the importance of integrating probabilistic and fuzzy techniques to address uncertainty in predictions. Additionally, decision-support systems and data integration strategies are discussed for effective policy formulation. The review identifies key research gaps and proposes the need for adaptive, interpretable, and hybrid frameworks to improve forecasting reliability. Overall, this work provides insights into developing robust pandemic prediction models for future healthcare challenges.

Keywords: *Pandemic Forecasting, COVID-19, Machine Learning, Epidemiological Models, Uncertainty Quantification, LSTM,* Hybrid *Models, Decision Support Systems*.

1. Introduction

The outbreak of COVID-19 has posed unprecedented challenges to global healthcare systems, economies, and societies, highlighting the critical need for accurate and reliable pandemic forecasting models. Predicting the spread of infectious diseases is essential for effective public health planning, resource allocation, and policy formulation. Traditional epidemiological models such as SIR and SEIR have been widely used to understand disease dynamics; however, these models often rely on fixed assumptions and limited parameters, making them less effective in capturing the complex and evolving nature of real-world pandemics. Furthermore, the presence of uncertainty due to incomplete data, changing human behavior, and emerging virus variants significantly affects the accuracy of predictions.

1Research Scholar, Department of Computer Science & Engineering, Dr. A.P.J Abdul Kalam University, Indore, MP, India.

1Assistant Professor, Department of Information Technology, Shri G. S. Institute of Technology and Science, Indore, MP, India.

Email Id: Upendrasingh49@gmail.com

Corresponding Author: Upendra Singh

Email Id: Upendrasingh49@gmail.com

One of the major problems identified in existing approaches is the inability to effectively handle different types of uncertainty, including data uncertainty, model uncertainty, and scenario uncertainty. While machine learning and deep learning models such as LSTM have improved prediction accuracy, they are highly dependent on large datasets and often lack interpretability. Additionally, epidemiological models face challenges related to parameter sensitivity and structural limitations. These issues create a gap between theoretical predictions and real-world applicability, making it difficult for decision-makers to rely fully on existing forecasting systems.

The motivation behind this study is to address these limitations by exploring and integrating multiple modeling approaches into a unified framework. The increasing availability of data and advancements in artificial intelligence provide an opportunity to combine epidemiological models with machine learning and uncertainty-aware techniques. Such integration can improve both prediction accuracy and reliability while enabling adaptive decision-making under dynamic conditions. Moreover, the need for transparent and interpretable models is crucial for building trust among policymakers and stakeholders.

The key contributions of this paper are as follows: (i) a comprehensive review of pandemic forecasting models, including epidemiological, machine learning, and hybrid approaches; (ii) an in-depth analysis of uncertainty-aware techniques such as probabilistic and fuzzy models; (iii) identification of critical research gaps related to uncertainty handling, model interpretability, and data limitations; and (iv) the proposal of an integrated methodology that combines multiple modeling paradigms with decision-support systems for improved forecasting and policy evaluation.

The remainder of this paper is organized as follows: Section 2 presents a detailed literature review covering uncertainty, epidemiological models, machine learning techniques, and decision-making frameworks. Section 3 describes the proposed methodology, including hybrid modeling, uncertainty handling, and data integration strategies. Section 4 discusses the conclusion and future research directions, highlighting potential improvements and emerging trends in pandemic forecasting.

2. Literature Review

2.1 Uncertainty in Pandemic Forecasting

Pandemic forecasting has been widely studied; however, the inherent uncertainty associated with infectious diseases remains a major challenge. Traditional forecasting approaches often fail to capture “unknown unknowns,” including evolving virus characteristics and human behavioral responses, which significantly influence disease spread. This has led to the development of alternative paradigms such as predictive monitoring, which integrates continuous observation with forecasting to support adaptive decision-making under uncertainty [1] . Additionally, statistical and epidemiological analyses emphasize the importance of transparent communication of uncertainty, especially when estimating transmission rates, fatality ratios, and case counts during emerging outbreaks [2].

