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

Tongzheng Pu 1, Chongxing Huang 2, Jingjing Yang 1,*, and Ming Huang 1,*

1 School of Information Science and Engineering, Yunnan University, Kunming 650500, China
2 Faculty of Social and Historical Sciences, University College London, London WC1E 6BT, UK
* Correspondence: yangjingjing@ynu.edu.cn (J.Y.); huangming@ynu.edu.cn (M.H.)

Abstract: As a fundamental, holistic, and strategic issue facing human society, human migration is a key factor affecting the development of countries and cities, given the 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. Based on the analysis, there is no "best" migration prediction model, and data are key to forecasting human migration. Social media’s popularity and its increase in data have enabled the application of artificial intelligence in population migration prediction, which has attracted the attention of researchers and government administrators. Future research will aim to incorporate uncertainty into the predictive analysis framework, and explore the characteristics of population migration behaviors and their interactions. The integration of machine-learning and traditional data-driven models will provide a breakthrough for this purpose.

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

1. Introduction

With rapid economic and social development occurring worldwide, 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: Geneva, Switzerland, 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 in the pandemic, and the adjustment of relevant policies, have once again led to global increases in internal 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. Policy-makers in labor, healthcare, education, and other such areas 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 industries. 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, policy importance, and greater visibility, 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 the HM drivers and the number of migrants, in order to predict future HM development trends [4]. Studies show that the prediction models used in internal 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, Zipf proposed a gravitational model to predict HM [6]. In the late 1950s, Bogue proposed the famous “push-pull theory” to describe the motives of HM from a kinematic perspective [7]. Later, Lee systematically summarized the push-pull hypothesis, and further outlined the factors influencing migration behavior [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, using approaches based on econometrics, time series analyses, and Bayesian statistics, among others. In recent years, with the rapid development of artificial intelligence (AI) technology, with big data and machine learning (ML) as its core, some scholars have attempted to use ML technology for HMP [11–18]; however, this approach is hampered by limited HM data, varying standards, and access difficulties. In addition, the uncertainty of HM drivers, and the difficulty in quantifying them, have led to the slow development of HMP research [4].

A number of systematic reviews of HMP research have been conducted, providing a comprehensive overview of the current status of HMP research, and a reference for further investigation [19–21]. Nevertheless, with the rapid development of artificial intelligence, and the publication of numerous related studies, it seems necessary to conduct a comprehensive review of population migration prediction methods. There have been many achievements in various migration fields over the past few years, given that population migration is a complex social phenomenon. By synthesizing studies on international and internal migration forecasting in a common context [22], this paper aims to provide a systematic overview of the academic literature and practice in migration modeling and forecasting. In line with this review, we also discuss traditional and machine-learning methods, as well as the advantages and disadvantages of each in population forecasting; and available public datasets, methods, and models are also provided for researchers with some basic knowledge interested of, and interest in, population migration forecasting.

In Section 2, we first review the development of HM theory, although the following is not a complete picture of all the immigration theories, but a basis for further reading. Next, we provide the basic process of population migration prediction, and the enrichment of artificial intelligence methods in HM prediction methods, including a list of the main methods, data sources, and spatiotemporal characteristics of migration prediction. In Section 3, we describe the state-of-the-art HMP methods, and expand on the AI methods. The state-of-the-art methods include the deterministic and stochastic methods. Machine-learning methods include the artificial neural network, random forest, support vector machine, recurrent neural network, and graph neural network models. We also provide guidance on the benefits and limitations of these methods. In Section 4, the uncertainties that lead to the slow growth in HM are analyzed. Finally, conclusions are drawn by analyzing the current status of HMP research, and pointing out the development direction of future research.

2. Research: Related Works and Methodology

2.1. Theoretical Review

The geographical or spatial flow of populations between two regions usually involves the change of a permanent residence from the place of departure to the location of arrival. This is called permanent migration [23]. With the rapid development of transportation and information networks, and the acceleration of globalization, the frequency and time spent traveling are increasing, travel distances are growing, and permanent residence is decreasing [24]. Globally, approximately one out of every 30 people is a cross-border migrant; in China, approximately one out of every 12 people is a trans-provincial migrant.
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. As a result, government organizations do not have a comprehensive and accurate understanding of the entire process of population movement [25]. 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 [26–28]. In addition, tourism migration, as another form of population flow, overlaps and has a causal relationship with migration in many respects [29]. It is difficult to draw a clear line between these two phenomena [24]. It is, therefore, necessary to dismiss the generally limited definition of HM, and adopt a systematic approach, in order to properly understand the interconnections in HM, and to ultimately better support policy. This will aid in comprehending the reasons for HM, and accurately predicting HM trends [30].

Compared with other areas of population research, the research on HM theory is relatively unitary, and can be traced back to British statistician Ravenstein’s Laws of Migration [31]. Since then, researchers and specialists in fields including population geography, socioeconomics, and political economy have combined their research to propose several related theories, including the neoclassical economic theory, labor market theory, world system theory, migration network theory, and cumulative causality [32,33]. 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, 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 such as big data and AI, researchers have increasingly been looking into this topic. Based on many empirical studies [34,35], they have developed theories such as the spatiotemporal network of HM, which constantly adds to the theoretical system of HM.

HM is a complex demographic phenomenon. Most existing migration theories only partially describe the migration phenomenon, and cannot comprehensively and systematically capture its patterns, as it is difficult to directly apply them in migration forecasting (except for the push–pull theory). Researchers commonly translate theories into mathematical models in migration forecasting. Most previous prediction studies have therefore been based on mathematical models, rather than on a particular theory [36]. This study thus provides only a brief description of some of the migration theories.

