

Review

Not peer-reviewed version

Transcending Time and Space: A Survey Methods, Uncertainty, and Development in Human Migration Prediction

TongZheng Pu , Chongxin Huang , [Jingjing Yang](#) ^{*} , [Ming Huang](#) ^{*}

Posted Date: 6 April 2023

doi: 10.20944/preprints202304.0075.v1

Keywords: human migration; prediction; methods; artificial intelligence; data; uncertainty



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Review

Transcending Time and Space: A Survey Methods, Uncertainty, and Development in Human Migration Prediction

TongZheng Pu ¹, ChongXin Huang ², JingJing Yang ^{1,*} and Ming Huang ^{1,*}

¹ School of Information Science and Engineering, Yunnan University, Kunming 650500, China; putongzheng@mail.ynu.edu.cn;

² Faculty of Social and Historical Sciences, University College London, London, United Kingdom WC1E 6BT; chongxing.huang.21@ucl.ac.uk

* Correspondence: jingjingyang@ynu.edu.cn (J.J.Y.); huangming@ynu.edu.cn (M.H.)

Abstract: As a fundamental, overall, and strategic issue facing human society, human migration is a key factor affecting the development of countries and cities given constantly changing population numbers. The fuzziness of the spatiotemporal attributes of human migration limits the pool of open-source data for human migration prediction, leading to a relative lag in human migration prediction algorithm research. This study expands the definition of human migration research, reviews the progress of research into human migration prediction, and classifies and compares human migration algorithms based on open-source data. It also explores the critical uncertainty factors restricting the development of human migration prediction. Given the effect of human migration prediction, in combination with artificial intelligence and big data technology, the paper concludes with specific suggestions and countermeasures aimed at enhancing human migration prediction research results to serve economic and social development and national strategy.

Keywords: human migration; prediction; methods; artificial intelligence; data; uncertainty

1. Introduction

With the worldwide occurrence of rapid economic and social development, human migration and mobility between urban and rural areas, between cities, and between countries have become more convenient, and human migration (HM) has become a universal phenomenon. For instance, in China alone, as of November 2021, the inter-provincial mobile population was 124,837,153 and the intra-provincial mobile population was 250,979,606[1]. A report issued by the United Nations Population Division (United Nations, 2020) predicted that the number of global migrants in 2020 would be 281 million [2]. While the pace of HM slowed due to the COVID-19 pandemic, the gradual improvement of the pandemic and the adjustment of relevant policies have once again led to an increase in domestic and international HM [3]. Given the significant decrease in the global population growth rate and the gradual increase in the ageing population, HM has become a significant component of population growth in many countries and regions. Labor, health care, education and other such policies must fully consider HM development dynamics and trends to ensure that the policies they formulate are forward-looking, targeted, and effective. To accurately understand HM trends, extensive research has been conducted by governmental organizations, academics, and industry. HM forecasting has, therefore, become a vital research hotspot in the field of population studies.

International migration has received greater attention than internal migration due to its multifaceted effects and policy importance, leading to greater theoretical and empirical results in that field. Early studies of human migration prediction (HMP) mainly focused on the analysis of HM drivers and the establishment of a relationship equation between HM drivers and number of

migrants to predict future HM development trends [4]. Studies show that prediction models used in domestic and international HM forecasting are remarkably similar, differing only in data sources and policies; models from both types of HM can be used to support each other in the development of improved forecasting methods [5]. In the 1940s, the sociologist Zipf first proposed a gravitational model to predict HM [6]. In the late 1950s, the demographer Bogue proposed the famous "push-pull theory" to describe the motives of HM from a kinematic perspective [7]. Later, the American scholar Lee added to and improved the push-pull theory of HM based on Bogue's thesis [8]. In recent years, scholars have conducted empirical analyses of these models and have continuously optimized them and improved their accuracy based on applied examples [9,10]. With the development of information technology and statistics, many forecasting methods have emerged, and these mainly focus on econometrics, time series, Bayesian statistics, etc. In recent years, with the rapid development of artificial intelligence (AI) technology with big data and machine learning (ML) as the core, some scholars have tried to use ML technology for HMP[67–73]; however, the limitations include the limited amount of HM data, different standards, and difficulties in access. In addition, the uncertainty of HM drivers and the difficulty of their quantification have led to the slow development of HMP research [4].

Many scholars respectively conducted systematic reviews of HMP research, which provided a comprehensive overview of the current status of HMP research and a reference for further investigation [11–13]. However, they mainly focused on traditional HMP methods and ignored the application and development of AI technology in HMP research. In this regard, this paper tracks the development of the application of AI technology to HMP and provides a review of the traditional and AI technology-based HMP research to provide a scientific basis for future HMP research.

This paper first reviews the development of HM theory and summarizes and analyzes the classification of HMP. Next, focus is placed on traditional HM forecasting methods, and their successes and shortcomings are analyzed. Then, the application and development of AI technology-based HMP are reviewed, and the uncertainties that lead to the slow growth of HM are analyzed. Finally, conclusions are drawn by analyzing the current status of HMP research and pointing out the development direction.

2. Human Migration Theory and Research Progress

2.1. Human Migration Theoretical Review

The geographical or spatial flow of population between two regions usually involves the change of a permanent residence from the place of departure to the location of arrival, which is called permanent migration [14]. With the rapid development of transportation and information networks and the acceleration of globalization, people's travel frequency and time are increasing, the travel distance is growing, and permanent residence is decreasing [20]. In HM projection, it is common practice to ignore internal migration projection studies and concentrate on international migration, despite the latter comprising only a small proportion of the mobile population. Because of this, the general public's comprehension of the entire process of population movement is less thorough and accurate [15]. Despite significant differences in the scale of these two phenomena, studies have shown that the evolution of international migration can be explained to a large extent by internal migration, as there is significant complementarity between the two [16–18]. In addition, tourism migration as another form of population flow, overlaps with migration in many respects and has a causal relationship [19]. It is difficult to draw a line between these two phenomena [20]. It is therefore necessary to dismiss the generally limited definition of HM and adopt a systematic approach to properly understand the interconnections in HM and ultimately better support policy. This will aid in the comprehension of the reasons for HM and the accurate prediction of its development trend [21].

Compared with other areas of population research, the research on HM theory is relatively unitary and can be traced back to the Laws of Migration written by Ravenstein [22], a British statistician. Since then, researchers and specialists in fields including population geography,

socioeconomics, and political economy have combined their research to propose several related theories, including neoclassical economic theory, labor market theory, world system theory, migration network theory, and cumulative causality [23,24]. These theories have helped to lay the groundwork for HMP. In traditional theoretical research, a relatively systematic and mature academic system has been established for HM. Many empirical studies have been carried out and have focused on analyzing the relationship between HM and urbanization, as well as the spatial characteristics, policies, causes, and influencing factors of HM. With the fast growth of technologies like big data and AI, researchers have increasingly been looking into this topic. Based on many empirical studies [25,26], they have developed theories like the spatiotemporal network of HM, which constantly adds to the theoretical system of HM.

2.2. Research Progress

The basic task of the HMP are to predict the scale, characteristics and development of HM in the future, on the assumptions of its influencing factors, current situation and development trend. Another task is to predict the future development changes of HM via space-time series abstraction, based on spatiotemporal information, population characteristics and events, and other information. In this regard, the traditional migration prediction methods mainly explore the targets of the migration population quantity (flow, stock, migration rate, etc.) and migration probability. Specifically, the human migration prediction process includes: data acquisition and pre-processing, feature extraction and correlation analysis, model construction, model application, and prediction result output, as shown in Figure 1.

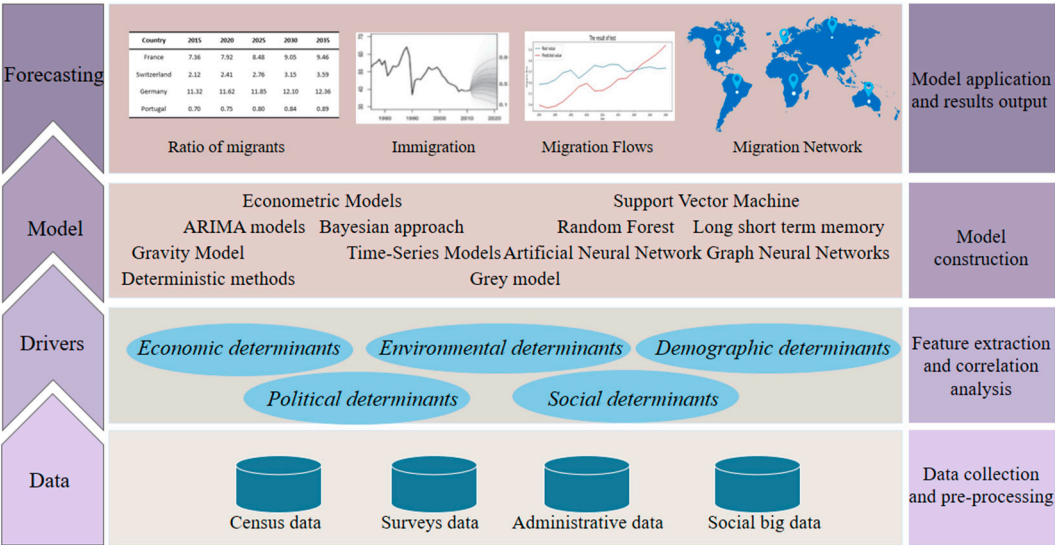


Figure 1. The flow of the human migration forecasting problem.