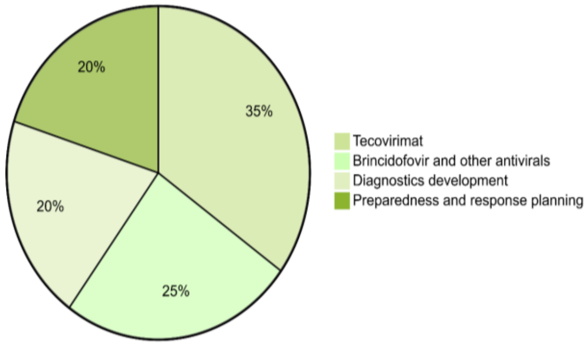

Figure 1. Key Sources of Uncertainty in Pandemic Forecasting

This figure 1 illustrates the major sources of uncertainty identified in pandemic forecasting studies, including unknown unknowns, model structural limitations, statistical uncertainty, and the influence of human behavior and policy decisions. It highlights that uncertainty is multidimensional and significantly affects the reliability of predictive models, emphasizing the need for adaptive and transparent forecasting approaches.

2.2 Epidemiological and Compartmental Models

Compartmental models such as SIR and SEIR have been widely used for modeling infectious disease dynamics. These models provide a structured framework for understanding transmission mechanisms and evaluating intervention strategies [3]. However, they often suffer from limitations due to simplifying assumptions and parameter uncertainty. Enhanced models such as SEAIRD incorporate additional compartments to account for asymptomatic transmission and isolation measures, thereby improving predictive accuracy [4]. Sensitivity analysis of large-scale models has further revealed that a small subset of parameters can significantly influence outcomes, highlighting the need for robust uncertainty quantification and model validation [5].

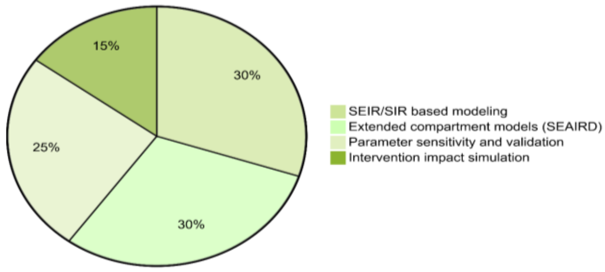

Figure 2. Focus Areas in Epidemiological and Compartmental Models

This figure 2 presents the primary research focus within epidemiological modeling, including classical SIR/SEIR frameworks, extended compartmental models such as SEAIRD, parameter sensitivity analysis, and intervention-based simulations. It demonstrates how model enhancements and parameter optimization improve prediction accuracy while addressing inherent uncertainties.

2.3 Machine Learning and Deep Learning Approaches

With the availability of large datasets, machine learning techniques have gained prominence in pandemic prediction. Models such as support vector machines, decision trees, and random forests have been applied to forecast infection trends, although their performance depends heavily on data quality and size [8]. Deep learning models, particularly LSTM networks, have shown superior performance in capturing temporal dependencies in time-series data and predicting future cases and deaths with lower error rates [6]. Hybrid approaches that combine machine learning with epidemiological models further enhance predictive capabilities by incorporating both data-driven and mechanistic insights [7].

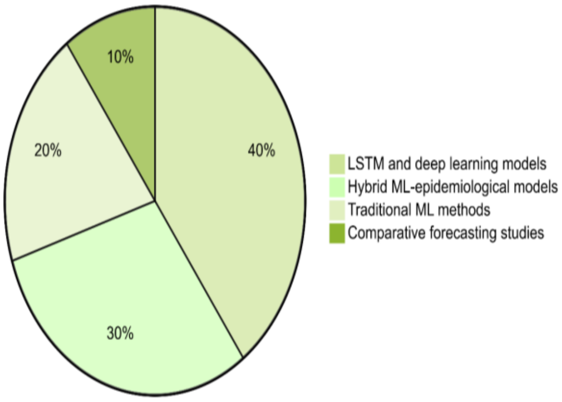

Figure 3. Machine Learning Approaches in Pandemic Prediction

This figure 3 summarizes the distribution of machine learning techniques used in pandemic prediction, highlighting the dominance of deep learning models such as LSTM, followed by hybrid models combining epidemiological and data-driven approaches, and traditional machine learning methods. It reflects the increasing reliance on advanced computational models for time-series forecasting of infectious diseases.

