# Migration Anticipation and Preparedness

## Making Migration Management Work

# Migration Anticipation and Preparedness

## MAKING MIGRATION MANAGEMENT WORK

OECD
BETTER POLICIES FOR BETTER LIVES

This work is published under the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of the Member countries of the OECD.

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

**Please cite this publication as:**
OECD (2026), *Migration Anticipation and Preparedness: Making Migration Management Work*, OECD Publishing, Paris, https://doi.org/10.1787/4161131f-en.

ISBN 978-92-64-64270-6 (print)
ISBN 978-92-64-55962-2 (PDF)
ISBN 978-92-64-85573-1 (HTML)

**Photo credits:** Cover © Summit Art Creations/Shutterstock.com.

Corrigenda to OECD publications may be found at: https://www.oecd.org/en/publications/support/corrigenda.html.

© OECD 2026

**Attribution 4.0 International (CC BY 4.0)**

This work is made available under the Creative Commons Attribution 4.0 International licence. By using this work, you accept to be bound by the terms of this licence (https://creativecommons.org/licenses/by/4.0/).

**Attribution** – you must cite the work.

**Translations** – you must cite the original work, identify changes to the original and add the following text: *In the event of any discrepancy between the original work and the translation, only the text of the original work should be considered valid.*

**Adaptations** – you must cite the original work and add the following text: *This is an adaptation of an original work by the OECD. The opinions expressed and arguments employed in this adaptation should not be reported as representing the official views of the OECD or of its Member countries.*

**Third-party material** – the licence does not apply to third-party material in the work. If using such material, you are responsible for obtaining permission from the third party and for any claims of infringement.

You must not use the OECD logo, visual identity or cover image without express permission or suggest the OECD endorses your use of the work.

Any dispute arising under this licence shall be settled by arbitration in accordance with the Permanent Court of Arbitration (PCA) Arbitration Rules 2012. The seat of arbitration shall be Paris (France). The number of arbitrators shall be one.

# Foreword

As international migration continues to evolve in scale, complexity and unpredictability, governments face increasing pressure to respond swiftly and effectively often under tight operational constraints and high public scrutiny. Anticipation is therefore becoming a core capability of migration governance. At the same time, forecasting migration remains inherently challenging: flows are shaped by rapidly changing geopolitical and economic conditions, policy interventions can alter the very outcomes that are being forecast, and data availability and quality vary widely across migration categories. In this context, forecasting is not about eliminating uncertainty; it is about making uncertainty actionable supporting preparedness, informing choices, and improving the timeliness and transparency of decisions.

This Handbook provides a practical guide to strengthening migration forecasting systems and preparedness strategies across OECD countries. It draws on country practices, state-of-the-art research and the lessons of the OECD's Migration Anticipation and Preparedness (MAP) Task Force. Bringing together perspectives from national institutions and academia, it addresses the questions that policymakers, analysts and practitioners consistently face: what forecasting should be used for, why migration needs to be forecast by category rather than only in aggregate, how to select models and data that fit specific purposes, how to evaluate performance credibly, and how to communicate results to policymakers, especially uncertainty so that forecasts can be used responsibly.

The Handbook is structured around nine chapters that follow the forecasting lifecycle from problem definition to communication and system learning. After the introduction and general overview, Chapter 3 clarifies the policy question that forecasting is expected to answer and the main purposes forecasts can serve. Chapter 4 discusses how to choose models that fit policymakers' needs, migration categories and operational constraints, including questions of model complexity and update cycles. Chapter 5 focusses on data foundations and system requirements, including ways to assess data sources, incorporate qualitative information, and operationalise policy indicators. Chapter 6 addresses model development, covering traditional time-series approaches, Bayesian modelling, structured integration of expert judgement (including Delphi surveys), and the role of machine learning and causal methods. Chapter 7 sets out how models can be evaluated, validated and adapted through back-testing, calibration and robustness checks. Chapter 8 focusses on communication and how to present forecasts and uncertainty to policymakers, and how to support effective use in practice. Finally, Chapter 9 looks ahead to emerging developments in migration forecasting, including new approaches to uncertainty quantification and the growing importance of interactions across categories and countries.

This project was financially supported by the Directorate General for foreign nationals in France (DGEF), Ministry of the Interior, France, whose contribution made this Handbook possible. The Handbook was edited by Yves Breem (DGEF, French Ministry of the Interior) and Taehoon Lee (OECD). It was written by Jakub Bijak (University of Southampton), Yves Breem, Marcello Carammia (University of Catania), Stefano M. Iacus (Joint Research Centre; formerly Harvard University), and Taehoon Lee, in alphabetical order. The Handbook also benefited from presentations, participation and written comments provided by members of the OECD Migration Anticipation and Preparedness (MAP) Task Force. In addition, it benefited

from comments by Jean-Christophe Dumont and Jonathan Chaloff (OECD) and Guillaume Mordant (DGEF, French Ministry of the Interior).

With practical tools, decision frameworks and real-world examples, this Handbook aims to support countries in building forecasting systems that are fit for purpose, resilient to changing conditions and connected to decision making. By strengthening the link between forecasting and preparedness, it seeks to help governments respond more effectively to uncertainty grounded in evidence, informed judgement and continuous learning.

# Table of contents

Foreword 3

Executive summary 8

1 Introduction 10
For whom is this handbook? 11
Forecasting for migration flows by categories versus (net) migration as a whole? 12
References 13

2 Current state of play 14
Where do OECD countries stand on migration forecasting? 14
Notes 15

3 What is the question to which forecasting is expected to respond? 16
What is the purpose of migration forecasting, and what are the associated challenges? What are policymakers looking for through forecasting? 16
References 19

4 Forecasters' toolkit: Choosing the right model for the task 20
What are the existing forecasting methods and models? 20
What is the appropriate forecasting model for each migration category? 24
What resources do countries need to run forecasting? 29
How complex should the model be? 31
How often should model results be updated? 32
References 33

5 What data are necessary to conduct the forecasting exercise? 37
How to assess the existing sources of data? 37
Where to find and how to incorporate qualitative data? 41
How to quantify policy indicators for statistical models? 44
Which software is the best fits for forecasting, causal effect estimation and projection? 47
References 49
Notes 52

6 How should the forecasting model be developed? 53
Is it still accurate to use traditional statistical and econometric time series analysis? 53
What are Bayesian Models, their principles, and key features? 55
How to include contextual knowledge in forecasting: Expert opinion and Delphi surveys 56

How to run a Delphi survey and how to include expert replies in the model? 57
How to develop machine learning models? 60
How to incorporate migration drivers into the machine learning models? 61
How to model causal influence in migration forecasting? 62
References 64