In the early stage, due to the limitation of technical capacity, HMP used a single mathematical method. With the development of information technology and computerized statistics, HMP based on statistical means has become mainstream. The use of mathematical tools, the accumulation of relevant data, and the use of small sample data have become the main means by which to predict HM. However, incomplete information may lead to prediction bias [27]. In recent years, prediction methods based on big data and AI technology have been proposed [28]. However, despite the difficulty in obtaining government data and consequently their incomplete utilization, these data can provide the government with personal migration preference information to a certain extent, thereby providing a reference for policy formulation and immigration governance. Figure 2 presents the development trend of the use of critical technologies for HMP in recent years. The results revealed that the use of classical prediction models has exhibited a downward trend in recent years, whereas the use of AI models has gradually exhibited an upward trend.

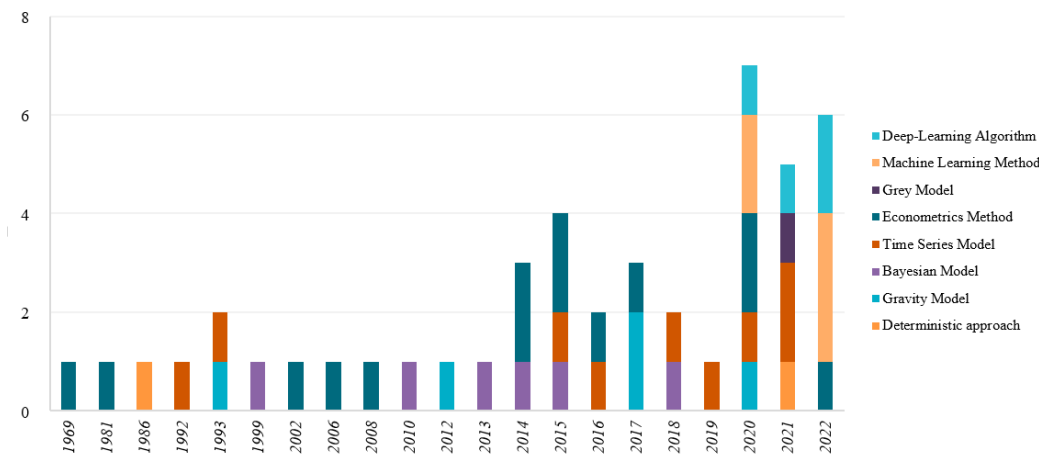


Figure 2. Key technology roadmap for population migration prediction.

In this review, HMP methods are divided into traditional and AI-based prediction methods. The conventional HMP methods include the cohort factor, scene judgment, expert, gravity, Bayesian statistical, and parameter models [11]. The AI prediction methods are mainly divided into ML and deep learning. According to the classification of the prediction results, the traditional prediction models can be divided into deterministic and stochastic approaches [13,26]. Table 1 in the Supplementary Material lists the main methods, data sources, and spatiotemporal characteristics of migration prediction. It can be seen that HM data are mainly government statistics. With the application of big data and AI technology, social media data with geographic characteristic information have been applied to HMP [28]. Based on the extant literature, it is common to see that many studies focus their attention on the prediction of international HM. Moreover, while most scholars have used the dichotomy method to forecast domestic and international HM, some have used the comprehensive prediction method. Due to the limitation of data, most scholars have chosen the use of traditional forecasting methods, mainly econometric models, for HMP, and have primarily focused on the long-term changes of HM. In contrast, new AI methods mainly focus on short-term forecasting.

Table 1. Classification of human migration forecasting methods.

Method	Data Source	Space-time Attributes and References
Deterministic Method	US Department of Health and Human Services and US Bureau of the Census statistics; Human Mortality Database; Federal Statistical Office of Germany.	internal migration, long-term, [29] ; international migration, long-term, [30]; international migration, long-term, [31].
	Statistics NewZealand online database INFOS and Australian Bureau of Statistics database; Eurostat, United Nations Statistics Division statistics and the Council of Europe's Demographic Yearbooks; Office for National Statistics census data; National Records of Scotland and the Office for National Statistics census data; United Nations Population Division's biennial World Population Prospects report; Statistics Sweden census data; Korean Statistical Information Service statistics,	international migration, short-term, [32]; international migration, long-term, [33]; international migration, long-term, [34]; international migration, long-term, [35]; international migration, long-term, [36]; international migration, long-term, [37].

	Australian Bureau of Statistics census data.	
Gravity Model	Internal Revenue Service statistics; Internal Revenue Service statistics; World Bank's Global Bilateral Migration database; Eurostat statistics, World DataBank statistical data(2016), Organization for Economic Co-Operation and Development statistics, French Centre d'Etudes Prospectives et d'Informations Internationales statistics; Google Trends data, Organization for Economic Co-Operation and Development statistics.	internal migration, long-term, [38]; international migration, long-term, [39]; international migration, long-term, [40] ; international migration, long-term, [41]; international migration, short-term, [42].
Time Series Model	Internal Revenue Service statistics; Netherlands Central Bureau of Statistics census data; Mexican demographic surveys of households and American Community Survey data; Household surveys data; Federal Statistical Office of Germany census data; Internal Revenue Service statistics; International migration report(2017); Author-collected datasets; Google Trends data.	internal migration, short-term, [43]; international migration, short-term, [44]; international migration, long-term, [45]; international migration, long-term, [46]; international migration, long-term, [47]; international migration, long and short-term, [48]; international migration, long-term, [49]; internal migration, short-term, [50]; internal migration, short and long-term, [51].
Econometrics Method	Trends in International Migration statistics, Migration Potential in Central and Eastern Europe statistics; Organization for Economic Co-Operation and Development statistics, world bank statistics, Federal Statistical Office census data; The Candidate Country Eurobarometer survey series; Household surveys data; Twitter; Statistics Norway's "Statbank"; World Population Prospects(2015); Russian Federation and Commonwealth of Independent States countries statistics; Google Trends data; Site of the Main Department of Statistics in the Khmelnytskyi Region statistics; World Bank Global Bilateral Migration; Twitter; United States Bureau of the Census data; Data Collected by Author.	international migration, long-term, [52]; international migration, short-term, [53]; international migration, short-term, [54]; international migration, short-term, [55]; international migration, short-term, [56]; international migration, long-term, [57]; international migration, long-term, [58]; international migration, long-term, [59]; international migration, short-term, [60]; internal migration, short-term, [61]; international migration, long-term, [62]; internal migration, short-term, [63]; internal migration, long-term, [64]; internal migration, short-term, [65].
Grey Model	China Statistical Yearbook; US Department of Health and Human Services and US Bureau of the Census statistics.	international migration, long-term, [66]; international migration, long-term, [67].
Machine Learning Method	Data Collected by Author; Mexican Migrant Project statistics; Department of Overseas Labor statistics; Household survey data; Google Trends Index.	internal migration, long-term, [68]; international migration, short-term, [69]; international migration, long-term, [70]; internal migration, --, [71];

		international migration, short-term, [72].
Deep-Learning Algorithm	Nationwide Human Mobility Dataset released by the Spanish Ministry of Transportation; Nationwide Human Mobility Dataset released by the Spanish Ministry of Transportation; Nationwide Human Mobility Dataset released by the Spanish Ministry of Transportation; Google Trends Index, Organization for Economic Co-Operation and Development International Migration Database.	internal migration, short-term, [73]; internal migration, short-term, [74]; internal migration, short-term, [75]; international migration, long-term, [76].

3. Traditional Human Migration Forecasting Approaches

3.1. Deterministic Methods

As a traditional population forecasting method, deterministic approaches are based on the assumption of relevant influencing factors, the independent belief of expert experience, or simple extrapolation, then limited predictions have been output(‘high’, ‘medium’ and ‘low’). The deterministic prediction method has the advantages of a simple concept; low data dependency and can fully utilize the subjective prediction results of experts. Moreover, the technique is simple in operation, suitable for medium- and long-term prediction, and widely used in many countries and international institutions, such as the United Nations (UN) and the European Union (EU) [77]. However, deterministic methods often ignore the influence of many uncertain factors. The limited forecast results do not answer the question "To what extent will the future development of population migration be medium (high or low)? " or "What is the probability of high (medium, low) population migration in the future?" which leads to an insufficient ability to interpret the prediction results. In addition, forecasting results are largely dependent on the knowledge of experts and can be influenced by the scope of expert knowledge, political stance or social attitudes, which can be easily misleading [78].