2.2. Problem Statement

Human migration can be measured in two ways: stocks and flows. Migration stocks refer to the number of migrants in a given place at a given time. Flows refer to events that occur within a given period of time (i.e., migratory movements). The migration stocks change over time, owing to the migrant inflow and outflow. Generally, stocks are easier to measure than flows, because migration dynamics are more difficult to capture [37]. Accordingly, HMP tasks can be divided into two main categories: forecasting HM, and predicting future HM development. The basic task of HMP is to forecast the scale, characteristics, and development of HM, based on its influencing factors, current situation, and developmental trend. Another task is predicting future HM development changes via space–time series abstraction, based on spatiotemporal information, population characteristics and events, and other information. Specifically, the HMP process framework includes data acquisition and pre-processing, feature extraction and correlation analysis, model construction, model application, and prediction result outputs, as shown in Figure 1.

In forecasting population movements, the most important task is to analyze data and to explore the driving factors based on different scenarios. Therefore, migration forecasting involves mining potential patterns from historical data, as well as analyzing driving factors. Economic, political, social, cultural, demographic, and environmental factors are often cited as driving forces for migration [38]. However, population migration is a decision
made in the context of individual or group needs, opportunities, challenges, constraints, urgency, and uncertainty. Therefore, migration drivers are situational and contextual. That is, complex migration scenario drivers are specific to the time and place when the desire to move is formed, and decisions are made. Often, a complex combination of economic, political, social, and other developments and events, not a single driver, dynamically affects migration opportunities, and migrants’ willingness and ability.

In the early stage, due to limited technical capacity, HMP used a single mathematical method. With the development of information technology and computerized statistics, HMP based on statistical methods has come to the fore, with an emphasis placed on the use of mathematical tools to collect relevant data, and the use of small-sample data to predict population movement issues. However, prediction bias and other problems can occur, due to incomplete information [39]. In recent years, prediction methods based on big data and AI technology have been proposed [40]. Artificial intelligence technology helps to inform immigration policy and management, using meaningful information extracted from social media data to predict individuals’ intentions of migrating, as well as predicting migratory movements in times of conflict or natural disaster [39,41].

2.3. Research Methodology

In this review, a systematic approach was used to collect, evaluate, and classify relevant studies. Firstly, in terms of research area, methodological scope, and research background, articles located in digital libraries and online search engines, including Web of Science, Springer JSTOR, Science Direct, Wiley Online Library, and Google Scholar, were searched using the custom keywords “migration”, “prediction”, and “method”. As of 2022, 527 results were retrieved for review. In the second step, all the research results were screened with customized excluded keywords (e.g., data migration, software migration, bird or fish migration). Literature tracking was performed, to add potentially relevant studies using the literature wizard linked paper, and a total of 320 papers were finally obtained. We then independently reviewed the resulting papers for consistency with the
application scenarios and classification methods of this study, selecting a total of 54 method-specific papers. As shown in Figure 2, we compared the publication trends in the field of population migration prediction, by plotting the number of references to the main prediction methods. As shown in the figure, there have been a considerable number of new research papers on population displacement forecasting, using the neural network family of techniques, since 2019.

Figure 2. Key technology roadmap for population migration prediction. Source: results of systematic literature search by author.

In this review, HMP methods are divided into traditional and machine-learning prediction methods. According to the classification of the prediction results, traditional prediction models can be divided into deterministic and stochastic approaches [19,21,35]. Deterministic methods are also called fixed-value forecasting methods, and scenario forecasting methods. Since deterministic forecasting methods are mainly applied in developed countries or international organizations (e.g., United Nations, Eurostat), and by few researchers, there are also few related research papers, so this paper does not list the specific methods. Machine-learning prediction methods are divided into the classical machine-learning method, and deep-learning methods. Table 1 lists the main methods, data sources, and spatiotemporal characteristics of migration prediction, and reveals that HM data are mainly in the form of government statistics, based on the extant literature; many studies focus on the prediction of international HM. Moreover, while most scholars have used the dichotomy method to forecast internal and international HM, some use the comprehensive prediction method. Due to the limitations of the data, most scholars have chosen traditional forecasting methods (econometric models, for the most part) for HMP, and have primarily focused on the long-term changes. In contrast, new AI methods mainly focus on short-term forecasting.

Table 1. Classification of human migration forecasting methods.

<table>
<thead>
<tr>
<th>Method</th>
<th>Model</th>
<th>Data Source</th>
<th>Spatial Attributes</th>
<th>Temporal Attributes</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deterministic method</td>
<td>--</td>
<td>US Department of Health and Human Services and US Bureau of the Census statistics</td>
<td>internal migration</td>
<td>long-term</td>
<td>[42]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Human Mortality Database; Federal Statistical Office of Germany</td>
<td>international migration</td>
<td>long-term</td>
<td>[43,44]</td>
</tr>
<tr>
<td>Method</td>
<td>Model</td>
<td>Data Source</td>
<td>Spatial Attributes</td>
<td>Temporal Attributes</td>
<td>References</td>
</tr>
<tr>
<td>------------------</td>
<td>------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>--------------------</td>
<td>---------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Bayesian Model</strong></td>
<td>Stochastic methods</td>
<td>Statistics New Zealand online database INFOS and Australian Bureau of Statistics database</td>
<td>international migration</td>
<td>short-term</td>
<td>[45]</td>
</tr>
<tr>
<td><strong>Stochastic methods</strong></td>
<td>Time Series Model</td>
<td>Internal Revenue Service statistics; Netherlands Central Bureau of Statistics census data</td>
<td>international migration</td>
<td>short-term</td>
<td>[52–54]</td>
</tr>
<tr>
<td></td>
<td>Mexican demographic surveys of households and American Community Survey data; Household surveys data; Federal Statistical Office of Germany census data</td>
<td>international migration</td>
<td>long-term</td>
<td>[55]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Internal Revenue Service statistics; International migration report (2017) internal migration short and long-term</td>
<td>international migration</td>
<td>long and short-term</td>
<td>[61,62]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Author-collected datasets; Google Trends data</td>
<td>internal migration</td>
<td>short and long-term</td>
<td>[63,64]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>United States Bureau of the Census da-ta;</td>
<td>internal migration</td>
<td>long-term</td>
<td>[74]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Site of the Main Department of Statistics in the Khmelnytskyi Region statistics; Twitter; data collected by the author.</td>
<td>internal migration</td>
<td>short-term</td>
<td>[11,75,76]</td>
<td></td>
</tr>
<tr>
<td><strong>Grey Model</strong></td>
<td>China Statistical Yearbook; US Department of Health and Human Services and US Bureau of the Census statistics.</td>
<td>international migration</td>
<td>long-term</td>
<td>[12,13]</td>
<td></td>
</tr>
</tbody>
</table>
3. Human Migration Forecasting Approaches