7 How can the model be evaluated? 67
How to validate model performance? 67
How to calibrate the model? 69
How to assess robustness? 71
How to back-test ML/Stat models? 72
How can expert knowledge and scenarios be back-tested? 72
References 74

8 How should forecasts be presented to policymakers and ensure effective interaction between policymaking and forecasting? 76
How to present forecasts and their inherent uncertainty to policymakers 76
How to use the forecast properly and effectively 82
References 84
Notes 85

9 What future developments can be expected in migration forecasting? 86
How to measure the potential effect of policies in countries of destination? 86
Is conformal prediction the final frontier in time series forecasting? 87
How can forecasting results in one migration category affect forecasting in other categories? 88
How can forecasts in one country affect forecasts in other countries (i.e. interaction across countries)? 89
References 90

Annex A. List of Migration Anticipation and Preparedness (MAP) task force members 92

Glossary 93

## FIGURES

Figure 4.1. Decision tree to identify the best migration forecasting models according to national practitioners' needs and resources 25
Figure 5.1. Decision tree to identify the best data according to migration forecasting models 37
Figure 6.1. A hypothetical DAG illustrating a sequence of causal relationships 63
Figure 8.1. Forecasts of total immigration to the United Kingdom, averaged univariate models (in thousands) 1975-2060 77
Figure 8.2. Immigration of EU Citizens to the United Kingdom with different probability intervals, 1975-2013 78
Figure 8.3. Probability Density Function for Prognosis – Forecasted first asylum applications 2024 in the Netherlands 79
Figure 8.4. Cumulative probability for the Dutch number of first asylum applications in 2024 80

## TABLES

Table 2.1. Overview of the current forecasting practices in OECD countries 15
Table 3.1. National institutions in OECD countries currently engaged in forecasting activities 17

Table 4.1. Stylised characteristics of seven main types of migration flows, typical data, and recommendations for selecting possible forecasting methods 27
Table 4.2. What models are currently used in OECD countries? 28
Table 5.1. What data sources are currently used in OECD countries and by international organisations and agencies to forecast migration flows? 39

# Executive summary

## Why predicting migration is hard – and yet necessary

Migration is one of the most complex social phenomena to anticipate. People move for deeply interconnected reasons: economic opportunity, political instability, safety concerns, education, family reasons, and personal aspirations. These drivers interact with policy changes, geopolitical crises, environmental shocks, and shifting labour markets. Sometimes flows change gradually; at other times, they surge abruptly in response to conflict or political events. The refugee movements following the Syrian conflict in 2015-2016 and Russia's large-scale invasion of Ukraine in 2022 revealed how difficult it is for governments to anticipate large-scale displacement.

Yet despite this inherent uncertainty, governments must attempt to forecast migration. Without forward-looking estimates, it becomes nearly impossible to properly allocate budgets, plan housing and integration services, manage asylum systems, or adjust labour market policies. Forecasting supports contingency planning, strengthens crisis preparedness, and enables more transparent public communication. It also facilitates international co-operation and burden-sharing by identifying potential pressure points before they escalate.

The OECD's Migration Anticipation and Preparedness (MAP) Task Force developed this Handbook to guide policymakers, civil servants, analysts, and modelers in building more robust forecasting systems. The document emphasises that forecasting is not about achieving perfect prediction. Rather, it is about giving the right tools to reduce uncertainty, prepare for multiple plausible scenarios, and embed anticipation into migration governance.

## One size does not fit all: Tailoring forecasts to migration categories

A central argument of the Handbook is that migration cannot be forecast as a single aggregate number. While demographic projections often rely on net migration figures, such aggregates are insufficient for operational policymaking. Effective governance requires understanding who is migrating, under which category, and why.

Different migration streams respond to different drivers and exhibit distinct statistical characteristics. So-called forced migration flows, such as asylum applications and irregular border crossings, are highly volatile and often non-stationary. They can shift rapidly in response to conflict, disasters, or policy changes. In contrast, labour, student, and family migration may follow more stable and structured patterns, often influenced by economic conditions or institutional frameworks.

Because of these differences, forecasting must be category specific. Time-series models such as ARIMA may work well for relatively stable flows. Econometric models can help analyse structured labour or student migration patterns, although they often struggle to anticipate sudden breaks. Bayesian methods help to improve forecasting results by incorporating prior knowledge (including expert opinion) and uncertainty.

More recently, machine learning techniques – including neural networks models – have shown promise for short-term forecasts of high-frequency data, particularly for asylum and irregular migration.

Across OECD countries, forecasting practices vary widely. Many countries regularly forecast asylum applications, but far fewer predict irregular border crossings or regulated migration streams such as labour or student migration. Most systems combine quantitative modelling with expert judgement, acknowledging that statistical models alone cannot capture sudden policy shifts or geopolitical developments. The appropriate level of model complexity depends on the volatility of the flow, the time horizon, data availability, and institutional capacity. In many cases, simpler models (whether augmented by expert elicitation or not) may perform nearly as well as more complex ones while remaining easier to maintain and communicate.

## From models to preparedness: Data, capacity, and continuous updating

Reliable data are the foundation of migration forecasting, yet they present significant challenges. Migration statistics often suffer from delays in publication, inconsistent definitions, limited frequency, and discontinuities over time. The Handbook proposes evaluating data sources according to criteria such as accuracy, timeliness, coverage, granularity, continuity, and clarity of definitions.

A key trade-off emerges between timeliness and accuracy. Highly validated data may arrive too late for operational use, while real-time or high-frequency indicators may lack full reliability. Increasingly, forecasters complement administrative data with alternative “digital trace” sources such as online search data, social media signals, mobile phone data, and event databases. These innovative sources can enhance short-term forecasting, particularly in more volatile contexts.

Beyond methodology and data, institutional capacity is critical. Effective forecasting requires secure IT infrastructure, computing resources, skilled analysts, and cross-ministerial co-ordination. Forecasting systems must be embedded within governance structures, supported by clear mandates and long-term political commitment. Data-sharing agreements and interagency collaboration are essential to ensure sustainability.

Forecasts must also be regularly updated. Migration systems are dynamic, and models must adapt to new events. Medium-term forecasts for a few years often require annual recalibration, while early warning systems based on high-frequency data and forecasting for a few months may need monthly revisions. Major geopolitical or policy changes may require immediate adjustments outside regular update cycles.