3.2. Stochastic Methods

In contrast to deterministic methods, the model parameters in stochastic methods are not fixed, but are considered as random variables. It is found that population migration has many random elements, so in order to get closer to reality, some unavoidable random factors are considered in the population migration model to build a random model. In this regard, many scholars use stochastic models to effectively predict population migration [57–59]. Of course, the analysis of stochastic systems is much more difficult than that of relatively deterministic systems. When using stochastic models to describe the migration process, the parameters of the model are difficult to deal with due to the lack of data, and econometric models based on sample data have come to be widely used to predict population migration. In addition, the development of linear and nonlinear migration theory has promoted the research of gravitational model, time series model, Bayesian model and other methods in population migration prediction, making the stochastic model one of the most effective methods for population migration prediction.

3.2.1. Econometric Forecasting Models

Econometric models are mainly used to study the causal relationships between variables of interest, and they are convenient for revealing the relationships between the relative amounts of change of variables; thus, they are widely used in HM forecasting studies. In the 1980s, Plaut used econometric models to forecast the number of net population movements in Texas, USA [64]. Later,

in response to the problem of expanding HM forecasting in the EU, Fertig and Schmidt provided a simple econometric model to forecast the rate of HM from the Czech Republic, Estonia, Hungary, and Poland (four candidate countries) to Germany [79]. Based on this, Dustmann et al. improved this method by adding a relative per capita income variable to predict European migration after the enlargement of the EU in 2004 [80]. However, because the model they used did not fully take into account the non-stationary of HM, but instead treated it as a stationary process, it was subject to significant errors in post hoc tests under the influence of relevant policies. Nevertheless, it was found to be valid for forecasting in Germany, where temporary restrictions on access to its labour market were imposed. A similar study was carried out by Alvarez-Plata et al. using actual income levels, employment rates, population size, and geographic and cultural similarity dummy variables to build a predictive model for HM from CEE10 to EU15 countries [81].

The determinants of HM can be selected and quantified. According to the theory of HM, Cappelen et al. from Statistics Norway used econometric models to set variables such as the income level, unemployment rate, and population size of Norway and the emigration countries, as well as the number of immigrants already living in Norway, to predict the overview of immigrants moving to Norway [82]. In addition, the change in income under the influence of exogenous variables was fully considered, and the corresponding prediction results were provided. In addition, econometric models have also been successfully applied to the prediction of various types of HM (such as skilled talents, college graduates, the labor force, etc.).

While many achievements have been made, the econometric model is also characterized by shortcomings when applied to HMP; in particular, the problem of missing variables often occurs during variable selection. In addition, to reduce the research complexity, researchers tend to ignore some essential characteristics of the population, such as the population size and age structure. Moreover, variable selection error is often a vital source of prediction error.

Given the shortcomings of the econometric model, Dao et al. selected and parameterized appropriate driving factors for international HM, built a structural equation model based on the social economy, solved the parameters via historical data, and then predicted the two-way trend of international HM [83]. The prediction results of the model were found to be consistent with the actual situation, which proved the correctness and feasibility of the model.

Burzynski et al. proposed a comparable institutional equation model by optimizing the model by Dao et al [84]. The model expanded the scope of international migration drivers by introducing fresh factors such as internal migration, technological change, and education. The quantitative analysis of the factors influencing the global distribution of highly skilled people (i.e., domestic educational opportunities, the sectorial distribution of workers, and international migration) conclusively demonstrated that the uneven distribution of labor is a significant contributor to global inequality, that HM is one of the most powerful ways to alter the global distribution of highly skilled people, and that economic inequality within regions can affect the global distribution of highly skilled people. According to the researchers, the model is equally appropriate for predicting the amount of international population flow and the working-age population.

Compared with other methods, the parametric model can be used to analyze the overall effects of individual indicators and the relationships between them, and has strong adaptability to complex systems.

3.2.2. Gravity Model

The gravity model is named for its morphology, similar to Newton's law of gravity, which can effectively explain spatial interaction. The model is suitable for the analysis of regional flow, and has been widely used in the research and prediction of HM. In the 1940s, Zipf proposed a gravity model for HMP; the model holds that the HM between region i and region j is directly proportional to the population of the two areas and inversely proportional to the distance between the two areas, as given by the following equation.

$$M_{ij} \sim \alpha \frac{P_i P_j}{d_{ij}^\beta}, \quad (1)$$

Where M_{ij} is the number of people migrating from region i to region j , P_i is the number of people in region i , P_j is the number of people in region j , d_{ij} is the distance between the two areas, and α and β are two parameters.

Based on the classical gravity model, Beine et al. optimized the model with the actual HM data [85]. The model is as follows:

$$E(m_{ij}) = S_i D_j \phi_{ij}, \quad (2)$$

where $E(m_{ij})$ represents the expected number of migrations from country i to country j , S_i represents the ability of country i to export immigrants, ϕ_{ij} represents bilateral accessibility, which can also be understood as the migration cost between countries i and j , and $D_i = \frac{y_i}{\Omega_i}$ denotes the relative attractiveness of destination j depending on the potential income y_i (wages or GDP) in country j and the relative cost Ω_i of emigrating to other destinations.

In gravity models, attractiveness generally refers to the economic beauty of a particular destination as compared to that of other countries. Because data on expected income are difficult to come by, economists usually use GDP levels or related indices instead [86,87]. When considering the transfer cost, various possible factors are regarded, such as the cost, psychological cost, new language learning cost, etc. [88]. However, in the classical gravity model, the distance parameter incorporates all of these factors, i.e., an increase in the distance between two countries leads to a rise in the cost of migration. In this regard, the fixed effects of similar countries can be regarded as dummy variables.

In addition, the gravity model can also be transformed into a multiple regression equation, and the migration flow at time t can be expressed as follows [85].

$$\ln(m_{ij,t}) = \beta_0 + \beta_1 \ln(GDP_{i,t}) + \beta_2 \ln(GDP_{j,t}) + \beta_3 \ln(distance_{ij,t}) + \beta_4 dummies_{ij,t} + \varepsilon_{ij,t} \quad (3)$$

In the formula, there is a linear relationship between all coefficients and migration flow, and each variable is independent of the others. While the principle of the model is straightforward, a critical problem is ignored, i.e., the individual and invisible characteristics in the process of HM are not considered. In this regard, Backhaus et al. studied the impact of climate change on bilateral migration by using the gravity model, and increased the average temperature and precipitation of the immigration country based on the classical model [89]. In addition, Friebe et al. added changes in smuggling routes based on the gravity model and studied the immigration costs that affect willingness to migrate to a specific location [90].

In summation, based on the interpretation of attraction and distance, the gravity model is applicable to the study of HM-related issues and has strong robustness. The flexible selection of parameters, such as the environment, politics, sociology, micro/macro-economy, geography, etc., enables the better understanding of the driving factors of cross-border migration flows. Although the gravity model provides a reasonable explanation for the spatial pattern of HM, Beyer et al. found that the gravity model based on the time dimension does not perform well [91]. The existing methods are all discussed based on historical data, and the relevant variables are ideal for long-term stability. Unpredictable impacts, such as those of financial crises, war, climate change, or technological progress, are not fully considered, and it is therefore difficult for the results to be convincing.

Because the gravity model contains parameters to be estimated, parameter estimation requires a large amount of historical data. The calculation process is complex, so Simini et al. proposed the radiation model [9].

$$M_{ij} \sim O_i \frac{P_i P_j}{(P_i + d_{ij})(P_j + d_{ij} + P_j)}, \quad (4)$$

Where M_{ij} is the number of migrations from region i to region j , P_i is the number of people in region i , P_j is the number of people in region j , d_{ij} is the distance between the two regions, and O_i

is the total number of trips in area i . This method is weak in has its limitations because it only focuses on the flow between two specific points.

3.2.3. Time-Series Models

In the traditional HMP method, the migration problem can also be abstracted into a time-series problem to be solved. The classical models of time-series prediction mainly include the autoregressive (AR) model, the moving average (MA) model, and the autoregressive integrated moving average (ARIMA) model [48]. Generally, the HM forecast is expressed by the AR (1) model as

$$m_{ij,t+1} = c + \varphi m_{ij,t} + \varepsilon_{t+1}, \quad (5)$$

where $m_{ij,t+1}$ represents the amount of HM between i and j at time $t+1$, $m_{ij,t}$ is the HM quantity between i and j at time t , and ε_{t+1} represents the error at time $t+1$. In the linear model, assuming that the error term satisfies a normal distribution, the parameters can be obtained using historical data φ . The uncertainty and randomness of HM are applied to the non-seasonal ARIMA model, and its differential form is expressed as follows.

$$y'_t = c + \phi_1 y'_{t-1} + \dots + \phi_p y'_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t, \quad (6)$$

In demographic applications, the order of the ARIMA model used for prediction usually does not exceed (1,1,1). Bijak used the ARIMA (0,1,0) model to predict international migration, and the model is expressed as follows [92]:

$$\ln(m_{t+1}) = c + \ln(m_t) + \varepsilon_t, \quad (7)$$

Where m_{t+1} is the predicted number of immigrants, m_t is the number of previous immigrants, c is the drift constant, and ε_t is the error term satisfying the independent normal distribution.