Research on migrant prediction is diverse, with a large number of researchers using qualitative research techniques (ethnographic fieldwork, in-depth interviews, etc.). The main focus of this paper is on quantitative methods for forecasting population migration, and only a brief description of the qualitative methods will be given. As mentioned in Section 2, we will introduce methods of predicting population migration, and provide empirical examples based on deterministic, stochastic, and artificial intelligence methods.

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, resulting in limited predictions (‘high’, ‘medium’, and ‘low’). The deterministic prediction method has the advantages of being a simple concept; it has 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) [81]. 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?” This 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 easily be misleading [65].

3.2. Stochastic Methods

In contrast to deterministic methods, stochastic method model parameters are not fixed, but are considered as random variables. Population migration consists of many random elements; therefore, 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 [66, 72, 73]. Of course, the analysis of stochastic systems is much more difficult than the analysis of relatively deterministic systems. Model parameters are difficult to deal with when using stochastic models to describe the migration process, owing to the lack of data, and econometric models based on sample data are now widely used to predict population migration. In addition, the development of linear and nonlinear migration theory has promoted research into the gravitational model, the time series model, the Bayesian model, and other population migration prediction methods, making the stochastic model one of the most effective methods for population migration prediction.

Table 1. Cont.

<table>
<thead>
<tr>
<th>Method</th>
<th>Model</th>
<th>Data Source</th>
<th>Spatial Attributes</th>
<th>Temporal Attributes</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Machine-learning Method</td>
<td>Classical Machine-Learning Method</td>
<td>Data collected by the author; Department of Overseas Labor statistics, Household survey data, Mexican Migrant Project statistics, Google Trends Index</td>
<td>internal migration</td>
<td>long-term</td>
<td>[14]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>international migration</td>
<td>long-term</td>
<td>[16]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>internal migration</td>
<td>--</td>
<td>[17]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>international migration</td>
<td>short-term</td>
<td>[15,18]</td>
</tr>
<tr>
<td></td>
<td>Deep-Learning Method</td>
<td>Nationwide Human Mobility Dataset released by the Spanish Ministry of Transportation, Google Trends Index, Organization for Economic Co-Operation and Development International Migration Database</td>
<td>internal migration</td>
<td>short-term</td>
<td>[77–79]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>international migration</td>
<td>long-term</td>
<td>[80]</td>
</tr>
</tbody>
</table>
3.2.1. Econometric Forecasting Models

Econometric models are mainly used to study causal relationships between variables of interest, and they are convenient for revealing the relationships between the relative amounts of variable change; thus, they are widely used in HM forecasting studies. Ordinary least squares (OLS), the generalized method of moments (GMM), several versions of the random effects (GLS), seemingly unrelated regressions (SUR), and others are among the most commonly used models for HMP. In 1981, Plaut used econometric models to forecast net population movements in Texas, USA [74]. The net civilian migration \( \text{NCM}_t \) in the time period \( t \) is modeled as:

\[
\ln \left[ 1 + \left( \frac{\text{NCM}_t}{\text{CPOP}_{t-1}} \right) \right] = \alpha \ln \lambda_0 + \lambda \beta_1 \ln \text{RW}_t + \lambda \beta_2 \ln \text{RVU}_t + \lambda \alpha_1 \ln \text{USY}_t - \lambda \ln \text{CPOP}_{t-1} + \mu_t
\]  

where \( \text{RW}_t \) is the real wage rate, \( \text{RVU}_t \) is the vacancy-to-unemployment ratio in Texas relative to the United States, \( \text{USY}_t \) is the real per capita income in the United States, and \( \text{CPOP}_{t-1} \) is the anticipated net civilian population of Texas.

Later, in response to the problem of HM forecasting in the EU-enlargement, 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 [82]. The migration rate \( m_{s,t} \) in the relevant age range for the origin country \( s \) and the period \( t \) is given by:

\[
m_{s,t} = \mu + \epsilon_s + \epsilon_t + \epsilon_{s,t}
\]

where \( \mu \) is an overall intercept term, \( \epsilon_s \) is a random component specific to country \( s \) but persistent over time \( t \), \( \epsilon_t \) is a component specific to time periods but relevant for all countries at this point in time \( t \), and \( \epsilon_{s,t} \) is an unpredictable white noise error term.

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 [83]. For a sending country \( s \) in the year \( t \), the aggregate migration rate \( m_{s,t} \) is given by:

\[
m_{s,t} = \mu_s + X_{s,t} \beta_s + \delta m_{s,t-1} + \epsilon_{s,t}
\]

where the parameter \( \mu_s \) captures all unobservable aspects of the process that are specific to country \( s \) but constant over time, \( X_{s,t} \) is the observable time-varying characteristics of country \( s \) at time \( t \), \( \beta_s \) are vectors of unknown parameters to be estimated, and \( \epsilon_{s,t} \) is the error term reflecting all unsystematic influences on the process. The method applied in that paper does not fully take into account the non-stationary nature of HM, but treats it as a stationary process. As a result, the model includes a large error in post hoc tests under the influence of relevant policies. Nevertheless, the method was found to be valid for forecasting in Germany, where temporary restrictions on access to the labor market were imposed.