Ultimately, the Handbook frames migration forecasting as a structured approach to managing uncertainty. Perfect prediction is neither possible nor the objective. Interacting regularly with policymakers to acknowledge that matter of fact, alongside disseminating forecasting results to the right network of decision makers on the ground, is key to make sure forecasting, despite its uncertainty, is relevant. By tailoring models to specific migration categories, investing in high-quality data and institutional capacity, combining quantitative analysis with expert insight, and maintaining continuous updates, governments can significantly strengthen their preparedness. Forecasting thus becomes not merely a technical exercise, but a central pillar of evidence-based and resilient migration governance.

# 1 Introduction

The multidimensional factors that drive migration trends are themselves difficult to predict. People migrate for a variety of interconnected reasons, which can include difficult-to-quantify considerations such as perceptions of safety or political freedom, changes in the economic situation at origin or destination, as well as personal reasons that cannot be easily captured. It is very difficult to look into the decision making of individuals before they move to pursue education, to start or unite with family, or to pursue individual aspirations. Migration also responds to policy changes in destination countries, and even in other countries as potential migration routes open or close. External factors may have an immediate impact on migration trends, while other times the impact may lag. Sometimes migration flows remain unchanged until a specific trigger point, while in other cases, emerging trends generate visible snowball effects. This extreme complexity makes predicting future flows particularly challenging.

Policymakers must still make an effort to anticipate migration flows, despite these difficulties. They are expected not only to exercise effective governance of migration but also to respond quickly and effectively to sudden changes. Designing effective policies for the future cannot be done without an attempt to understand what future flows might look like.

Forced migration (e.g. asylum, irregular migration, and movements combining elements of irregular and regular migration, such as mixed migration) in particular requires governments to conduct robust contingency planning and to be ready to rapidly adapt asylum and reception services. While regulated migration is by nature easier to predict – and less likely to catch policymakers off guard –, forecasting exercises are nonetheless important for programming and budgeting. Forecasts provide valuable insights for improved programming, financing, anticipating labour market needs, and making operational decisions, including staffing, integration support, facilities planning, establishing numerical limits or quotas coherent with policy objectives, and more. Improved anticipation of future migration trends also allows policymakers to inform public debate, explain choices and justify the costs of proactive investments and actions.

The track record for migration forecasting is improving, especially after several notable missed scenarios in recent memory. For example, the 2015/16 refugee surge from Syria in Europe and Russia's war of aggression against Ukraine in February 2022 found most destination countries off guard. Looking back, however, visible early warning signals of mass displacement were missed. It is not just forced migration that reveals examples of missed signals: legal migration has been expanding recently to most OECD countries, sometimes putting unexpected pressure on housing or social services.

Enormous progress has been made in the past decade. Recognising the need for better forecasting, countries and international organisations or agencies have developed tools to predict short-term forced migration movements and assess migration uncertainties and risks. The EU, in particular, has established a robust early warning framework, which includes the Integrated Political Crisis Response (IPRC) Blueprint network, the well-established Frontex Common Integrated Risk Analysis Model (CIRAM), the European Union Asylum Agency (EUAA) Early Warning and Forecasting approach, and the most recent Situational Awareness, Early Warning and Forecasting Capacity Development Project, from the EU Joint Research Centre (JRC) in support of the implementation of the EU Pact on Migration and Asylum. Additionally, various national forecasting approaches in countries such as the Netherlands, Sweden, Switzerland and the United States have been developed over the last two decades.

Forecasting at first was at most notional or mechanical, relying on migration theories and causal (gravitational) models unable to fully capture the complexity of migration phenomenon. Today, it has increased in sophistication along with the development of the discipline, the massive expansion in computational capacity and the increasing availability and use of big data. However, these new forecasting efforts remain scattered, hard to generalise and narrowly focussed on short term effects, or even just on nowcasting – even with the right information systems in place, useful predictions are currently limited to a 12- to 18-month time horizon. Moreover, for obvious reasons, most progress has been made on forced displacement (e.g. asylum applications). Regulated migration flows – such as family, work, and student migration – have generally not been the object of forecasting. Finally, much of the progress in modelling and data use has been occurring behind the scenes, for internal use by governments which understandably avoid publicising not only their data sources and the scenarios but often the models themselves.

The need for more, and better, forecasting, is clear. The request often comes down to civil servants and to modelers with shorter timelines and high expectations. There are no models that can be taken directly from the shelf and applied.

To provide more clarity, the OECD, with the financial support from the General Directorate for Foreign Nationals in France (DGEF), established the MAP (Migration Anticipation and Preparedness) task force. It aims to improve the anticipation of and enhance forecasting capabilities for both forced migration movements and regulated migration flows. The MAP Task Force complements other collaborative efforts cited above. The Task Force focusses on forecasting and does not engage in nowcasting (see Box 4.3), early warning, or foresight exercises. This Handbook is meant to support the growing demand for more timely and robust migration forecasting by providing practical guidance.

## For whom is this handbook?

Migration forecasting involves many actors – the politicians and policymakers who ask for forecasts, the civil servants who must figure out how to produce forecasts, the modellers who must identify and secure data and choose and apply methodologies, the officials responsible for explaining the results of the forecasts, and the policymakers, politicians and communication teams who make use of them.

This handbook attempts to provide a guide for all these actors. In particular, it has been drafted primarily with five categories of users in mind:

- The civil servant, policymaker or official who has been tasked with producing forecasts, who needs to understand the state of the art of forecasting and how to meet the request with the resources at hand
- The modeller who has experience with forecasting in different domains but is relatively new to migration forecasting, who must identify data sets and understand specificities of migration
- The migration data analyst with little to no experience in time series modelling, who needs to better understand the different approaches and tools necessary for forecasting migration
- The migration forecaster who has been asked to improve the accuracy of their forecasts using new techniques or drawing on developing practices
- The civil servant, policymaker or official who receives the output of the modeller and must translate it into a form that can be used correctly, with all the uncertainties inherent to forecasting.

In some cases, the user of this Handbook may find themselves playing more than one role – both modeller and communicator, for example, or both commissioner and communicator. Policymakers who are considering whether and how to establish a forecasting approach may also draw on this Handbook in order to understand better the need for – and limits of – forecasting in migration.

The Handbook is structured into questions, which can be grouped into four general categories. First, what is migration forecasting, why is it useful and how is it currently being used? Second, how can a forecast be conducted, in terms of models, resources, data, development and evaluation? Third, how can forecasts be presented to policymakers so that they understand the complex results and the uncertainties and make proper use of the forecast? Finally, the Handbook points to some emerging questions in forecasting that may help policymakers and modellers alike to understand how the discipline may evolve in the coming years.