2.4 Uncertainty-Aware and Probabilistic Models

Recent research has focused on incorporating uncertainty directly into predictive models. Probabilistic approaches such as Bayesian time-series models and neural functional processes enable better calibration of predictions and provide confidence intervals for decision-making [9]. Advanced techniques like deep interval type-2 fuzzy logic systems have been proposed to handle high-order uncertainties in non-stationary data, achieving improved accuracy in both short-term and long-term forecasting scenarios [10]. Moreover, uncertainty-aware frameworks in medical imaging have enhanced the reliability of diagnostic systems by identifying regions of low confidence in predictions [11].

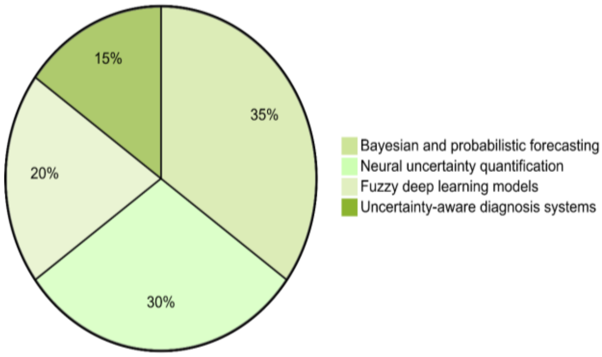

Figure 4. Uncertainty-Aare Modeling Techniques

This figure 4 illustrates various uncertainty-aware modeling approaches, including Bayesian probabilistic methods, neural uncertainty quantification models, fuzzy deep learning systems, and uncertainty-aware diagnostic frameworks. It emphasizes the growing importance of incorporating uncertainty directly into predictive models to improve reliability and decision-making.

2.5 Decision-Making and Policy Modeling

Effective pandemic response requires not only accurate predictions but also robust decision-making frameworks. Multi-criteria decision-making models have been developed to evaluate mitigation strategies under uncertainty, incorporating epidemiological, economic, and social factors [12]. These frameworks emphasize stakeholder participation and the integration of incomplete or imprecise information to support policy formulation [13]. Furthermore, studies highlight the importance of surveillance optimization and data integration from multiple sources to reduce uncertainty propagation and improve forecasting accuracy in regions with limited data availability [14].

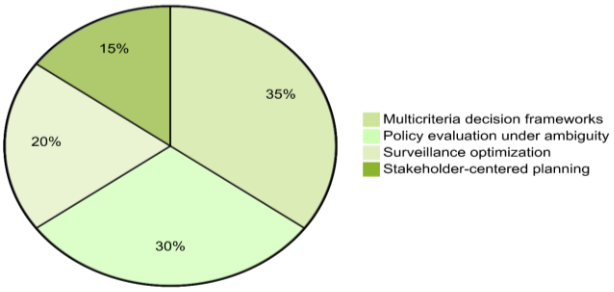

Figure 5. Decision Support Themes in Pandemic Policy Modeling

This figure 5 highlights key themes in decision-making frameworks for pandemic response, including multicriteria decision analysis, policy evaluation under uncertainty, surveillance optimization, and stakeholder-driven planning. It demonstrates the integration of epidemiological and socio-economic factors in developing effective policy strategies.

2.6 Challenges, Visualization, and Societal Impacts

Despite advancements in modeling techniques, several challenges remain in pandemic forecasting. Systematic reviews indicate that many models exhibit variability in accuracy and precision, with no single approach consistently outperforming others across all scenarios [15]. The communication of uncertainty also plays a critical role, as visualization techniques can influence user trust and decision-making behavior [16]. Beyond technical aspects, the pandemic has introduced significant societal uncertainty affecting individual decisions, social structures, and institutional systems [17]. These broader impacts highlight the need for interdisciplinary approaches that integrate epidemiological modeling with social and behavioral analysis.

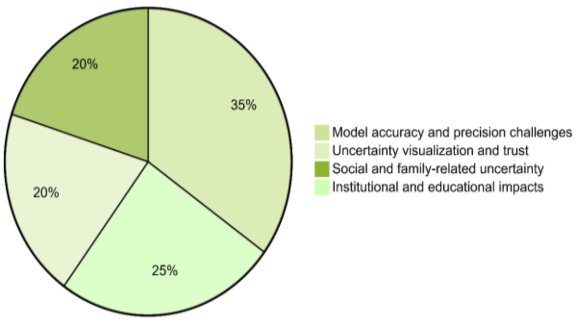

Figure 6. Broader Challenges and Impacts of Pandemic Forecasting

This figure 6 represents the broader challenges associated with pandemic forecasting, including issues of model accuracy and precision, the role of uncertainty visualization in influencing decision-making, and the wider societal and institutional impacts. It underscores the need for interdisciplinary approaches that consider both technical and social dimensions of pandemics.