A similar study, carried out by Alvarez-Plata et al., used actual income levels, employment rates, population size, and geographic and cultural similarity dummy variables to build a predictive model for HM from central and eastern European (CEE10) to European Union (EU15) countries [84]. Specifically, they model the share of migrants from country \( h \) residing in country \( f \) expressed as \( \text{mst}_{fh} \text{t} \) :

\[
\text{mst}_{fh} = \alpha + (1 - \delta) \text{mst}_{fh, t-1} + \beta_1 \ln \left( \frac{w_{fh}}{\text{w}_{ht}} \right) + \beta_2 \ln (\text{w}_{ht}) + \beta_3 \ln \left( \frac{e_{fh}}{e_{ht}} \right) + \beta_4 \ln \left( \frac{\text{p}_{fh}}{\text{p}_{ht}} \right) + Z_{fh} \gamma + u_{fh} \text{t}
\]

where \( w \) is the wage, \( e \) is the employment rate, and \( \text{p}_{fh} \) is the population in the home country. \( u_{fh} = \mu_{fh} + v_{fh} \text{t} \), \( \mu_{fh} \) denotes a country-specific effect, and \( v_{fh} \text{t} \) is white noise.

The determinants of HM can be selected and quantified. Using HM theory, Cappelen et al. used econometric models to set variables, such as the income level, unemployment rate, and population size of Norway and of migrant countries of origin, as well as the number of...
immigrants already living in Norway, to create an overview prediction of immigrants to Norway [66]. The model is given by:

$$
\log \left( \frac{M_j}{POP_j} \right) = \beta_0 + \beta_1 \log \left( \frac{Y_{NOR}}{Y_j} \right) + \beta_2 \log \left( \frac{IS_j}{POP_{NOR}} \right) + \beta_3 U_{NOR} + \beta_4 U_j + \beta_5 D_j \tag{5}
$$

where the left-hand-side variable is the log of the migration group $j$, $Y_{NOR}/Y_j$ is the relative per cap income, $IS_j/POP_{NOR}$ is a proxy for the migrant network, $U_{NOR}$ is the unemployment rate in Norway, $U_j$ is the unemployment rate in the sending areas, and $D_j$ is a vector of intervention dummy variables that captures political events relating to changes in either country that are relevant to people who try to enter Norway.

In addition, the change in income under the influence of exogenous variables was fully considered, leading to corresponding predictions. Econometric models have also been successfully applied to the prediction of various types of HM (such as skilled, college graduate, and labor force migration) [85].

While there have been many achievements, the econometric model has certain shortcomings when applied to HMP. In particular, missing variables are often a problem during variable selection. In addition, to reduce the research complexity, researchers tend to ignore some essential population characteristics, such as the population size and age structure. Moreover, the 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 socio-economic structural equation model, solved the question of parameters using historical data, and then predicted the two-way trend of international HM [85]. At time $t$, the utility of a type-$s$ individual born in country $I$ and living in country $j$ is given by:

$$
u_{ij,s,t} = \gamma \ln w_{j,s,t} + \ln v_{ij,s,t} + \xi_{ij,s,t} \tag{6}
$$

where $w_{j,s,t}$ is the wage rate attainable in the destination country $j$; $\gamma$ is a parameter governing the marginal utility of income; $v_{ij,s,t}$ stands for the nonwage income and amenities in country $j$ (public goods, non-monetary amenities, and transfers minus taxes), and is netted from the legal and private costs of moving from $i$ to $j$; $\xi_{ij,s,t}$ is the random taste component, capturing heterogeneity in the preferences for alternative locations, in mobility costs, in assimilation costs, etc. The prediction results were found to be consistent with the actual situation, which supported the correctness and feasibility of the model.

Burzynski et al. optimized Dao et al.’s model to propose a comparable institutional equation model [86]. The model expanded the scope of international migration drivers by introducing new factors, including internal migration, technological change, and education. The quantitative analysis of the factors influencing the global distribution of highly skilled people (i.e., internal educational opportunities, the sectorial distribution of workers, and international migration) 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 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 a 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 which holds that the HM between two regions is
directly proportional to the population of the two areas, and inversely proportional to the distance between them, as given by the following equation:

\[ M_{ij} \sim \frac{P_i P_j}{d_{ij}^\beta} \tag{7} \]

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 \( \alpha \) and \( \beta \) are two undetermined parameters.

Based on the classical gravity model, people have primarily concentrated on the study of models based on actual HM data, which includes model optimization, empirical research, the explanation of global migration patterns, and the impact of economy, society, politics, culture, and climate on immigration, among others [87].

Beine et al. optimized the model with actual HM data [88]. The expected number of migrations between two countries, as a function of the source country’s ability to send migrants, the cost of migration between countries, and the relative attractiveness of the destination country. 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 [89, 90]. When considering the transfer cost, various possible factors are considered, such as the cost, psychological cost, new-language-learning cost, etc. [91]. 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 [88]. In the formula, there is a linear relationship between all coefficients and the 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 using the gravity model, and increased the average temperature and precipitation of the immigration country based on the classical model [92]. Migration inflows from country \( i \) to country \( j \) in year \( t \) were modeled as \( M_{ijt} \):

\[
\ln M_{ijt} = a_0 + a_1 wtemp_{ijt} + a_2 wpre_{ijt} + a_3 GDP_{it} + a_4 GDP_{jt} + a_5 DemPres_{it} + a_6 \ln \text{Population}_{it} + a_7 U_{jt} + a_8 Trade_{it} + \gamma_1 + \omega_{ijt} + \epsilon_{ijt} \tag{8}
\]

where \( wtemp_{ijt} \) is the population-weighted average annual temperature in degrees Celsius, \( wpre_{ijt} \) is the average annual precipitation in millimeters, \( GDP_{it} \) is the PPP-adjusted GDP per capita divided by a factor 1000 in the origin (destination) country in year \( t \), \( DemPres_{it} \) is the share of young people in the country of origin’s working-age population, \( U_{jt} \) is the unemployment rate in the country of destination at time \( t \), \( Trade_{it} \) is the openness ratio in the country of origin at time \( t \); \( \omega_{ijt} \) are the country-pair characteristics, \( \gamma_1 \) are a set of year dummies that captures global shocks, and \( \epsilon_{ijt} \) is the error term.