## Forecasting for migration flows by categories versus (net) migration as a whole?

Deterministic scenarios of international migration are used by the vast majority of the statistical institutes in the developed countries notably to estimate population forecast. These efforts most often focus on net migration (immigration of both national and foreigners minus emigration of both groups) offering only aggregate trends without disaggregation by category. These are usually based on past demographic trends and does not attempt to predict the impact of policies or economic changes or of other factors potentially affecting migration.

Recently, as a result of rapidly increasing asylum applications, some efforts have been made to better anticipate specifically forced migration. The very nature of this migration component requires taking into account specific factors that go beyond demographic variables and are linked notably to the situation in countries of origin and transit. This has led to new forms of modelling that can not be directly applied to forecast other categories of migration. Several key factors explain why forecasting needs to distinguish between different migration categories.

First, different migration flows respond to different driving forces and show very different structural characteristics. For example, forced displacement due to conflict or natural disasters tends to be highly volatile and non-stationary. In contrast, labour migration between two high-income countries often follows more regular, predictable patterns (Bijak et al., 2019[1]). Forecasting models must reflect these differences. Time-series models that assume stationarity are unsuitable for forecasting highly volatile flows such as asylum applications; while applying a non-stationary model to more stable, rule-based labour migration may lead to overestimated variability and uncertainty.

Second, data availability and structure vary significantly across migration categories. Very high frequency data (daily or weekly) are typically available on asylum applications or irregular border crossings, while other administrative data based on permits are at best available monthly or just annually. The level of information attached to the migration headcounts also varies, with more or less details on socio economic profiles of migrants. These differences affect the choice of models and the robustness of forecasts. Selecting an appropriate forecasting approach requires an understanding of the scope, granularity, and quality of available data.

Third, while total net migration projections – which balance immigration and emigration figures – can be informative for long-term demographic planning, they offer limited utility for migration policymaking. Effective migration management depends on knowing *who* is migrating and *why*, not just *how many*. For example, preparing integration services, setting visa quotas, or adjusting labour market programmes all require category-specific insights. Forecasting future changes of status (for example transitions from study permits to work permits) might also require specific models (see Chapter 9 for a discussion on category interactions). Aggregated figures, while useful for demographic purposes are often too coarse to guide operational or strategic decisions. Tailoring models to the specific features of each migration category is therefore essential for producing meaningful and actionable forecasts.

## References

Bijak, J. et al. (2019), “Assessing time series models for forecasting international migration: Lessons from the United Kingdom”, *Journal of Forecasting*, Vol. 38/5, pp. 470-487, https://doi.org/10.1002/for.2576. [1]

# 2 Current state of play

## Where do OECD countries stand on migration forecasting?

16 OECD Member countries have actively participated in the Migration Anticipation and Preparedness (MAP) Task Force. The current overview of forecasting practices is mainly based on responses to a questionnaire distributed to these countries, supplemented by selected examples from other OECD Members. Table 2.1 provides a snapshot of the current forecasting practice, detailing which countries are engaged in forecasting and in which migration areas. At least 13 OECD countries are conducting regular forecasting exercises to predict asylum applications. Despite growing interest on the policy agenda in many OECD countries, forecasting irregular border crossings remains underdeveloped. To date, only Sweden and the United States provide such predictions. Most EU countries, however, use the Frontex Common Integrated Risk Assessment Model (CIRAM) for anticipating irregular migration, but it does not estimate future flows directly.

Eight countries forecast labour migration on an annual basis, while five do so for international students and only four for family migrants. Sweden and the United States are conducting regular forecasting exercises for all categories and Norway for all but irregular border crossings. The United Kingdom has also done forecasting for almost all categories but as one-off exercises. Canada and France are considering expanding their routine forecasting work to all categories, while Japan anticipate doing this only for regular migration categories.

Additional information drawn from an Ad Hoc Query launched by Finland in 2024 on “forecasting methods of future migration trends” to which 21 EU member states have responded including 12 OECD Member or candidate countries (BEL, BUL, HRV, CZE, EST, FIN, HUN, LVA, LTU, PRT, SVK, SVN).[1] From this list the majority of countries, except Bulgaria, Hungary, Latvia and Lithuania make use of methods or models for predicting or analysing future migration trends but most of them are only focussing on Asylum application. Czechia, Slovenia and the Slovak Republic use Frontex’s CIRAM (Common Integrated Risk Analysis Model) 2.1 or 3.0 to analyse migration risks.[2] Belgium has developed its own model. No country reported modelling regulated migration flows such as study, labour or family migration on a regular basis.

## Table 2.1. Overview of the current forecasting practices in OECD countries

| Migration Category | Currently conducting on a regular basis | Conducted in the past as a one-off exercise | Considering conducting |
|---|---|---|---|
| Irregular Border Crossings | Sweden, United States (MAC) | Canada[1], Germany, United Kingdom[2] | Canada, France, Germany, and Greece |
| Asylum Applications | Austria, Belgium, Canada, Finland, France, Germany, Ireland, Netherlands, Norway, Poland, Sweden, Switzerland, United States | United Kingdom | |
| Family Migration | Canada, Norway, Sweden, the United States (CIS) | France, United Kingdom | France, Japan, Poland |
| International Student Mobility | Canada, Finland, Norway, Sweden, United States (CIS) | France, United Kingdom | France, Japan, Poland |
| Labour Migration | Australia, Canada, Finland, Korea[3], Norway, Poland, Sweden, United States (CIS) | France, Japan, Türkiye, United Kingdom | France, Japan, Türkiye |

1. Only irregular border crossings followed by an asylum application. Canada developed forecasts for asylum claims across provinces, broken down by claim type; Irregular claims by irregular border crossers were included as one of the categories.
2. Climate migration.
3. Instead of conducting a forecasting exercise on migration flows per se (see Box 8.2), the Korea Immigration Service (KIS) has implemented the Employment Visa Quota Pre-Announcement System since 2024. This system pre-announces the total number (quota) of work visas available for both skilled and low-skilled foreign workers, designed to consider labour market forecasts and the impact of the foreign workers on the domestic labour market. The KIS plans to announce visa quotas annually through this system.
Note: Ireland and the United States have also conducted literature review exercises on asylum applications. Japan, Türkiye, and the United States have undertaken similar efforts focussed on labour migration.

## Notes

[1] https://emnbelgium.be/publication/ad-hoc-query-forecasting-methods-future-migration-trends.

[2] https://www.frontex.europa.eu/what-we-do/monitoring-and-risk-analysis/ciram/.

# 3 What is the question to which forecasting is expected to respond?

## What is the purpose of migration forecasting, and what are the associated challenges? What are policymakers looking for through forecasting?