The time-series prediction model can determine the characteristics, trends, and development rules of HM changes according to the time series to effectively predict the future modifications of the HM. However, because the time-series forecasting method does not consider outside factors, there is a defective prediction error; when significant changes take place in HM policy, they tend to have a more substantial deviation, and will produce predicted results that do not tally with the actual situation [12]. Therefore, the effect of the time-series prediction method for short-term prediction is better than that for long-term forecasts.

3.2.4. Bayesian Prediction Model

Bayesian models are considered to be an extension of univariate time-series models, which use probabilistic methods as inputs. In HMP, the number of historical population movements is the only influence; therefore, the method is also considered a purely data-driven approach. Research results have shown that the Bayesian model is more flexible and practical for migration data deficiencies [93]. Due to the incompleteness of HMP data, Bayesian models can all be represented in a probabilistic manner, in which historical trends, expert judgments, and various models are combined in a probabilistic way. In the combined Bayesian model, expert judgment can be used as a prior distribution of the different parameters [33]. The parameters are then updated according to the data.

Azose and Raftery used a Bayesian hierarchical first-order autoregressive model or AR (1) model to achieve a fitted forecast of global HM rates [36]. The model is expressed as

$$(r_{c,t} - \mu_c) = \phi_c (r_{c,t-1} - \mu_c) + \varepsilon_{c,t} \quad (8)$$

where $r_{c,t}$ denotes the HM rate of country c at time t , and $\varepsilon_{c,t}$ is a normally distributed random deviation with mean 0 and variance σ_c^2 . The uncertainty of international migration is quantified based on the posterior distribution by inputting demographic variables. The model enables the long-term forecasting of international migration without causing an explosion of tension.

The advantage of Bayesian models is that probabilistic, rather than quantitative, assessment is used for the estimation of the model parameters, i.e., there is a complete distribution in the Bayesian analysis, not just a parameter. In a Bayesian model, the parameters are considered variables of a random distribution that are extracted from a specific distribution, and the type of distribution of the parameters is used as an additional input variable for the input data. By using this distribution, it is possible to simulate the data following a stochastic process and to derive possible values for the parameters from the assumed distribution using a data generation process.

3.2.5. Expert Prediction Model

In the context of traditional temporal probabilistic prediction, Lutz and Goldstein proposed an expert-based probabilistic population prediction method [94], which can be expressed as

$$v_t = \bar{v}_t + \varepsilon_t \quad (9)$$

Where v_t is the number of population movements studied, \bar{v}_t the average trajectory of the population movement process, which is an a priori assumption derived from the subjective judgment of experts, and ε_t is the chosen stochastic process. However, for exceptional cases (e.g., war, disaster, etc.), expert experience may lose its usefulness and lead to invalid or opposite prediction results. In response, an expert-based algorithm for forecasting population composition (including net migration, etc.) under Bayesian models was proposed by Billari et al. [95]. The method was then further extended by the researchers to population in-migration and out-migration forecasting [96]. However, purely expert methods are limited by the use of too little data, and rely entirely on the subjective judgments of experts. The problem of bimodality may arise when experts make errors in their emotional decisions or when there are differences of opinion among expert groups.

In addition to the lack of a large amount of temporal data on HM, some scholars use the grey model to take a portion of HM information as the research object [66,67]. A grey model is established by extracting sufficient information from known data to achieve an accurate description and grasp of HM development trends.

4. Machine Learning Prediction Methods

Machine learning (ML) is an important branch of artificial intelligence (AI). Its basic principle is to study how computers simulate human learning patterns to automatically acquire knowledge and continuously upgrade their performance [97]. Compared with the traditional research model based on statistics and simulation, the machine learning method can quickly and accurately extract effective associative information from historical data. In recent years, it has been widely used in population migration prediction research.

According to different learning methods, machine learning can be divided into classical machine learning and deep learning. As shown in Figure 3, a range of machine learning methods have been applied in population migration prediction research, including illegal migration prediction, conventional migration prediction, labour migration prediction, migration flow data generation, migration trend prediction, international migration drivers, and asylum seeker prediction [68–76].

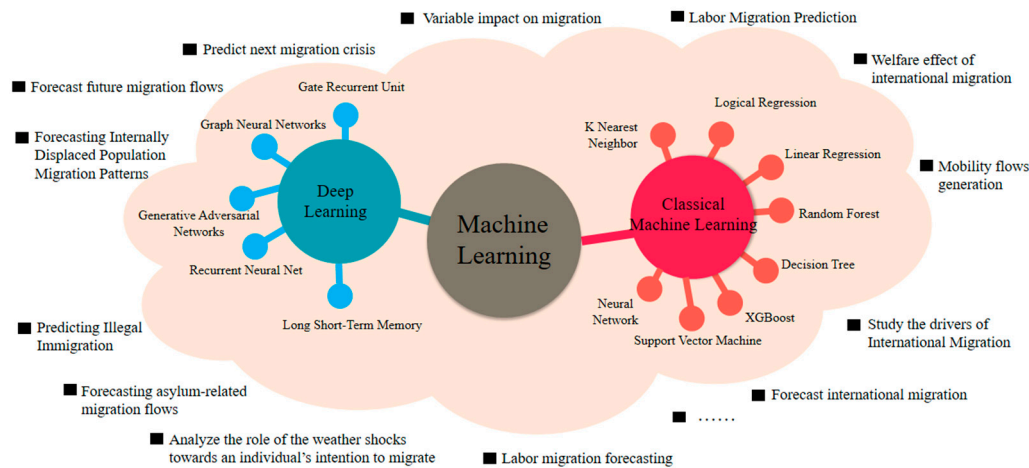


Figure 3. Human migration research using various machine learning methods.

4.1. Classical Machine Learning Prediction Method

4.1.1. Artificial Neural Network

As a model to simulate the structure function and computation of biological neural network, artificial neural network aims to achieve certain functions by simulating some mechanisms and mechanisms of the brain, such as image recognition, speech recognition, etc. Its main structure includes input layer, hidden layer and output layer. After years of efforts and research, machine learning has shown strong advantages in the field of population migration prediction. The prediction method of population migration based on neural network takes the original data or the features extracted based on the original measurement data as the input of the neural network, and constantly adjusts the structure and parameters of the network by a certain training algorithm, and uses the optimized network to predict the development trend of population migration.

Robinson and Dilkina were probably the first to use machine learning models to predict population migration; addressing the inability of traditional linear models to model the non-linear relationship between population migration and its characteristics, while proposing a comprehensive solution to the problems of data imbalance, hyperparameter tuning and performance evaluation in model training, providing a new tool and instrument for population migration prediction [98]. The study successfully uses machine learning as an emerging tool to predict the development trend of domestic and international population migration, demonstrating the advancement and generalizability of the prediction tool, and providing a new reliable tool for the assessment of the future development of population migration and the evaluation of migration management policies.

Tarashev et al. constructed a multi-regional migration-unemployment-wage model by selecting the distribution of migrants, age structure of migrants, wage level, cost of migrating, labor market conditions, regional employment and unemployment information, climatic conditions and distance between countries of origin and destination, etc [99]. An inductive machine learning approach is used to explore the trend of labour migration.

Subsequently, to improve the interpretability of machine learning models, Kiossou et al. used an interpretable machine learning approach to study the drivers of international migration with higher accuracy than the classical gravity model [100]. It also provides a deeper understanding of how migration is affected by its drivers, effectively revealing the interesting non-linear relationship between covariates and outcome variables. To solve the problem of predicting illegal migration, Azizi, S. and Yektansani, K. established a machine learning model that is different from the traditional one [101]. They used eight machine learning techniques to effectively predict the legal status of individuals from Mexico in the United States using data available from Princeton University's Mexican Immigration Project. Based on an adaptive machine learning algorithm, Carammia et al.

developed a Dynamic Elastic Net Model that integrates government statistics and social media data to effectively predict asylum-related migration flows [72]. Giang, N.H. et al. proposed a BPNN model for forecasting labour production and labour migration, and the results show that this method can improve the forecasting performance compared to K-nearest neighbour (kNN) and Random Forest Regression (RFR) models [70].

4.1.2. Random Forest

Random forest is an algorithm that uses multiple decision trees to train, classify and predict samples. It was first proposed by Breiman, L. in 2001 and was mainly applied to regression and classification scenarios [102]. The predicted value of Random Forest is the calculation result based on multiple decision trees (forests), which is usually the mean or mode of the output value of all decision trees. Random forest has the advantages of simple operation, fast training speed and easy to fall into overfitting, so it has become a hot tool for population migration modelling. The focus of population migration prediction research is to use random forest to solve the regression problem. Its basic idea is as follows: One, the autonomous sampling method is used to extract k samples from the original training set, and the sample size of each sample is the same as that of the original training set. Then, k decision tree models are constructed for k samples and k regression results are obtained. Finally, k decision tree results are combined by taking the average values.