In addition, Friebel et al. added changes in smuggling routes, based on the gravity model, and studied the immigration costs that affect the willingness to migrate to a specific location [93]. For the individual \( i \) to move from country of origin \( o \) to country of destination \( d \) at time \( t \), the binary indicator of migration intention (\( M_{iodt} \)) is:

\[
M_{iodt} = \beta_0 + \beta_1 X_{iot} + \beta_2 D_{odt} + v_{ot} + w_{dt} + u_{od} + \epsilon_{iodt} \tag{9}
\]

where \( D_{odt} \) is the time-varying (log of) distance, measuring the distance between the origin and destination country along migration routes. \( X_{iot} \) is a vector of individual covariates, which include age, gender, education, household size, wealth, urban city residence, and
satisfaction with local amenities. \( u_{od} \) is the country-pair fixed effects together with a full set of origin-by-year \((v_{od})\) and destination-by-year \((w_{dt})\).

Campos proposed an extended gravity equation that allows projections to be made for all pairs of countries in the world using small explanatory variables. The results showed that the number of migrants is projected to increase from 2.8% of the world population in 2010, to around 3.5% in 2050. Due to the assumption of migration–population relationships, without consideration of future policy changes and unexpected events, the prediction results of this method may be biased [53].

Nicolaie et al. proposed a simplified gravity model to study emigration in Romania. Because of limited access or incomplete statistical data, Romanian migration flows were analyzed only to certain EU countries, and at the EU-27 level, for the period 1995–2014. The analysis showed that for every 1% increase in unemployment, the number of emigrants increased by 2.57 [54]. The flow of migrants from region/country \( i \) to region/country \( j \) in time \( t \) is given by:

\[
\ln M_{ijt} = \beta_0 \ln M_{ijt} + \beta_1 \ln A_{it} + \gamma_1 A_{jt} + \beta_2 \ln Y_{it} + \gamma_2 Y_{jt} + \beta_3 \ln C_{ijt} + g_t + f_{ij} + \epsilon_{ijt} \tag{10}
\]

where \( Y_{it}/Y_{jt} \) means the income in location \( i/j \) at time \( t \), \( A_{it}/A_{jt} \) are amenities in \( i/j \) at time \( t \), \( C_{ijt} \) are the costs of moving from location \( i \) to \( j \), \( g_t \) is a time effect variable, \( f_{ij} \) are specific effects, and \( \epsilon_{ijt} \) is the error term.

The nonlinear gravity model by Rikani and Schewe (2021), used to project future global migration trajectories, matches spatio-temporally pooled observed flows well. Although this model does a good job of reproducing past patterns and trends using fewer parameters, a gap remains between the predictions and real values at the country level [94].

In summary, 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, and geography, among others, enables a better understanding of the driving factors in 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 [95]. 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]. In this model, when choosing a destination, the total number of individuals is proportional to the population of the source and the destination, and decreases with distance. The movement is based on this criterion. This method is weak, and has limitations, because it only focuses on the flow between two specific points.

3.2.3. Time-Series Models

In the traditional HMP method, time-series analysis and extrapolation is another important method for migration forecasting. 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 [61].

Generally, for a long enough series, the different migration processes \( (m_t) \) can be expressed by the unconstrained autoregressive model of the first order (AR (1)), with a constant \( c \) describing a stationary process whenever the autoregression parameter \( \Theta \in (-1, 1) \) applies:

\[
\ln(m_t) = c + \varphi \ln(m_{t-1}) + \epsilon_t \tag{11}
\]
In addition, ARMA (1, 1) is also in effect, with a moving average element added to the AR (1) model above, with an additional parameter $\theta$:

$$\ln(m_t) = c + \varphi \ln(m_{t-1}) + \epsilon_t + \theta \epsilon_{t-1}$$ (12)

AR (1) modeling was used to forecast state-to-state migration rates from Free, based on the recently revised and updated US Bureau of the Census database. From this, it follows that recent rates are important in predicting internal migration in the near term [56]. Later, Beer used an ARIMA model to specify a statistical forecast interval in the Netherlands. The results showed that the uncertainty of immigration has been underestimated in official forecasts in the past [57].

To forecast international migration to and from the UK, Bijak et al. used time-series models with and without expert opinion, including ARIMA models, autoregressive distributed lag models, and past-error propagation. It was found that the low order ARIMA models perform better with stationary data [61].

A SARIMA model for predicting the entry into, and exit from Tokyo was developed by Shimizu and Shin. To improve the accuracy, factors such as the COVID-19 crisis were added. The results reflected that the model could be sufficiently used for such short-to-medium-term time-series data [63]. Similarly, three classes of model were employed by Fantazzini et al. for out-of-sample forecasting of interregional migration in Russia; these include short-term forecasting using ARIMA and Google-augmented ARIMA models, as well as multivariate models for long-term forecasting. The empirical analysis finds that including Google Trends data in a model enhances the prediction of migration flows [64].

The time-series prediction model can determine the characteristics, trends, and development rules of HM changes according to the time series, to effectively predict future modifications in 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 show a more substantial deviation, and will produce predicted results that do not tally with the actual situation [20]. 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 [96]. 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.

To forecast immigration for seven European countries, Bijak presented a Bayesian model, based on quantitative data and qualitative knowledge elicited from country-specific migration experts in a two-round Delphi survey. In the combined Bayesian model, expert judgment can be used as a prior distribution for the different parameters [46]. The parameters are then updated according to the data.