Accurate migration forecasting is critical for effective policymaking, as it provides essential insights into the volume, composition and timing of migration flows. It plays a crucial role in guiding policy decisions, resource allocation, and preparedness for potential migration surges, including contingency planning and possible counter measures.

The key purposes of migration forecasting include:

1. **Guiding Budget and Resource Planning**

   Forecasts help governments anticipate the number and profile and needs of incoming migrants, enabling precise allocation of budgets for housing, health and education services, as well as integration support. They also inform staffing needs for border management and visa requests, processing of asylum applications, and provision of relevant integration services including language course. Additionally, they inform financing needs related to allowances for asylum seekers, housing, and other forms of integration support more broadly.

2. **Supporting Contingency Planning and Crisis Preparedness**

   Forecasting enables authorities to prepare for the unexpected by developing contingency plans or by taking proactive measures to minimise the risk of certain scenarios. This is particularly important when countries are at risk of instrumentalisation of migration. This is also relevant for countries experiencing transformative policy changes with unexpected consequences on regular and irregular migration (e.g. change in membership to the EU free mobility area).

3. **Informing Fee-Based Migration Schemes**

   A number of countries use visa fees to finance their migration and integration activities. One extreme example is the United States where the budget of the Department of Homeland Security is directly linked to the number of entries in the current year. This may be the case as well for a number of specialised agencies (e.g. for international students) or specific support services that are fee based (For example, on recognition of foreign qualification). In such case, it is important to anticipate future flows for budgetary purpose or to shape future recruitment strategies (see Box 3.1 for an example).

4. **Enhancing Public Communication**

   Migration forecasting also supports informed public debate by grounding migration discussions in data. By sharing forecast-based expectations and planned responses, authorities can demonstrate preparedness, reduce misinformation, and help frame migration governance as manageable rather than a crisis mode. This contributes to reinforcing public trust and strengthening political stability.

### 5. Strengthening International Co-operation and Burden Sharing

Forecasts play an important role in co-ordinating responses across destination countries but also with countries of origin and transit. Forecasting helps anticipate pressure points and enables fair burden sharing mechanisms such as relocation or financial contributions (for example, the Pact on Migration and Asylum in the EU context). At the global level, better forecasting can also contribute to addressing the root causes of forced displacement.

### Box 3.1. Classifying potential countries for future student migration

Another study conducted in the United Kingdom (British Council, 2024[1]) used econometric modelling not to directly forecast flows, but to classify potential of origin countries for future international student migration. The results indicated that, at the individual country level, GDP growth and the expansion of the middle class were strongly correlated with, and statistically significant in explaining, the variation in outbound student volumes by countries of origin over time. Additionally, several of the destination-weighted exchange rates calculated in the research demonstrated a striking year-on-year correlation with outbound student growth, particularly in more price-sensitive markets such as Nepal, Nigeria, and Viet Nam.[1] Based on these findings, prospects for student mobility were assessed according to the "growth potential" and "risk profile" of sending countries. The study also developed an index classifying future sending countries up to 2030, aimed at helping universities shape their international recruitment strategies.

1. A prime example is the collapse of the Nigerian Naira, first in mid-2023 and then again in early 2024, which effectively priced many Nigerian students out of overseas higher education markets.

Various national institutions across OECD countries are involved in migration forecasting to support policymaking and operational planning (see Table 3.1).

### Table 3.1. National institutions in OECD countries currently engaged in forecasting activities

| Country | National institution | Purpose |
|---|---|---|
| Australia | Immigration Planning and Policy Framework Branch, Department of Home Affairs | Provide policy advice to the Australian Government on the optimal size and composition of the permanent Migration Program |
| Austria | Asylum and migration affairs, Federal Ministry of the Interior | |
| Belgium | Federal Ministry of Interior | Forecasts are discussed within the framework of the monitoring of the asylum chain (co-ordinated by the cabinet for asylum and Migration). It determines the necessary housing capacity, and the workload of the Commissioner General for Refugees and Stateless Persons (CGRS) and Immigration Office. |
| Canada | Immigration, Refugees and Citizenship Canada (IRCC) | IRCC develops demand forecasts across its main business lines, including Asylum, Temporary Residents (visitors, workers, and students), Biometric Enrolments, Permanent Resident admission, Permanent Resident cards, Citizenship Grants and Proofs, and Passports. These forecasts are actively used to support departmental planning and decision making, including policy planning, budget planning, operational capacity, staffing, intake management, and levels planning. Forecast outputs inform workload and staffing requirements, intake management strategies and service delivery models. They also support levels and volumes planning by informing the development of Canada's annual Immigration Levels Plan and associated multi- |

| Country | National institution | Purpose |
|---|---|---|
| | | year planning assumptions. IRCC's forecasts are embedded in the department's and the government of Canada's planning cycles. Migration forecasts have implications beyond IRCC, affecting planning in areas such as housing, healthcare, education, labour markets, and central agency oversight. As such, there is an emphasis on translating forecasts into clear, decision-ready products that can be shared across sectors for co-ordinated planning. The forecasts are reviewed and approved through governance committees from the working level to the Deputy Minister level and shared interdepartmentally through Canada's Interdepartmental Working Group on Migration-Related Forecasting. This ensures alignment between anticipated demand, available resources, and policy priorities. |
| Canada (Quebec) | Ministry of Immigration, Francisation and Integration (MIFI) | Permanent immigration planning. Forecasting supports planning for temporary immigration and accommodation capacity. |
| Finland | Analysis and Foresight Team, Finnish Immigration Services (FIS), Ministry of Interior | Supports operational management and financial planning of FIS. |
| France | Migration Analysis National Hub (PONAM), Office for the Struggle against Migrant Smuggling (OLTIM), Central Directorate of the French Border Police (DCPAF) | Risk assessments to adapt border control staff allocation and operating procedures. |
| France | *Département des Statistiques, des Etudes et de la Documentation* (DSED)/DGEF, Ministry of the Interior | Better budgeting, planning and managing resources (staff, infrastructure, housing facilities and number of beds in reception centres) and quickly adapt migration policies to future migration shocks. |
| France | Office for the Protection of Refugees and Stateless Persons (OFPRA) | Planning interviews and translation services, reallocating staff based on changing migrant profiles. |
| Germany | Federal Office for Migration and Refugees (Bamf) | Planning resources in the asylum system and supporting international policy actions. |
| Ireland | Ministry of Justice | Planning for migration-related requirements, including accommodation and service capacity. |
| Japan | Japan International Co-operation Agency (JICA) | Investigating labour migration projection to support migrant worker programmes. |
| Korea | Seoul National University, commissioned by the Korean Immigration Service (KIS) of the Ministry of Justice | Planning the Employment visa quota pre-announcement system. |
| Netherlands | Dutch Ministry of Justice | Adapting budgets and staffing throughout the asylum process. |
| Norway | Norwegian Directorate of Immigration (UDI) | Forecasts are used across institutions to adapt the state budget and better manage housing and caseloads. |
| Poland | Migration Analysis and Statistics Division, Office for Foreigners (UDSC) | Forecasts are prepared for external needs, particularly related to asylum and migration processing. |
| Sweden | Planning Department, National Operations Department, Swedish Migration Agency (SMA) | Prognosis papers used for planning, including contingency planning and budget forecasting for migration services. SMA interacts with policymakers every two weeks. |
| Switzerland | State Secretariat for Migration | Forecasts to support planning for finances, staffing, and accommodations. |
| Türkiye | Directorate for International Labour Force (DGILF), Ministry of Labour and Social Security | Forecasting exercises help define policies, legislation, and co-ordination arrangements based on future migration scenarios. |
| United Kingdom | Centre for Economic Performance, London School of Economics and Political Science | |
| United States | Migration Analysis Center, Office of Homeland Security Statistics, Department of Homeland Security (MAC DHS OHSS) | Forecasting for planning purposes and to provide decision support. |
| United States | Citizenship and Immigration Services, Department of Homeland Security (CIS DHS) | Predicting incoming workloads to determine future budgets and staffing needs for processing applications. |