To address the problem of forecasting environmental migration, Kelsea Best et al. proposed a random forest model that can effectively identify significant variables from large social surveys [71]. This model can identify the most important predictors of migration from around 2000 original factors, allowing regression analysis with fewer variables and more degrees of freedom.

Aoga et al. proposed a tree-based machine learning (ML) method to predict the impact of weather shocks on individual migration tendencies in six agriculture-dependent economies: Burkina Faso, Cote d'Ivoire, Mali, Mauritania, Niger and Senegal [103]. The results show that climatic factors have positive and significant effects on the predictive performance of individual migration intentions.

4.1.3. Support Vector Machine

SVM was first proposed by Cortes and Vapnik in 1995 [104]. It is mainly used to solve the classification and regression problems of machine learning, and is suitable for analysing small samples and multidimensional data. Different from the traditional neural network learning methods, SVM is developed based on the Vapnik-Chervonenkis (VC) dimension theory and the structural risk minimization principle, which realizes the structural risk minimization (SRM) principle. It also minimizes the empirical risk, has good generalization performance for future samples, and has less samples. The main idea of SVM-based HM forecasting method research is to use the actual HM data to train the support vector machine model, determine the model parameters (insensitivity coefficient, penalty factor, kernel function parameters, etc.), forecasting the future state based on the trained SVM model, and obtain the predicted value of HM by comparing with the pre-set failure threshold.

Lan Zhang et al. employed the support vector machine (SVM) algorithm to create a classification model for the migration of the unregistered resident population in Beijing, and conducted an empirical analysis of migration data from various surveys in Beijing [105]. The results show that SVM performs better than BP neural network and logical regression in accuracy and generalization for these specific classification tasks, and can forecasting migration trends with higher accuracy.

4.2. Deep Learning Prediction Method

The traditional shallow machine learning algorithm relies heavily on expert prior knowledge and signal processing technology, and it is difficult to automatically process and explore massive monitoring data. As a new technology developed from neural networks, deep learning provides a solution for training massive data with its powerful feature extraction capability. However, in the field of migration research, due to differences in relevant concepts and insufficient human, material

and financial resources, the data collected on population migration is sparse or missing, which makes the understanding of population migration lagging behind and inaccurate. Social media data offers a new way to expand the timely perception of complete information on population migration. With the acceleration of globalization, social media and traditional data complement each other, which can further meet the needs of efficient management of population migration.

4.2.1. Recurrent neural networks

Recurrent Neural Network (RNN) is a sort of neural network with short-term memory capability. In RNN, neurons can not only accept information from other neurons, but also accept their own information, forming a network structure with loops. Compared with feedforward neural network, RNN is more consistent with the structure of biological neural network. In tasks such as language modelling and natural language generation, recurrent neural networks (RNN) are suitable for processing sequence data with back-and-forth dependency because of the feedforward and feedback connections between neurons at each layer. The parameters of the recurrent neural network can be learned over time using the back-propagation algorithm. The back-propagation algorithm with time is to transmit the error information step by step in the reverse order of time. If the input sequence is relatively long, there will be gradient explosion and disappearance problems. To solve the problem of RNN, some improved RNN structures have been proposed. Among them, gate current unit (GRU) and long short term memory (LSTM) are typical representatives, which overcome the problems of gradient disappearance and gradient explosion of traditional RNN networks, so that the structure of RNN can not only extract the deep characteristics of time series, but also consider the long-term dependence of time series. It also makes the model obtain better prediction results.

Nicolas Golenvaux et al. use the Long and Short Term Memory Network (LSTM) to predict international migration based on Google trend data, using the thermal coding vector input tag to incorporate more complex time-invariant factors (such as the distance between the two countries and the common language) [76]. The results show that the LSTM method is significantly superior to the standard artificial neural network and the traditional gravity model; in addition, this paper adjusts the LSTM structural model and adds crises to improve the accuracy of the model according to the abnormal impact of a particular year.

4.2.2. Graph Neural Networks

Graph Neural Networks (GNNs), as a generalization of cyclic neural networks, are widely used due to their powerful ability to process complex graph data. By formulating certain strategies on the nodes and edges of the graph, GNN converts the graph structure data into a standardized and standard representation and inputs it into a variety of different neural networks for training, achieving good results in node classification, link prediction, graph clustering and other tasks. In the actual network, the network can be mapped to the relationship between nodes and edges, and the GNN can be used to generate a graph from unstructured data. Its output does not change with the input order of the nodes. The edges represent the dependency between two nodes, and can update the state of nodes by relying on the surrounding state.

Terroso-Sáenz and Muñoz proposed a method for predicting population movement at the national level [73]. This method uses a Graphical Neural Network (GNN) to consider the potential relationship between large geographical regions, and realizes the prediction of population movement between cities on the national spatial granularity; In addition, the author also introduced the impact of climatic factors on population flows. The results show that the effect of weather factors is not obvious due to the mismatch between climate data and model data.

Although ML methods have made some achievements in HM research in recent years, they are currently at a preliminary stage overall, and have had a limited impact on HMP research. To give full play to the predictive power of ML, massive amounts of data are required, and the current sample size of HM research is far from reaching the lower limit that can allow it to yield accurate predictions. In addition, with the complexity of cross-border HM policy changes, the related models lack robustness, and the generalization ability of ML models is poor.

Moreover, the feasibility of the use of Twitter data for predicting domestic HM has been explored to address the problem of an insufficient data volume for deep learning prediction models [63,74]. The results showed that Twitter data have considerable value in HMP. In future studies, focus will be placed on the selection of pertinent data and the design of efficient feature models to further the research on deep learning-based HMP.

5. Uncertainty in Population Migration Projections

As the key to the success of HMP, datasets are traditionally collected using official statistics or survey data compiled and published by relevant organizations. However, different countries have different definitions of HM statistics during collection. During the collection process, government convenience is followed and tools developed for other purposes, rather than specifically designed to measure HM and its outcomes, are used [106]. Even in developed countries like EU nations, information on migrating populations remains sketchy. As a result, HM definitions and data quality issues often cause forecasting methods to be incomparable, forecasts are not accurate, forecasting models are not robust, and there is significant uncertainty [106,107]. Moreover, the numerous social, political, demographic, economic, environmental, and technological drivers in HM forecasting are highly uncertain and difficult to quantify. Their interactions lead to different migration outcomes and bring significant uncertainty to HMP.

Though there has been an accumulation of theories and empirical studies related to HM, no single idea has proven comprehensive enough to cover the multiple forms of migration. The push and pull factors (determinants) or the drivers of migration and non-migration interact with each other, thus making a comprehensive explanation of the migration process impossible even if they cover most cases. Therefore, HM theory has a limited role in the interpretation for the results. Moreover, HM forecasting involves different disciplines, and other experts have entirely different expectations about HM changes. The cumulative effect of all these uncertainties hampers the development of HMP. Therefore, uncertainty discussions should be given attention in all HM projections; otherwise, such uncertainties will spread.

6. Conclusion and Outlook

This paper examined the work on HM forecasting in the past 80 years. The main conclusions include the following: (1) the lack of necessary data will cause misjudgement and affect the development of the industry; (2) deep learning is the future development trend; (3) currently, more consideration is given to the macro level than to the individual level; more research is conducted on the static process than on the dynamic process; finally, there has been more research on international HM than on domestic HM. While the scale and frequency of HM are increasing, the development of HM forecasting research is slow due to the constraints of data and existing technology. The future development of HMP can be carried out from the following aspects.

First, global population growth is gradually slowing down. The driving force of population growth in many countries is mainly rely on HM, which has consequently become an important issue related to the population and political security of some countries. Only by mastering the actual data of HM can scientific and practical measures be taken to ensure the stable development of economies and societies. To ensure the comparability and uniformity of data, statistical authorities of all countries should strengthen cooperation and follow uniform statistical standards. In addition, to ensure the comprehensiveness and adequacy of data, data from transit countries or regions should be collected in addition to data from countries or areas where people move in and out. Data should be made freely available to the public, and government agencies, industry, and academia should collaborate to facilitate rapid research on migration prediction.

Second, with the development of big data and AI technology, the number, type, and fineness of data will be continuously improved, which will challenge the ability of prediction models. The dynamic change of HM depends not only on the differences in the temporal dimension, but also on the transformation of the characteristics in the spatial dimension [108]. The traditional model-driven methods cannot capture the hidden nonlinear features in the spatiotemporal series, so they are

unsuitable for processing these series. As an efficient deep learning framework based on the graph data structure, the GNN is widely used in various fields and has achieved remarkable results. The data of HM flow have the natural attribute of a graph data structure, and the application of GNNs will be an inevitable choice in the future. Therefore, HMP models based on GNNs will be an essential future development direction.