Responding to the potential impact of the environment on migration, Abel investigated how to use Bayesian modeling to predict uncertainty about the level of immigration to the UK resulting from environmental factors elsewhere [47]. They considered a set of autoregressive (AR) time-series models based on the $k$-year history of immigration, AR $(k)$, defined as follows:

$$\begin{align*}
    m_t &= \mu + \sum_{i=1}^{k} \varphi_i (m_{t-i} - \mu) + \epsilon_t, & t < 2010 \\
    m_t &= v + \sum_{i=1}^{k} \varphi_i (m_{t-i} - v) + \epsilon_t, & t \geq 2010
\end{align*}$$ (13)
where \( m_t \) is the transformed immigration in year \( t \), and \( \mu \) and \( \nu \) are the mean level of \( m_t \) from the observed data series and forecasted future data series. The parameters for \( \phi_i \) are the ensemble of the autoregression coefficients of \( m_t \), related to its past history up to \( k \) periods (years) before \( \varepsilon_t \sim N(0, \sigma^2_c) \).

Later, the possible effects of Scottish independence on internal and international migration were studied by Wisniowski et al. [48]. The predictions presented in this paper were obtained from the results of the Bayesian forecasting model, and take into account different sources of uncertainty of future migration flows.

Azose and Raftery used a Bayesian hierarchical first-order autoregressive model or AR (1) model to achieve a fitted forecast for global HM rates [49]. In this model, 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 author modelled the migration rate \( r_{c,t} \) in country \( c \) and time period \( t \) as follows:

\[
(r_{c,t} - \mu_c) = \phi_c(r_{c,t-1} - \mu_c) + \varepsilon_{c,t}
\]

where \( \varepsilon_{c,t} \) is a normally distributed random with a mean of zero and a variance of \( \sigma^2_c \), \( \mu_c \) is the long-term average migration rate, and \( \phi_c \) is the uniform prior on the autoregressive parameter, \( \varepsilon_{c,t} \sim N(0, \sigma^2_c) \).

Wisniowski et al. developed a Bayesian approach to forecast immigration (counts) and emigration (rates) by age and sex for the UK, which can be adapted to different data types and information sources [97].

In the short term, the applicability of the model to other country situations and types of data has been tested by Raymer and Wisniowski. Time-series data for Sweden, South Korea, and Australia were also used to validate the predictive accuracy and generalizability of the model [50].

The advantage of Bayesian models is that probabilistic, rather than quantitative, assessment is used to estimate the model parameters; i.e., there is a complete distribution in the Bayesian analysis, not just a single 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

Expert forecasting allows for timely adjustments to policy releases and expected changes, contains more information than model forecasting, and makes it easier to recreate the forecasting process, which has become an important method for population forecasting. A single forecasting model often contains only part of the information of the forecast object, and combining each single model according to certain rules can improve the forecast accuracy, by including more comprehensive forecast information. In this regard, in population migration prediction, expert prediction models rarely appear alone, and are often combined with other prediction models, in a process that can combine the advantages of both, and improve the prediction accuracy.

In the context of traditional temporal probabilistic prediction, Lutz and Goldstein proposed an expert-based probabilistic population prediction method [98]. The number of population movements can be obtained by the average trajectory of the population movement process, an a priori assumption derived from the subjective judgment of experts, and a selected 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) under Bayesian models was proposed by Billari et al. [99]. This method was then further extended by the researchers, to population in-migration and out-migration.
forecasting [100]. 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 decisions, or when there are differences of opinion among expert groups.

In addition to the lack of a large amount of temporal HM data, some scholars use the grey model to take a portion of HM information as the research object [12,13]. A grey model is established by extracting sufficient information from known data to achieve an accurate description and grasp of HM development trends.

3.3. Machine-Learning Prediction Methods

Machine learning (ML) is an important branch of artificial intelligence (AI), the basic principle of which is to study how computers simulate human learning patterns, to automatically acquire knowledge and continuously upgrade their performance [101]. Compared with a 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 HMP 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 HMP research, including illegal migration prediction, conventional migration prediction, labor migration prediction, migration flow data generation, migration trend prediction, international migration drivers, and asylum seeker prediction [14–18,77–80].

![Figure 3. Human migration research using various machine-learning methods.](image)

3.3.1. Classical Machine-Learning Prediction Method

1. Artificial Neural Network

As a model to simulate the structure, function, and computation of the biological neural network, an artificial neural network aims to achieve certain functions by simulating some mechanisms of the brain, such as image recognition and speech recognition, among others. Its main structure includes an input layer, a hidden layer, and an output layer. After years of effort and research, machine learning has shown strong advantages in the field of population migration prediction. The prediction of population migration based on the 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 using a certain training algorithm. The optimized network is then used to predict the development trend in HM.

Robinson and Dilkina were likely the first to use ML models in HMP. They used these models to address the inability of traditional linear models to model the non-linear relationship between population migration and its characteristics, while proposing a comprehen-
sive solution to the problems of data imbalance, hyperparameter tuning, and performance evaluation in model training, thus providing a new tool and instrument for HMP [102]. The study successfully used machine learning as an emerging tool to predict development trends in internal and international population migration, demonstrating the advances in, and generalizability of, the prediction tool. Machine learning, therefore, offers a new reliable tool for the assessment of future developments in population migration, and the evaluation of migration management policies.

Tarasiev et al. constructed a multi-regional migration-unemployment-wage model that uses an inductive ML approach to explore labor migration trends, based on the migrant distribution, age structure, income level, cost of migrating, labor market conditions, regional employment and unemployment information, climatic conditions, and the distance between the countries of origin and destination, among other variables [103].

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 greater accuracy than the classical gravity model [104]. This approach also provides a deeper understanding of how migration is affected by drivers, effectively revealing the non-linear relationship between covariates and outcome variables. To solve the problem of predicting illegal migration, Azizi and Yektansani established a machine-learning model based on eight machine-learning techniques that effectively predict the legal status of individuals from Mexico in the United States, using data available from Princeton University’s Mexican Immigration Project [16]. 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 [77]. Giang, et al. proposed a BPNN model for forecasting labor production and labor migration, and the results show that this method can improve the forecasting performance, compared to the K-nearest neighbor (kNN) and random forest regression (RFR) models [17].