Source: Based on a questionnaire to MAP participants, EMN Ad-Hoc Query on 2024.1 Forecasting methods of future migration trends, https://home-affairs.ec.europa.eu/document/download/af40f7e5-d819-442b-b3ea-f5142e033c2c_en, and additional research.

## References

British Council (2024), *The outlook for international student mobility: amidst a changing global macroeconomic landscape*. [1]

# 4 Forecasters' toolkit: Choosing the right model for the task

## What are the existing forecasting methods and models?

There are various approaches to migration forecasting, both quantitative and qualitative, each with distinct methodological assumptions, strengths and shortcomings. This chapter presents an overview of key forecasting methods, which serve as the methodological foundation for models featured throughout this handbook. For comprehensive overviews, see Bijak (2011[1]), Sardoschau (2020[2]) Wiśniowski (2021[3]), Disney et al., (2015[4]), Bijak et al (2019[5])). For a detailed description of the terms used in this publication, please refer to the glossary.

**Judgemental scenarios or deterministic models** are commonly used in demographic forecasting to outline plausible future trajectories for population components – fertility, mortality, and migration. They can also be used, for example to anticipate future family migration based on simple chain migration models. These approaches are typically grounded in past demographic trends and can be informed further by qualitative insights in case of major institutional changes (e.g. 2004 EU enlargement). They produce baseline, high and low trajectories, but these are not probabilistic forecasts (see Box 4.2 for a recent example of demographic analysis based on probabilistic projections of migration). These methods generally do not account for unforeseen changes in government policy, economic shocks, or environmental disruptions, and they do not explicitly model uncertainty.

**Expert-based forecasts** incorporate qualitative insights on future developments – migration policies, economic trends, environmental changes, and technological advances – that may not be reflected in historical data (see for example Lutz and Scherbov (1998[6]) or Sohst, Acostamadiedo and Tjaden (2023[7]) see also Box 4.1). Expert judgement may be provided directly by the forecasters or elicited through structured expert surveys such as the Delphi method (see Chapter 6 and Box 6.2). The Delphi method facilitates iterative feedback and consensus building in an anonymous setting. While useful, such methods can suffer from overconfidence and a tendency to understate uncertainty – especially when trends have been relatively stable.

### Box 4.1. Expert-based asylum forecasts: The example of Switzerland

Switzerland's (State Secretariat for Migration) has had a migration analysis unit specialising in asylum applications forecasts since 2005. Since asylum migration can change almost overnight by events that are nearly impossible to predict (COVID-19, Araspring, etc.), Switzerland chose to establish migration flow scenarios based on expert judgment without using data models. Scenarios are built from six broad factors:

1. The situation in countries of origin (e.g. conflict, human rights violation, economic pressure to migrate) is collected by migration analysts through a survey. Country-of-origin experts from the SEM are asked to evaluate the likely evolution of flows to Europe. Further reports of Migration liaison officers stationed in Swiss embassies are taken into consideration;
2. The situation on migration routes (e.g. quantitative information on irregular border crossings, migration control measures taken at the EU external border, legal travel options) are gathered from countries and other international stakeholders;
3. The economic situation in EU countries of first entry vis-à-vis Switzerland;
4. Measures taken by European countries in terms of migration and asylum policy;
5. Measures taken by Switzerland in terms of migration and asylum policy;
6. Future Swiss implementation of the Dublin procedure and the European Pact on Migration and Asylum.

These six factors are given a different "weight" e by migration analysts. At the beginning of each year, based on all these information inputs, the migration analysis unit estimates four scenarios. The experts assign different probabilities of occurrence, as well as the main nationalities of origin of applicants. This yearly prognosis and its scenarios are then re-evaluated by mid-year.

**Time series models**, particularly autoregressive integrated moving average (ARIMA) models, extrapolate past migration trends into the future. These models assume that future values depend on past values, a drift term, and an error component (see for example de Beer, J. (1997[8]) or Keilman et al. (2001[9])). For example, a random walk with drift (ARIMA(0,1,0)) is often used to model migration trends: $\ln(m_{t+1}) = c + \ln(m_t) + \varepsilon_t$, where the logarithm of predicted migration, $m_{t+1}$, depends on the value from the preceding period, $m_t$, a drift constant c, and an error term, $\varepsilon_t$, usually assumed to follow independent normal centred distributions. Multivariate extensions, such as ARIMAX or Vector Auto-Regressive (VAR) models allow inclusion of additional interdependent variables such as GDP growth or unemployment rates. Time series models are attractive due to their well-established statistical foundations and available software tools. However, they are mostly extrapolative and cannot easily incorporate policy or structural changes unless pre-specified.

**Expert-based probabilistic population projections** have the form where the migration flow under study, $v_t$ is composed of the average trajectory of the process, $\bar{v}$, which is an a priori assumption taken from the subjective expert judgment and the error term, $\varepsilon_t$, following the chosen stochastic process (see for example Lutz et al. (2000[10])). Using expert elicited opinions, stochastic processes are then run to generate a range of possible future population scenarios. Each simulation accounts for random variations in demographic rates, reflecting the inherent uncertainty in population dynamics. These forecasts provide a full probability distribution of future population sizes and structures, rather than a single deterministic prediction. Such models are especially limited by their lack of empirical grounding, though adjustments may be made when structural breaks (e.g. EU enlargement) are anticipated.