Once again, research on HMP will lead to peaks and valleys in the application development of big data and AI technology. Countries should use the chance for growth to solve the talent problem that has been a longstanding issue in research on HM. Governments should plan to facilitate the migration of skilled people, rely on universities and institutions that carry out population research, and work hard to train reserve talent with an AI technology base for the research and management of HM. Also, different models can be used to improve the skills and abilities of government management teams on the job, thereby creating a talent pool that can be used to manage HM and carry out research in the new era.

Finally, in the traditional process of HM management, although relevant government agencies have obtained a large amount of stock data, they are unable to make full use of these data due to technical limitations. Moreover, no technical support is formed, and the management method is often based on the personal intentions and subjective opinions of managers. With the development of big data and AI technology, the acquisition, preservation, and processing of big data on HM have become possible. Important directions for future development will be to use big data to improve the management level of HM, to establish and improve the mechanism of the scientific decision-making and social management of HM, to promote the innovation of government management and the social governance mode, and to strengthen the research and development of HMP and intelligent auxiliary decision-making systems.

Author Contributions: Tong Zhen Pu, Chong Xin Huang, Jing Jing Yang, and Ming Huang contributed equally to the authorship of the manuscript, including the research design, conducting the research, performing the analysis, and writing the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (NSFC), grant number 61963037, 61863035, 62261059.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. China's State Council. *China population census yearbook 2020*. China Statistics Press: Beijing, China, 2022; pp. 10-23.
2. United Nations. *International Migrant Stock 2019*; (United Nations Database, POP/DB/MIG/Stock/Rev.2019); Department of Economic and Social Affairs, United Nations: Geneva, Switzerland, 2019.
3. Sharma, D. R. , & Kandpal, D. V.. Covid 19 pandemic and international migration: an initial view. *Sustainable Operations and Computers*, **2021**, 2, pp.122-126. <https://doi.org/10.1016/j.susoc.2021.06.002>.
4. Willekens, F., Massey, D., Raymer, J., & Beauchemin, C.. International migration under the microscope. *Science* **2016**, 352(6288), 897-899. <https://doi.org/10.1126/science.aaf6545>.
5. King, R., & Skeldon, R.. "Mind the gap!" integrating approaches to internal and international migration. *Journal of Ethnic and Migration Studies* **2010**, 36(10), 1619-1646. <https://doi.org/10.1080/1369183X.2010.489380>.
6. Zipf, G. K.. The P 1 P 2/D hypothesis: on the intercity movement of persons. *American sociological review* **1946**, 11(6), 677-686. <https://doi.org/10.2307/2087063>.
7. Bogue, D. J.. Internal migration. In *The study of population*. Edited by P. M. Hauser and O. D. Duncan. Chicago: University of Chicago Press, 1959.
8. Lee, E. S.. A theory of migration. *Demography* **1966**, 3(1), 47-57. <https://doi.org/10.2307/2060063>.
9. Simini, F., González, M. C., Maritan, A., & Barabási, A. L.. A universal model for mobility and migration patterns. *Nature* **2012**, 484(7392), 96-100. <https://doi.org/10.1038/nature10856>.
10. Brockmann, D., Hufnagel, L., & Geisel, T.. The scaling laws of human travel. *Nature* **2006**, 439(7075), 462-465. <https://doi.org/10.1038/nature04292>.
11. Disney, G., A. Wiśniowski, J.J. Forster, et al. Evaluation of existing migration forecasting methods and models. Report for the Migration Advisory Committee, 2015; Southampton: University of Southampton. <https://www.gov.uk/government/publications/evaluationof-existing-migration-forecasting-methods-and-models>.

12. Sardoschau, S.. The Future of Migration to Germany. Assessing Methods in Migration Forecasting. DeZIM Project Report #DPR 1|20 Berlin, 2020.
13. Vanella, Patrizio, Philipp Deschermeier, and Christina B. Wilke. An Overview of Population Projections—Methodological Concepts, International Data Availability, and Use Cases. *Forecasting* **2020**, 2(3): 346-363. <https://doi.org/10.3390/forecast2030019>.
14. Van de Walle, E., & Henry, L.. Multilingual demographic dictionary. English Section. Ed. Liege, Belgium: Ordina, 1982.
15. Skeldon, R. *International Migration, Internal Migration, Mobility and Urbanization: Towards more Integrated Approaches*; United Nations: New York, NY, USA, 2018.
16. Otoi, A., Titan, E., & Dumitrescu, R.. Internal and international migration: Is a dichotomous approach justified?. *Procedia-Social and Behavioral Sciences* **2014**, 109, 1011-1015. <https://doi.org/10.1016/j.sbspro.2013.12.581>.
17. Cirillo, M., Cattaneo, A., Miller, M., & Sadiddin, A.. Establishing the link between internal and international migration: Evidence from Sub-Saharan Africa. *World Development* **2022**, 157, 105943. <https://doi.org/10.1016/j.worlddev.2022.105943>.
18. Bernard, A., & Perales, F.. Linking internal and international migration in 13 European countries: complementarity or substitution?. *Journal of Ethnic and Migration Studies* **2022**, 48(3), 655-675. <https://doi.org/10.1080/1369183X.2020.1868983>.
19. Provenzano, D., & Baggio, R.. The contribution of human migration to tourism: The VFR travel between the EU 28 member states. *International Journal of Tourism Research* **2017**, 19(4), 412-420. <https://doi.org/10.1002/jtr.2127>.
20. Möhring, M.. Tourism and Migration: Interrelated Forms of Mobility. *Comparativ* **2014**, 24(2), 116-123. <https://doi.org/10.26014/j.comp.2014.02.08>
21. Marschall, Sabine. "Memory, Migration and Travel." In: S Marschall (ed.). *Memory, Migration and Travel*. Routledge: London, 2018; pp.1–23.
22. Ravenstein, E. G.. The laws of migration. *Journal of the statistical society of London* **1885**, 48(2), 167-235. <https://doi.org/10.2307/2979181>.
23. O'Reilly, K. (2015). *Migration theories: a critical overview*; Routledge handbook of immigration and refugee studies. Abingdon: Routledge, 2015. pp. 25-33.
24. Arango, J.. *Theories of international migration*; International migration in the new millennium, Aldershot: Ashgate, 2017; pp. 25-45.
25. Lewis, G.J. *Human Migration: A Geographical Perspective*; Croom Helm: London, UK, 1982.
26. Sohst, R.; Tjaden, J.; de Valk, H.; Melde, S. *The Future of Migration to Europe: A Systematic Review of the Literature on Migration Scenarios and Forecasts*; International Organization for Migration: Geneva, Germany, 2020.
27. Demirel, D. F., & Basak, M.. A fuzzy bi-level method for modeling age-specific migration. *Socio-Economic Planning Sciences* **2018**, <https://doi.org/10.1016/j.seps.2018.11.001>.
28. Ana Beduschi. International migration management in the age of artificial intelligence, *Migration Studies* **2021**, 9(3):576–596. <https://doi.org/10.1093/migration/mnaa003>.
29. Smith, S. K.. Accounting for migration in cohort-component projections of state and local populations. *Demography* **1986**, 23(1), 127-135. <https://doi.org/10.2307/2061413>.
30. Hyndman, R. J., & Booth, H.. Stochastic population forecasts using functional data models for mortality, fertility and migration. *International Journal of Forecasting* **2008**, 24(3), 323-342. <https://doi.org/10.1016/j.ijforecast.2008.02.009>.
31. Fuchs, J., Söhnlein, D., & Vanella, P.. Migration Forecasting—Significance and Approaches. *Encyclopedia* **2021**, 1(3), 689-709. <https://doi.org/10.3390/encyclopedia1030054>.
32. Gorbey, S., James, D., & Poot, J.. Population forecasting with endogenous migration: An application to trans-Tasman migration. *International Regional Science Review* **1999**, 22(1), 69-101. <https://doi.org/10.1177/0160017997610122>.
33. Bijak J and Wiśniowski A (2010). Bayesian forecasting of immigration to selected European countries by using expert knowledge. *Journal of the Royal Statistical Society A* **2010**, 173 (4): 775-796. <https://doi.org/10.1111/j.1467-985X.2009.00635.x>.
34. Abel, G., Bijak, J., Findlay, A., McCollum, D., & Wiśniowski, A.. Forecasting environmental migration to the United Kingdom: an exploration using Bayesian models. *Population and Environment* **2013**, 35(2), 183-203. <https://doi.org/10.1007/s11111-013-0186-8>.
35. Wiśniowski, A., Bijak, J., & Shang, H. L.. Forecasting Scottish migration in the context of the 2014 constitutional change debate. *Population, Space and Place* **2014**, 20(5), 455-464. <https://doi.org/10.1002/psp.1856>.
36. Azose, J. J., & Raftery, A. E.. Bayesian probabilistic projection of international migration. *Demography* **2015**, 52(5), 1627-1650. <https://doi.org/10.1007/s13524-015-0415-0>.