2. Random Forest

Random forest is an algorithm that uses multiple decision trees to train, classify, and predict samples. First proposed by Breiman in 2001, it has mainly been applied to regression and classification scenarios [105]. The predicted value of random forest is the calculation result of multiple decision trees (forests), which is usually the mean or mode of the output value of all decision trees. The advantages of random forest include that it is simple to operate, has a fast training speed, and is not easy to fall into overfitting [106]. It has, therefore, become a popular tool for population migration modelling. In population migration prediction research, random forest is used to solve the regression problem in the following way. First, 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. Next, k decision tree models are constructed for the k samples, and the k regression results are obtained. Finally, the k decision tree results are combined, by taking the average values.

To address the problem of forecasting environmental migration, Best et al. proposed a random forest model that could effectively identify significant variables from large social surveys [18]. Random forests enable the ranking of variables by importance. In regression random forest models, the importance is calculated by the node impurity, which is a calculation of to what extent the variance of the result can be reduced by splitting the decision trees by a particular variable [15]. In this regard, this model can identify the most important predictors of migration from around 2000 original factors, allowing a 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 [107]. The results show that climatic factors have positive and significant effects on the predictive performance of individual migration intentions.
3. Support Vector Machine

The support vector machine (SVM) has mainly been used to solve the classification and regression problems of machine learning, and is suitable for analyzing small samples and multidimensional data [108]. SVM differs from traditional neural network learning methods, and is based on the Vapnik–Chervonenkis dimension theory, and the structural risk minimization principle, which realizes the structural risk minimization principle. It also minimizes empirical risk, has a good generalization performance for future samples, and has many advantages, such as a simple structure, good adaptability, global optima property, fast training speed, and more generalized applications. The SVM-based HM forecasting method uses actual HM data to train the support vector machine model, determine the model parameters (e.g., the insensitivity coefficient, penalty factor, and kernel function parameters), forecast the future state based on the trained SVM model, and obtain the predicted value of HM, by comparing it with the pre-set failure threshold.

Zhang et al. employed the SVM algorithm to create a classification model for the migration of Beijing’s unregistered resident population, and conducted an empirical analysis of migration data from various surveys in Beijing [109]. The results show that SVM is more accurate and generalizable than the basic BP neural network and logical regression for these specific classification tasks, and can forecast migration trends with greater accuracy.

3.3.2. 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, with its powerful feature-extraction capability, provides a solution for training massive data. However, in the field of migration research, due to differences in the relevant concepts, and insufficient human, material, and financial resources, the data collected on population migration are sparse or entirely lacking, leading to an inaccurate and out-of-date understanding of population migration. Social media data offer a new way to expand the timely perception of complete population migration information. With the acceleration of globalization, social media and traditional data complement each other, further meeting the needs of the efficient management of population migration.

1. Recurrent Neural Networks

A recurrent neural network (RNN) is a neural network with short-term memory capability. In an RNN, neurons can both accept information from other neurons, and accept their own information, forming a looped network structure. Compared with a feedforward neural network, an RNN is more consistent with the structure of the biological neural network. In tasks such as language modelling and natural language generation, because of the feedforward and feedback connections between neurons at each layer, RNNs are suitable for processing sequence data with back-and-forth dependency. 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 step-by-step error information, in reverse chronological order. A relatively long input sequence causes gradient explosion and disappearance problems. To solve the RNN problem, some improved structures have been proposed. Among these, gate current unit and long short-term memory (LSTM) are typical representatives that overcome the problems of gradient disappearance and gradient explosion of traditional RNN networks, leading the RNN structure to extract the deep characteristics of the time series, and also consider the long-term dependence of the time series. This allows the model to obtain better prediction results.

Golenvaux et al. used LSTM to predict international migration based on Google Trends data, using the thermal coding vector input tag to incorporate more complex time-invariant factors (such as the distance between two countries, and common language) [110]. 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.

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, a GNN converts the graph structure data into a standardized and standard representation, which it inputs into a variety of different neural networks for training, achieving good results in node classification, link prediction, graph clustering, and other tasks. The actual network can be mapped to the relationship between nodes and edges, and the GNN can be used to generate a graph from the unstructured data. The 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 [78]. This method uses GNNs to consider the potential relationship between large geographical regions, and realizes the prediction of population movement between cities on national spatial granularity. In addition, they 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.

Additionally, the feasibility of the use of Twitter data to predict internal HM has been explored, to address the problem of an insufficient data volume for deep-learning prediction models [76,78]. The results showed that Twitter data have considerable value in HMP. Future studies could focus on the selection of pertinent data, and the design of efficient feature models to further the research on deep-learning-based HMP.

Although ML methods have seen 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. Massive amounts of data are required to fully realize the predictive power of ML, and the current sample size of HM research is far from reaching the lower limit that would allow for accurate predictions. In addition, with the complexity of cross-border HM policy changes, the related models lack robustness, and the generalization of ML models is poor. Once again, neural networks are essentially a black box, and their reliability is difficult to guarantee. To promote the application of neural network methods in population migration prediction, it is possible, first, to use new media data to replace or augment traditional data, in addition to using the idea of migration-learning to reduce the amount of data required for model training. The next step is to take advantage of machine learning’s ability to parse large amounts of text, conduct sentiment analysis to reduce the impact of macroeconomic policies, and improve the robustness and generalizability of the models. Moreover, exploiting the model interpretability further improves predictive reliability; for example, Simini et al. used SHapley Additive exPlanations (SHAP) values to understand how different geographic features play a role in the model [111].

4. 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, definitions of HM statistics differ from country to country. During the collection process, governments take a convenience approach, and tools developed for other purposes, rather than specifically designed to measure HM and its outcomes, are used [112]. Even in developed countries such as EU nations, information on migrating populations remains rudimentary. For instance, Cyprus, Ireland, France, Portugal, and the United Kingdom use surveys. However, Romania and Greece use other data sources, including mirror statistics from other countries. The remaining countries (e.g., Germany) use population registers, partly including registers for foreign citizens. The data do not capture the migration processes well, and are often not comparable in migration study. The quality of the data can be very problematic.
As a result, HM definitions and data quality issues often cause forecasting methods to be incomparable, with inaccurate forecasts, forecasting models that lack robustness, and significant uncertainty [67,112]. 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.