**Agent based models** simulate individual behaviours and decisions within a system of interacting agents – migrants, institutions, or other actors. Agents respond to information (e.g. wage differentials) and make decisions stochastically in an imperfect information environment (see for example Willekens (2019[11]) and Klabunde and Willekens (2016[12]) or McAlpine et al. (2020[13])). These models allow for rich causal modelling of migration dynamics, life-course trajectories, and interactions between micro- and macro-level factors. ABMs can incorporate feedback between individual decisions and institutional responses, offering a powerful lens for causal and scenario-based forecasting. The main drawback of these models is that they rely on extensive hypotheses and parameters that are hard to estimate and have a large impact on outcomes. They are usually more useful to identify transmission channels to dependent variables in complex interactive systems rather than produce reliable forecast.

**Econometric or causal models**, such as logit or probit models, estimate migration as a function of explanatory variables, enabling theory testing and policy analysis. Migration theories are however at this stage too fragmented and unsettled to support forecasting.

Gravity Models are among the most widely used causal type of models (see Beine et al. (2015[14]) for a comprehensive overview). They are based on distance, previous migration waves as well as origin destination macroeconomic characteristics and linkages. They explain past migration flows with a satisfactory degree of accuracy but usually fail to anticipate breaks in migration trends and are a poor fit for forced displacement. Multilateral resistance in gravity models makes it possible account for alternative destinations, which may be particularly relevant for some categories of migration where competition between destination countries is strong (e.g. international students or highly skilled migration). Recent work (Beyer, Schewe and Lotze-Campen (2022[15]), Welch and Raftery (2022[16])), however, show that gravity models can neither explain nor correctly predict the temporal dynamics of human movement.

Another example of econometric models used in migration forecasting is Simultaneous Equations Models (SEMs) which model multiple interdependent relationships between migration and other socio-economic factors, offering richer system-level analysis (see for example Burzynski et al. (2020[17]); Dao et al. (2021[18])). While promising for regional forecasting, SEMs are constrained by data comparability and availability across countries.

**Bayesian approach** incorporate prior information – including expert judgement – into statistical models (see for example model (Bijak and Wiśniowski (2010[19])). Forecasting in the Bayesian approach is based on the construction of a probability distribution of the vector of future values of the dependant variable, $\mathbf{X_f}$, conditional on the vector of past (observed) values, $\mathbf{X}$, and taking into account the posterior knowledge on the parameters of the forecasting model, $\theta$. The predictive probability distribution of $\mathbf{X_p}$ can be interpreted as an average from the conditional predictive distribution $p(\mathbf{X_p}|\theta, \mathbf{X})$, weighted with the posterior probabilities of the parameters (see Bijak (2005[20]) for more detailed explanation). An important issue is the selection of the prior probability distribution of the estimated parameters, $p(\theta)$, reflecting the knowledge of the researcher. This approach allows coherent integration of model uncertainty, data uncertainty and inherent uncertainty about future conditions. Bayesian methods are especially valuable when historical data are sparse or noisy. For example, if an expert believes a given flow is non-stationary, the prior distribution for an autoregressive term in the model takes a value of 1 or above. See Chapter 6 for more details).

**Machine learning models** are data-driven rather than assumption-driven and are increasingly used for short-term (days to months) migration forecasting (see for example Carammia et al. (2022[21]); Robinson and Dilkina (2018[22])) For example, Dynamic Elastic Net (DynENet) is a relatively new type of regularisation method that combines model selection and estimation within a single framework. It offers: parsimony (selects only relevant variables), flow-specific adaptability, resilience to missing data, time adaptability (via moving training windows) and interpretability through its linear functional form (Carammia, Iacus and Wilkin, 2022[21]).

Another example of machine learning models is Long Short-Term Memory (LSTM) Networks, which is a form of recurrent neural networks (RNN) capable of learning long-term dependencies in sequential data. They maintain and update “memory” states to inform predictions. Though powerful in capturing non-linearity, LSTMs are often criticised for limited interpretability and a tendency to overfit.

### Box 4.2. Probabilistic projections of future trends in international migration in World Population Prospects 2024

For the first time, the medium variant of net international migration in the 2024 revision of the United Nations World Population Prospects was projected using probabilistic methods. A Bayesian hierarchical approach was used to model future trajectories of the crude net migration rate and to account for uncertainties and correlations across countries and over time. The analytical approach follows these 10 steps:

1. Using the annual estimates of crude net migration rates from 1950 to 2023 for all countries from the 2024 revision, a Bayesian hierarchical first-order autoregressive model, or AR(1), was fitted to crude net migration rate data for all countries to estimate the parameters of the Azose and Raftery (2015[23]) migration model.
2. By default, the estimation period used for model fitting referred to 1950 to 2023, but the application of autoregressive statistical methods in this context faced two types of challenges for some countries: (1) the country-specific overall rate of change and long-term mean net migration rate are influenced by the reference period used to fit the model, which might need to be adjusted depending on the country's historical context (e.g. conflicts or political and economic changes), and (2) the short-term and long-term projections are highly influenced by the most recent situation (e.g. countries experiencing large in/out flows of migrants or refugees in 2023). Expert knowledge was incorporated to inform prior distributions and provide further guidance to the model fitting with respect to the reference period used for a specific set of locations.
3. To address the challenges with the set of countries experiencing recent large movements of net international migrants or refugees, expert knowledge was used to define the expected target years when the situation in these locations is expected to return to some baseline level. For a subset of those that experienced large in/outflow of recent refugees, expertise was also used to determine what proportions of refugees are expected to return to their country of origin by the target year. For each of these cases, some linearly interpolated values were estimated between the 2023 baseline and the expected median value for the user-defined target year. All probabilistic trajectories were adjusted proportionately based on the expected values for the median trajectory of each country.
4. Using the results from the model fitting (step 1), and the additional constraints (steps 2 and 3), the R package ***bayesMig*** (Azose, Sevcikova and Raftery, 2023[24]) was used to create projections of net migration rates, and generate 2 000 trajectories from 2024 to 2100 for each country.
5. Upon review of the median projection results for the net migration rates for each country, for a subset of locations, the default level for the long-run projection up to 2100 was considered to deviate too much from prior expectations and expert knowledge about country history and context. For each of these cases, user-defined target values were defined and used as constraints for the user-defined projection period. All probabilistic trajectories were adjusted based on these expected values for the median trajectory of each country.
6. Additional constraints were also imposed on the cumulative migration rates over different time periods to prevent net migration from growing indefinitely, which would generate unrealistic population growth. The following upper-bound constraints were imposed on cumulative migration for each of the 2 000 trajectories of several countries of the Gulf region, where temporary labour migrants comprise a large share of the population. These limits were set

based on past fluctuations experienced by these countries and the unconstrained 95th percentiles.