37. Raymer, J., & Wiśniowski, A. (2018). Applying and testing a forecasting model for age and sex patterns of immigration and emigration. *Population studies* **2018**, 72(3), 339-355. <https://doi.org/10.1080/00324728.2018.1469784>.
38. Frees, E. W.. Short-Term Forecasting of Internal Migration. *Environment and Planning A: Economy and Space* **1993**, 25(11), 1593–1606. <https://doi.org/10.1068/a251593>.
39. Ramos, Raul; Surinach, Jordi. *A Gravity Model of Migration between ENC and EU*, IZA Discussion Papers, No. 7700, Institute for the Study of Labor (IZA), Bonn, 2013.
40. Campos, R. G.. *Migratory pressures in the long run: international migration projections to 2050*. Banco de Espana Article, 2017, 38, 17.
41. Iancu, N., Badulescu, A., Urziceanu, R. M., Iancu, E. A., & Simut, R.. The use of the gravity model in forecasting the flows of emigrants in EU countries. *Technological and economic development of economy* **2017**, 23(2), 392-409. <https://doi.org/10.3846/20294913.2016.1213194>.
42. Böhme, M. H., Gröger, A., & Stöhr, T. (2020). Searching for a better life: Predicting international migration with online search keywords. *Journal of Development Economics* **2020**, 142, 102347. <https://doi.org/10.1016/j.jdeveco.2019.04.002>
43. Frees, E. W.. Forecasting state-to-state migration rates. *Journal of Business & Economic Statistics* **1992**, 10(2), 153-167. <https://doi.org/10.1080/07350015.1992.10509895>.
44. Beer, J. D.. Forecast intervals of net migration: The case of the Netherlands. *Journal of Forecasting* **1993**, 12(7), 585-599. <https://doi.org/10.1002/for.3980120705>
45. García-Guerrero, V. M.. A probabilistic method to forecast the international migration of Mexico by age and sex. *Papeles de población* **2016**, 22(88), 113-140.
46. Schoumaker, B., & Beauchemin, C.. Reconstructing trends in international migration with three questions in household surveys: Lessons from the MAFE project. *Demographic Research* **2015**, 32, 983-1030. <https://doi.org/10.4054/DemRes.2015.32.35>.
47. Vanella, P.; Deschermeier, P. A stochastic Forecasting Model of international Migration in Germany. In *Familie—Bildung—Migration. Familienforschung Im Spannungsfeld Zwischen Wissenschaft, Politik Und Praxis. Tagungsband Zum 5. Europäischen Fachkongress Familienforschung*; Kapella, O., Schneider, N.F., Rost, H., Eds.; Verlag Barbara Budrich: Opladen, Berlin, Germany; Toronto, ON, Canada, 2018; pp. 261–280.
48. Bijak, J., Disney, G., Findlay, A. M., Forster, J. J., Smith, P. W., & Wiśniowski, A.. Assessing time series models for forecasting international migration: Lessons from the United Kingdom. *Journal of Forecasting* **2019**, 38(5), 470-487. <https://doi.org/10.1002/for.2576>
49. Vollset, S. E., Goren, E., Yuan, C. W., Cao, J., Smith, A. E., Hsiao, T., ... & Murray, C. J.. Fertility, mortality, migration, and population scenarios for 195 countries and territories from 2017 to 2100: a forecasting analysis for the Global Burden of Disease Study. *The Lancet* **2020**, 396(10258), 1285-1306. [https://doi.org/10.1016/S0140-6736\(20\)30677-2](https://doi.org/10.1016/S0140-6736(20)30677-2).
50. Shimizu, S., & Shin, S. Applicability of SARIMA Model in Tokyo Population Migration Forecast. In *Proceedings of 2021 14th International Conference on Human System Interaction (HSI)*, Gdańsk-Wrzeszcz, Poland, 2021, pp. 1-4. <https://doi.org/10.1109/HSI52170.2021.9538690>
51. Fantazzini, D., Pushchelenko, J., Mironenkov, A., & Kurbatskii, A.. Forecasting Internal Migration in Russia Using Google Trends: Evidence from Moscow and Saint Petersburg. *Forecasting* **2021**, 3(4), 774-803. <https://doi.org/10.3390/forecast3040048>.
52. Marek Kupiszewski. How trustworthy are forecasts of international migration between Poland and the European Union?, *Journal of Ethnic and Migration Studies* **2002**, 28:4,627-645. <https://doi.org/10.1080/1369183021000032236236>.
53. Brücker, H., & Siliverstovs, B.. On the estimation and forecasting of international migration: how relevant is heterogeneity across countries?. *Empirical Economics* **2006**, 31(3), 735-754. <https://doi.org/10.1007/s00181-005-0049-y>.
54. Bahna, M.. Predictions of Migration from the New Member States after Their Accession into the European Union: Successes and Failures. *International Migration Review* **2008**, 42(4), 844–860. <https://doi.org/10.1111/j.1747-7379.2008.00149.x>.
55. Rogers, T.W. . Migration Prediction On The Basis Of Prior Migratory Behavior: A Methodological Note. *International Migration* 1969, 7: 13-19. <https://doi.org/10.1111/j.1468-2435.1969.tb00286.x>.
56. Zagheni, E.; Garimella, V.R.K.; Weber, I.; State, B. Inferring international and internal migration patterns from Twitter data. In *Proceedings of the 23rd International Conference on World Wide Web*, Seoul, Korea, 7–11 April 2014; pp. 439–444.
57. Cappelen, Å., Skjerpen, T., & Tønnessen, M.. Forecasting Immigration in Official Population Projections Using an Econometric Model. *International Migration Review* **2015**, 49(4), 945–980. <https://doi.org/10.1111/imre.12092>.
58. Azose, J.J.; Sevcikova, H.; Raftery, A.E. Probabilistic population projections with migration uncertainty. *Proc. Natl. Acad. Sci. USA* **2016**, 113, 6460–6465.

59. Vasilyeva, A. V. (2017). The Forecast of Labour Migration, Reproduction of the Population and Economic Development of Russia. *Economy of Regions* **2017**, 13(3):812-826. <https://doi.org/10.17059/2017-3-14>.
60. Böhme, M. H., Gröger, A., & Stöhr, T.. Searching for a better life: Predicting international migration with online search keywords. *Journal of Development Economics* **2020**, 142, 102347.
61. Ovchinnikova, O., Nahornova, O., Mylko, I., Begun, S., Buniak, N., & Kolenda, N. Forecasting Regional Migration Flows. In Proceedings of 10th International Conference on Advanced Computer Information Technologies (ACIT), Deggendorf, Germany, 2020, pp. 165-169).
62. Shayegh, S.; Emmerling, J.; Tavoni, M. International Migration Projections across Skill Levels in the Shared Socioeconomic Pathways. *Sustainability* **2022**, 14, 4757. <https://doi.org/10.3390/su14084757>.
63. Jurdak, R., Zhao, K., Liu, J., AbouJaoude, M., Cameron, M., & Newth, D. (2015). Understanding Human Mobility from Twitter. *PloS one* **2015**, 10(7), e0131469. <https://doi.org/10.1371/journal.pone.0131469>
64. Plaut, T. R.. An econometric model for forecasting regional population growth. *International Regional Science Review* **1981**, 6(1), 53-70. <https://doi.org/10.1177/016001768100600104>.
65. Sun, Y., & Pan, K.. Prediction of the intercity migration of Chinese graduates. *Journal of Statistical Mechanics: Theory and Experiment* **2014**, 12, P12022. <https://doi.org/10.1088/1742-5468/2014/12/P12022>.
66. Geng, Y., Wang, R., Wei, Z., & Zhai, Q.. Temporal-spatial measurement and prediction between air environment and inbound tourism: Case of China. *Journal of Cleaner Production* **2021**, 287, 125486. <https://doi.org/10.1016/j.jclepro.2020.125486>.
67. Tongzheng Pu, Ming Huang, and Jingjing Yang. 2022. Forecasting international migrants using grey model with heat label. In Proceedings of the 5th International Conference on Computer Science and Software Engineering (CSSE, Guilin, China, pp.652–656. <https://doi.org/10.1145/3569966.3571179>.
68. Weber, H.. How well can the migration component of regional population change be predicted? A machine learning approach applied to German municipalities. *Comparative Population Studies* **2020**, 45, 143-178. <https://doi.org/10.12765/CPoS-2020-08>.
69. Azizi, S., & Yektansani, K. (2020). Artificial intelligence and predicting illegal immigration to the USA. *International Migration*, 58(5), 183-193. <https://doi.org/10.1111/imig.12695>.
70. Giang, N.H.; Nguyen, T.-T.; Tay, C.C.; Phuong, L.A.; Dang, T.-T. Towards Predictive Vietnamese Human Resource Migration by Machine Learning: A Case Study in Northeast Asian Countries. *Axioms* **2022**, 11, 151. <https://doi.org/10.3390/axioms11040151>.
71. Best, K., Gilligan, J., Baroud, H., Carrico, A., Donato, K., & Mallick, B.. Applying machine learning to social datasets: a study of migration in southwestern Bangladesh using random forests. *Regional Environmental Change* **2022**, 22(2), 1-12. <https://doi.org/10.1007/s10113-022-01915-1>.
72. Carammia, M., Iacus, S. M., & Wilkin, T.. Forecasting asylum-related migration flows with machine learning and data at scale. *Scientific reports* **2022**, 12(1), 1-16. <https://doi.org/10.1038/s41598-022-05241-8>.
73. Terroso-Sáenz, F., & Muñoz, A.. Nation-wide human mobility prediction based on graph neural networks. *Applied Intelligence* **2022**, 52(4), 4144-4160. <https://doi.org/10.1007/s10489-021-02645-3>.
74. Terroso-Saenz, F., Flores, R., & Muñoz, A.. Human mobility forecasting with region-based flows and geotagged Twitter data. *Expert Systems with Applications* **2022**, 117477. <https://doi.org/10.1016/j.eswa.2022.117477>.
75. Terroso-Sáenz, F., Muñoz, A., Fernández-Pedauye, J., & Cecilia, J. M.. Human Mobility Prediction With Region-Based Flows and Water Consumption. *IEEE Access* **2021**, 9, 88651-88663. <https://doi.org/10.1109/ACCESS.2021.3090582>.
76. Golenvaux, N., Alvarez, P. G., Kiossou, H. S., & Schaus, P.. An LSTM approach to Forecast Migration using Google Trends. *arXiv preprint arXiv:2005.09902*, 2020.
77. Gaigbe-Togbe, V.; Bassarsky, L.; Gu, D.; Spoorenberg, T.; Zeifman, L. World Population Prospects 2022; United Nations: New York, NY, USA, 2022; ISBN 978-92-1-148373-4.
78. Kupiszewski, M.. How trustworthy are forecasts of international migration between Poland and the European Union?. *Journal of Ethnic and Migration Studies* **2002**, 28(4), 627-645. <https://doi.org/10.1080/1369183021000032236236>.
79. Fertig, Michael, and Christoph M. Schmidt. 2005. Aggregate-level migration studies as a tool for forecasting future migration streams. In *International Migration: Trends, Policy and Economic Impact*. London and New York: Institute for the Study of Labor, pp. 110–36.
80. Dustmann, C.; Casanova, M.; Fertig, M.; Preston, I.; Schmidt, C.M. *The Impact of EU Enlargement on Migration Flows*; Home Office Online Report 25/03; Research Development and Statistics Directorate, Home Office: London, UK, 2003; pp. 1–76.
81. Alvarez-Plata, P., Brücker, H., & Siliverstovs, B.. Potential migration from Central and Eastern Europe into the EU-15: An update. European Commission, Directorate-General for Employment and Social Affairs, 2003, Unit A. 1.
82. Cappelen, Å., Skjerpen, T., & Tønnessen, M. (2015). Forecasting Immigration in Official Population Projections Using an Econometric Model. *International Migration Review* **2015**, 49(4), 945–980. <https://doi.org/10.1111/imre.12092>.