Although 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 rendering a comprehensive explanation of the migration process impossible, even if they cover most cases. Therefore, HM theory has a limited role in the interpretation of the results. Moreover, HM forecasting involves different disciplines, and other experts have entirely different expectations about HM changes. The cumulative effect of all of these uncertainties hampers the development of HMP. Therefore, attention should be given to uncertainty discussions in all HM projections; otherwise, such uncertainties will spread.

5. Conclusions

This paper examined the work on HM forecasting, up to 2022. The results show that HMP is undergoing a rapid development, indicating that the input data have changed from simple stock information to multi-source spatiotemporal data, including spatiotemporal scene characteristics, and individual behavioral characteristics. The data also indicate that the prediction model and algorithm are evolving from a simple linear model to a nonlinear model. Most importantly, the prediction effect is becoming more refined and precise. Currently, HMP can be broadly classified into three types: deterministic, stochastic, and machine-learning HMP. Deterministic HMP estimates the future population migration development by setting model parameters in one or more scenarios; it has a low dependence on data, and can fully utilize the subjective prediction results of experts. The method is simple to operate, suitable for medium- and long-range forecasting, and widely used in many countries and international institutions (e.g., the United Nations, the European Union) [96]. However, the uncertainty of evaluation in the consistent and explicit quantitative sense has led to uncertain interpretations of results; the medium-range solution is the most likely forecasting method, which has also led to errors. Unlike the deterministic model, the stochastic model’s parameters are not fixed, but random variables. The stochastic prediction model describes the influence of driving factors on population migration by a statistical analysis or an assumed probability distribution based on historical data. The uncertainty in population movements is fully considered. As a result, the most probable situation and prediction range can be forecasted. In addition, the model’s prediction ability can be improved by introducing expert knowledge. However, stochastic models are highly dependent on variables and available data. They involve a high degree of subjectivity, and the operational difficulties make them unsuitable for beginners. With the development of big data and artificial intelligence technology, machine learning has enabled great progress in HMP. However, there is a long way to go in its further application. An important reason for this is that most machine-learning methods use social media data, which cannot effectively explain the causality of population migration, although their flexible choice of functional forms to fit data can effectively improve their prediction ability. In addition, data privacy and data protection problems have restricted the machine-learning method’s use in predicting the development of population migration. In the future, more data on population migration are expected to become available, so the model will become more sophisticated. Regardless of how population migration prediction evolves, the data and model algorithms must be reasonably selected, based on clear application scenarios and uncertainty, to achieve better results. The analysis and summary of population displacement projections presented in this paper can help develop and improve the field to some extent.
6. Outlook

Overall, HMP is a multidisciplinary research and application area. However, it has limitations, including a low data quantity and quality, insufficient support of system theory, insufficient innovative methods, and imbalanced research areas. Combined with the developing trend of big data and artificial intelligence in recent years, this paper has established that more steps need to be taken to ensure the future development of HMP.

Firstly, global population growth is gradually slowing. The driving force of population growth in many countries is primarily reliant on HM, which has consequently become an important issue related to the population and political security of some countries. Only by mastering actual HM data can scientific and practical measures be taken to ensure the stable development of economies and societies. To ensure the comparability and uniformity of data, the statistical institutions in 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 regions. There should be free public availability of data, and government agencies, industry, and academia should collaborate, to facilitate rapid research on migration prediction.

Secondly, with the development of big data and AI technology, the number, type, and fineness of data will be continuously improved, thus challenging the ability of prediction models. The dynamic change in HM depends not only on temporal differences, but also on the transformation of spatial characteristics [113]. Traditional model-driven methods are unsuitable for processing spatiotemporal series, because they cannot capture their hidden nonlinear features. As an efficient deep-learning framework based on the graph data structure, GNN is widely used in various fields, and has achieved remarkable results. HM flow data 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.

The development of HMP research will experience peaks and troughs in the application of big data and AI technology. Countries should use the opportunity for growth to solve the talent problem that has been a longstanding issue in HM research. Governments should plan to facilitate the migration of skilled people, rely on those universities and institutions that carry out population research, and work to train reserve talent with an AI technology base for the research and management of HM. In addition, 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, while relevant government agencies have obtained a large amount of stock data in the traditional process of HM management, they have been unable to make full use of these data, due to technical limitations. With a lack of technical support, management methods are often based on personal intentions, and the subjective opinions of managers. With the development of big data and AI technology, the acquisition, preservation, and processing of HM big data has 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 scientific decision-making and social management of HM, to promote the innovation of government management and the social governance mode, and to strengthen research into, and development of, HMP and intelligent auxiliary decision-making systems.

Author Contributions: T.P., C.H., J.Y. and M.H. contributed equally to the authorship of the manuscript, including the research design, conducting the research, performing the analysis, and writing the manuscript. All authors have read and agreed to the published version of the manuscript.

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

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

Sustainability 2023, 15, 10584


69. Shayegh, S.; Emmerling, J.; Tavoni, M. International Migration Projections across Skill Levels in the Shared Socioeconomic Pathways. *Sustainability* 2022, 14, 4757. [CrossRef]


86. Burzynski, M.; Deuster, C.; Doccquier, F. Geography of skills and global inequality. *J. Dev. Econ.* 2020, 142, 102333. [CrossRef]


89. Hanson, G.; McIntosh, C. Is the Mediterranean the new Rio Grande? US and EU immigration pressures in the long run. *J. Econ. Perspect.* 2016, 30, 57–82. [CrossRef]


106. Blau, G.; Scornet, E. A random forest guided tour. TEST 2015, 25, 197–227. [CrossRef]


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.