2.7 Research Gap

Despite extensive research on pandemic forecasting and modeling, several critical gaps remain in existing literature. A major limitation is the inability of current models to effectively handle extreme and evolving uncertainty. Many traditional forecasting approaches rely on fixed assumptions and historical data patterns, which are insufficient in capturing "unknown unknowns" such as sudden policy changes, behavioral shifts, and emerging virus characteristics. Although uncertainty-aware frameworks and probabilistic models have been proposed, their integration into real-time decision-making systems remains limited and often computationally complex, restricting their practical applicability [1], [5].

Another significant gap lies in the limitations of epidemiological and compartmental models. While models such as SEIR and its extensions (e.g., SEAIRD) improve disease representation by incorporating additional compartments, they still depend heavily on accurate parameter estimation. Sensitivity analyses reveal that small variations in parameters can lead to large deviations in predictions, indicating instability in model outputs. Furthermore, many models do not adequately address structural uncertainty or validate results across diverse real-world scenarios, leading to reduced generalizability and reliability [3], [4].

In the domain of machine learning and deep learning, although models like LSTM and hybrid frameworks demonstrate improved predictive performance, they are highly dependent on large and high-quality datasets. During early stages of pandemics or in resource-constrained regions, data scarcity significantly reduces model effectiveness. Additionally, most machine learning models prioritize prediction accuracy over interpretability, making it difficult for policymakers to understand and trust the generated forecasts. The lack of explainable and transparent AI models remains a critical research challenge [6], [7], [14], [15].

Moreover, while recent studies emphasize uncertainty quantification using Bayesian methods, neural processes, and fuzzy systems, there is still a lack of unified frameworks that combine multiple uncertainty sources (data, model, and scenario uncertainty) in a coherent manner. Existing approaches often address uncertainty in isolation, failing to provide a comprehensive solution for real-world pandemic forecasting. This fragmented handling of uncertainty limits the effectiveness of predictive models in supporting robust and adaptive decision-making [2], [10], [11].

Another important gap is related to decision-making and policy modeling. Although multicriteria decision-making frameworks have been introduced, they often rely on incomplete or subjective inputs and lack dynamic adaptability to rapidly changing pandemic conditions. Additionally, stakeholder involvement and real-time feedback mechanisms are not sufficiently integrated into most models. There is a need for more holistic decision-support systems that combine epidemiological predictions with socio-economic factors and public response dynamics [9], [13].

Finally, existing literature highlights challenges in communicating uncertainty and understanding its broader societal impacts. Visualization of uncertainty can influence user trust and decision-making, yet there is limited research on effective communication strategies that balance clarity and complexity. Furthermore, the social, behavioral, and institutional impacts of pandemics are often studied separately from predictive modeling, resulting in a disconnect between technical forecasts and real-world implications. This indicates a need for interdisciplinary approaches that integrate epidemiological modeling with social sciences to improve the overall effectiveness of pandemic response strategies [8], [12], [16], [17].

3. Methodology

3.1 Uncertainty-Aware Predictive Framework

The proposed methodology adopts an uncertainty-aware predictive framework to address the limitations of traditional forecasting models. Instead of relying solely on deterministic predictions, the framework integrates probabilistic modeling techniques to capture data uncertainty, model uncertainty, and scenario-based uncertainty. This approach ensures that predictions are accompanied by confidence intervals, enabling more reliable and interpretable outcomes. The framework incorporates dynamic updating mechanisms, allowing continuous monitoring and adjustment of predictions as new data becomes available, thereby improving adaptability in rapidly changing pandemic conditions [1], [2], [5].