83. Dao, T. H., Docquier, F., Maurel, M., & Schaus, P.. Global migration in the twentieth and twenty-first centuries: The unstoppable force of demography. *Review of World Economics* **2021**, 157(2), 417-449. <https://doi.org/10.1007/s10290-020-00402-1>.
84. Burzynski, M., Deuster, C., & Docquier, F.. Geography of skills and global inequality. *Journal of Development Economics* **2020**, 142, 102333. <https://doi.org/10.1016/j.jdeveco.2019.02.003>.
85. Beine, M., Bertoli, S., & Fernández-Huertas Moraga, J.. A practitioners' guide to gravity models of international migration. *The World Economy* **2016**, 39(4), 496-512. <https://doi.org/10.1111/twec.12265>.
86. Hanson, G., & McIntosh, C.. Is the Mediterranean the new Rio Grande? US and EU immigration pressures in the long run. *Journal of Economic Perspectives* **2016**, 30(4), 57-82. <https://doi.org/10.1257/jep.30.4.57>.
87. Bertoli, S., Brücker, H., & Moraga, J. F. H.. The European crisis and migration to Germany. *Regional Science and Urban Economics* **2016**, 60, 61-72. <https://doi.org/10.1016/j.regsciurbeco.2016.06.012>.
88. Sjaastad, L. A.. The costs and returns of human migration. *Journal of political Economy* **1962**, 70(5, Part 2), 80-93. <https://doi.org/10.1086/258726>.
89. Backhaus, A., Martinez-Zarzoso, I., & Muris, C. (2015). Do climate variations explain bilateral migration? A gravity model analysis. *IZA Journal of Migration* **2015**, 4(1), 1-15. <https://doi.org/10.1186/s40176-014-0026-3>.
90. Friebel, G., Manchin, M., Mendola, M., & Prarolo, G.. International migration intentions and illegal costs: Evidence using Africa-to-Europe smuggling routes. CEPR Discussion Paper No. DP13326, 2018, Available at SSRN: <https://ssrn.com/abstract=3290517>.
91. Beyer, R. M., Schewe, J., & Lotze-Campen, H.. Gravity models do not explain, and cannot predict, international migration dynamics. *Humanities and Social Sciences Communications* **2022**, 9(1), 1-10. <https://doi.org/10.1057/s41599-022-01067-x>.
92. Bijak, J. *Migration Assumptions in the UK National Population Projections: Methodology Review*; University of Southampton: Southampton, UK, 2012.
93. Bijak, J. *Forecasting International Migration in Europe: A Bayesian View*; Springer Science+Business Media: Dordrecht, The Netherlands; Heidelberg, Germany; London, UK; New York, NY, USA, 2011.
94. Lutz W and Goldstein JR . Introduction: how to deal with uncertainty in population forecasting? *International Statistical Review* **2004**, 72 (1): 1-4. <https://doi.org/10.1111/j.1751-5823.2004.tb00219.x>.
95. Billari FC, Graziani R, Melilli E. Stochastic population forecasts based on conditional expert opinions. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **2012**, 175 (2):491-511. <https://doi.org/10.1111/j.1467-985X.2011.01015.x>.
96. Billari FC, Graziani R, Melilli E.. Stochastic Population Forecasting Based on Combinations of Expert Evaluations Within the Bayesian Paradigm. *Demography* **2014**, 51 (5):1933-1954. <https://doi.org/10.1007/s13524-014-0318-5>.
97. Mitchell, T. *Machine Learning*; McGraw Hill: New York, NY, USA, 1997.
98. Robinson, C., & Dilkina, B. (2018, June). A machine learning approach to modeling human migration. In Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies, Menlo Park and San Jose CA, USA, pp. 1-8. <https://doi.org/10.1145/3209811.3209868>.
99. Tarasyev, A. A., Agarkov, G. A., & Hosseini, S. I.. Machine learning in labor migration prediction. In Proceedings of AIP Conference, 2018, 1978:440004. AIP Publishing LLC. <https://doi.org/10.1063/1.5044033>.
100. Kiossou, H. S., Schenk, Y., Docquier, F., Houndji, V. R., Nijssen, S., & Schaus, P.. Using an interpretable Machine Learning approach to study the drivers of International Migration. *arXiv preprint* **2020**, arXiv:2006.03560.
101. Azizi, S. and Yektansani, K.. Artificial Intelligence and Predicting Illegal Immigration to the USA. *Int Migr* **2020**, 58: 183-193. <https://doi.org/10.1111/imig.12695>.
102. Breiman, L. Random Forests. *Machine Learning* **2001**, 45(1), 5-32. <https://doi.org/10.1023/A:1010933404324>.
103. Aoga, J., Bae, J., Veljanoska, S., Nijssen, S., & Schaus, P.. Impact of weather factors on migration intention using machine learning algorithms. *arXiv preprint*, 2020, arXiv:2012.02794.
104. Cherkassky, V., & Ma, Y. (2004). Practical selection of SVM parameters and noise estimation for SVM regression. *Neural networks* **2004**, 17(1), 113-126. [https://doi.org/10.1016/S0893-6080\(03\)00169-2](https://doi.org/10.1016/S0893-6080(03)00169-2).
105. ZHANG L, LUO L, HU L, et al. An SVM-based classification model for migration prediction of Beijing. *Engineering Letters*, **2020**, 28(4):1023-1030.
106. Bijak, J.. Migration forecasting: Beyond the limits of uncertainty. Global Migration Data Analysis Centre Data Briefing Series, 2016, (6).

107. Azose, J. J., Ševčíková, H., & Raftery, A. E.. Probabilistic population projections with migration uncertainty. In *Proceedings of the National Academy of Sciences*, 2016, 113(23), 6460-6465. <https://doi.org/10.1073/pnas.1606119113>.
108. Srbu, A., Andrienko, G., Andrienko, N., Boldrini, C., & Sharma, R.. Human migration: the big data perspective. *International Journal of Data Science and Analytics* **2021**, 11(1). <https://doi.org/10.1007/s41060-020-00213-5>.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.