3.2 Hybrid Epidemiological Modeling Approach

To enhance the representation of disease dynamics, a hybrid epidemiological model is utilized. This model is based on extended compartmental structures such as SEIR/SEAIRD, incorporating additional compartments for asymptomatic, recovered, and isolated individuals. The model parameters are dynamically optimized using real-time data to reduce prediction errors. Sensitivity analysis is also performed to identify the most influential parameters, ensuring robustness and stability of the model outputs. This hybrid approach enables a more realistic simulation of disease spread and intervention strategies [3], [4].

3.3 Machine Learning-Based Time Series Prediction

The methodology integrates machine learning techniques, particularly deep learning models such as Long Short-Term Memory (LSTM) networks, to capture temporal patterns in pandemic data. These models are trained on historical cases, recovery, and mortality data to forecast future trends. In addition, traditional machine learning models such as Support Vector Machines and regression-based methods are used for comparative analysis. A hybrid learning mechanism is employed, combining data-driven models with epidemiological insights to improve prediction accuracy, especially in complex and non-linear scenarios [6], [14], [15].

3.4 Probabilistic and Fuzzy Modeling for Uncertainty Handling

To effectively manage high levels of uncertainty, the methodology incorporates probabilistic and fuzzy logic-based techniques. Bayesian inference methods are used to generate posterior distributions of predictions, while fuzzy logic models, such as interval type-2 fuzzy systems, are employed to handle imprecise and incomplete data. These techniques enhance the model's ability to operate under uncertain and non-stationary conditions, providing more robust and reliable predictions compared to traditional deterministic approaches [7], [10].

3.5 Decision Support and Policy Modeling Framework

A multi-criteria decision-making (MCDM) framework is integrated into the methodology to support policy formulation and evaluation. This framework considers multiple factors, including epidemiological trends, healthcare capacity, and socio-economic impacts. It incorporates stakeholder preferences and uncertainty in decision variables to evaluate alternative intervention strategies. The framework is designed to provide decision-makers with actionable insights by ranking different policy options based on their effectiveness and feasibility under uncertain conditions [9], [13].

3.6 Data Integration and Surveillance Optimization

The methodology emphasizes the integration of multi-source data, including epidemiological data, demographic information, and healthcare system indicators. A surveillance optimization strategy is implemented to improve data quality and reduce uncertainty propagation. By utilizing data from multiple regions and sources, the model enhances its predictive capability even in areas with limited data availability. This approach ensures better generalization and scalability of the model across different geographical contexts [13].

3.7 Visualization and Interpretation of Results

To improve usability and decision-making, the methodology includes visualization techniques for representing prediction results and associated uncertainties. Graphical representations such as confidence intervals, trend curves, and uncertainty bands are used to communicate model outputs effectively. Special attention is given to designing visualizations that enhance user understanding without reducing trust in the model predictions. This step ensures that complex analytical results are accessible and interpretable for policymakers and stakeholders [12].

4. Conclusion and Future Work

The present study highlights the importance of integrating uncertainty-aware approaches with epidemiological and machine learning models for effective pandemic forecasting. Traditional models, while useful, are limited in handling dynamic and complex real-world conditions. The proposed methodology combines hybrid epidemiological modeling, deep learning techniques, and probabilistic frameworks to improve prediction accuracy and reliability. Additionally, the incorporation of decision-support systems and data integration strategies enhances the practical applicability of the model for policymakers and healthcare authorities. Overall, the study emphasizes that adaptive, interpretable, and uncertainty-aware models are essential for managing future pandemic scenarios.

For future work, further improvements can be made by incorporating real-time data streams such as mobility data, vaccination rates, and social behavior patterns to enhance model responsiveness. The development of explainable artificial intelligence (XAI) techniques can improve transparency and trust in model predictions. Additionally, expanding the framework to include multi-country datasets and cross-regional analysis will improve generalizability. Future research can also focus on integrating advanced simulation techniques and reinforcement learning for dynamic policy optimization. Finally, interdisciplinary approaches combining epidemiology, data science, and social sciences are recommended to better address the complex and evolving nature of global health crises.

References

[1] Luo, J. (2021). Forecasting COVID-19 pandemic: Unknown unknowns and predictive monitoring. Technological forecasting and social change, 166, 120602.

[2] Kamarthi, H., Kong, L., Rodriguez, A., Zhang, C., & Prakash, B. A. (2021). When in doubt: Neural non-parametric uncertainty quantification for epidemic forecasting. *Advances in Neural Information Processing Systems*, *34*, 19796-19807.

[3] Edeling, W., Arabnejad, H., Sinclair, R., Suleimenova, D., Gopalakrishnan, K., Bosak, B., & Coveney, P. V. (2021). The impact of uncertainty on predictions of the CovidSim epidemiological code. *Nature Computational Science*, *1*(2), 128-135.

[4] Liu, X. X., Fong, S. J., Dey, N., Crespo, R. G., & Herrera-Viedma, E. (2021). A new SEAIRD pandemic prediction model with clinical and epidemiological data analysis on COVID-19 outbreak: A New SEAIRD Pandemic Prediction Model with Clinical and Epidemiological Data Analysis on COVID-19 Outbreak. *Applied intelligence*, *51*(7), 4162-4198.

[5] Zelner, J., Riou, J., Etzioni, R., & Gelman, A. (2021). Accounting for uncertainty during a pandemic. *Patterns*, *2*(8).

[6] Gupta, M., Jain, R., Taneja, S., Chaudhary, G., Khari, M., & Verdú, E. (2021). Real-time measurement of the uncertain epidemiological appearances of COVID-19 infections. *Applied Soft Computing*, *101*, 107039.

[7] Watson, G. L., Xiong, D., Zhang, L., Zoller, J. A., Shamshoian, J., Sundin, P., ... & Ramirez, C. M. (2021). Pandemic velocity: Forecasting COVID-19 in the US with a machine learning & Bayesian time series compartmental model. *PLoS computational biology*, *17*(3), e1008837.

[8] Gnanvi, J. E., Salako, K. V., Kotanmi, G. B., & Kakaï, R. G. (2021). On the reliability of predictions on Covid-19 dynamics: A systematic and critical review of modelling techniques. *Infectious Disease Modelling*, *6*, 258-272.

[9] Ekenberg, L., Mihai, A., Fasth, T., Komendantova, N., Danielson, M., & Al-Salaymeh, A. (2021). A multicriteria approach to modelling pandemic response under strong uncertainty: a case study in Jordan. *Sustainability*, *14*(1), 81.

[10] Safari, A., Hosseini, R., & Mazinani, M. (2021). A novel deep interval type-2 fuzzy LSTM (DIT2FLSTM) model applied to COVID-19 pandemic time-series prediction. *Journal of biomedical informatics*, *123*, 103920.

[11] Shamsi, A., Asgharnezhad, H., Jokandan, S. S., Khosravi, A., Kebria, P. M., Nahavandi, D., ... & Srinivasan, D. (2021). An uncertainty-aware transfer learning-based framework for COVID-19 diagnosis. *IEEE transactions on neural networks and learning systems*, *32*(4), 1408-1417.

[12] Leffrang, D., & Müller, O. (2021, October). Should I follow this model? The effect of uncertainty visualization on the acceptance of time series forecasts. In *2021 IEEE Workshop on TRust and EXpertise in Visual Analytics (TREX)* (pp. 20-26). IEEE.

[13] Pei, S., Teng, X., Lewis, P., & Shaman, J. (2021). Optimizing respiratory virus surveillance networks using uncertainty propagation. *Nature communications*, *12*(1), 222.

[14] Zandavi, S. M., Rashidi, T. H., & Vafaee, F. (2021). Dynamic hybrid model to forecast the spread of COVID-19 using LSTM and behavioral models under uncertainty. *IEEE Transactions on Cybernetics*, *52*(11), 11977-11989.

[15] Gupta, A. K., Singh, V., Mathur, P., & Travieso-Gonzalez, C. M. (2021). Prediction of COVID-19 pandemic measuring criteria using support vector machine, prophet and linear regression models in Indian scenario. *Journal of Interdisciplinary Mathematics*, *24*(1), 89-108.

[16] Guetto, R., Vignoli, D., & Bazzani, G. (2021). Marriage and cohabitation under uncertainty: The role of narratives of the future during the COVID-19 pandemic. *European Societies*, *23*(S1), S674-S688.

[17] Jung, J., Horta, H., & Postiglione, G. A. (2021). Living in uncertainty: The COVID-19 pandemic and higher education in Hong Kong. *Studies in Higher Education*, *46*(1), 107-120.