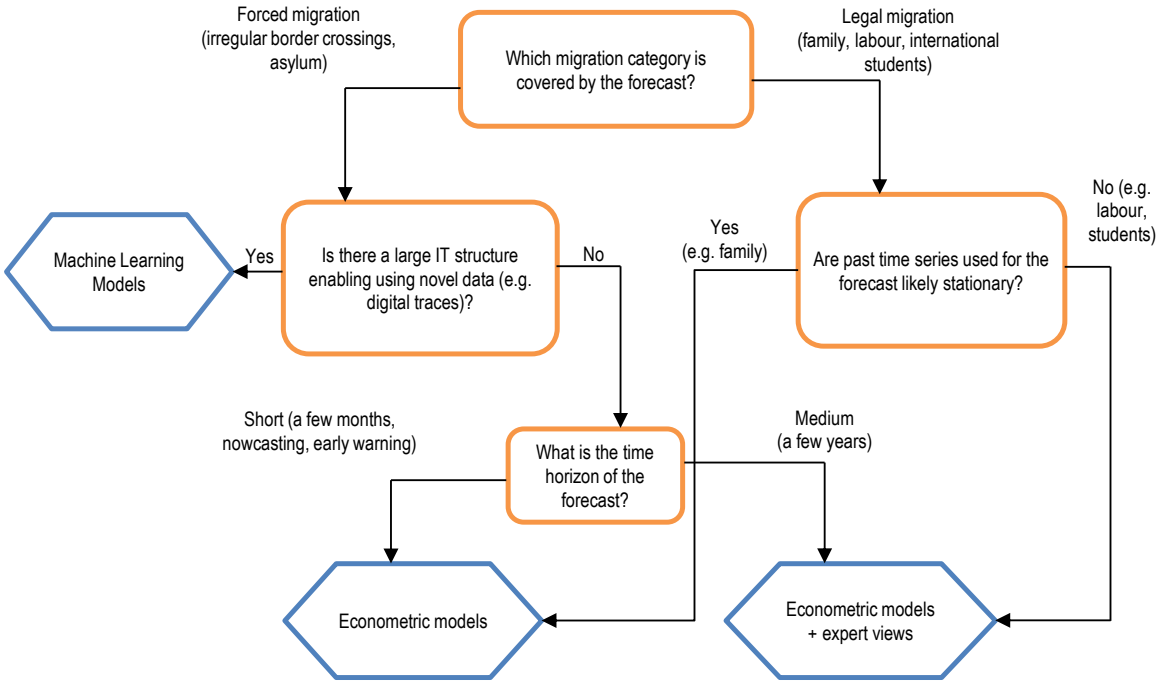
7. Based on the additional constraints from steps 5 and 6, a new set of projections of net migration counts was generated to consider these set of constraints in addition to those from steps to 2-3, and 2 000 trajectories were exported for each country.
8. Using these 2 000 trajectories of net migration rates, an initial set of net migration counts was computed to generate probabilistic population projections based on the population, fertility and mortality data by age and sex from the 2024 revision, and the 2 000 trajectories of total fertility rates, life expectancy at birth by sex, and net migration rates (from step 6) combined with model-based assumptions about the expected distribution of migrants by sex and age.
9. As with estimates, net migration balancing was carried out for projections as a final step to ensure that, at the global level, the sum of all net international migration flows equalled zero for each year of the projection period. To ensure that for each of the 2 000 trajectories the sum annually of net migration counts across all countries is zero, an extra step was implemented to "balance" total net migration counts.
10. The 2 000 adjusted and balanced trajectories of total net migration counts (from step 9) for each country were exported and used as inputs to the final run of the probabilistic population projections. They were also used to generate the median projected values for each country that were used for the medium deterministic scenario of the 2024 revision (after applying an additional rebalancing step as described in step 9 because the annual sum of all the country medians is initially unbalanced).

Source: UN DESA/POP/2024/DC/NO.10 (2024) Methodology of the United Nations population estimates and projections, https://population.un.org/wpp/assets/Files/WPP2024_Methodology.pdf.

## What is the appropriate forecasting model for each migration category?

The previous sections provide an overview of the different forecasting methods. Various criteria – such as the type of migration flow, the stationarity of the data series, the time horizon, and available resources – help determine the most appropriate forecasting model among them (see Figure 4.1). For example, for forced migration (irregular border crossings, asylum, as well as categories in-between regular and irregular, such as mixed migration phenomena), machine learning models using both traditional and innovative data are preferred, whereas for regular migration, standalone econometric models or econometric models with expert opinions can be utilised.

## Figure 4.1. Decision tree to identify the best migration forecasting models according to national practitioners' needs and resources




Where previously the dominant strand of the literature was on overall migration trends, in the last decade, attention has shifted towards forced displacement and irregular border crossings. In the meantime, the deterministic models and expert forecast as well as time series and causal models have been progressively supplemented by machine learning and Bayesian models that have become more operational as big data and large computational power became more accessible. This is true for short term forecasts such as nowcasting (Box 4.3) as much as for medium (12-18 months) forecast.

### Box 4.3. Nowcasting: Short-time horizon forecasts and early warning systems

A special category of migration forecasting methods concerns approaches to *nowcasting* – predicting migration flows for the current period or the nearest future. This strand of research has seen rapid developments in the recent years, amplified by the proliferation of large-scale data sources, including mobile phone data (Lai et al. (2019[25])), social media (Alexander et al. (2019[26]); Culora et al. (2021[27]); Leasure et al. (2023[28]); Rampazzo et al. (2024[29])), online search data (Böhme et al. (2020[30]); Avramescu and Wiśniowski (2021[31])) or experimental prediction markets (Morgenstern and Strijbis (2024[32])). Many of these methods have successfully attempted combining the advantages of the non-traditional sources, such as the timeliness of mobile phone or social media data, with the known features of more traditional data, such as censuses or surveys. This area of application of novel methods to predicting current or near-future migration flows is very promising, and has already gained traction also within the user community: organisations such as the International Displacement Monitoring Centre (IDMC (2022[33])) or the UN Refugee Agency (UNHCR (2025[34])) are developing their own nowcasting models for the populations of interest – particularly important for such rapidly-changing contexts as asylum or other forced displacement categories.

A special group of related methods, closely linked with the notion of preparedness for rare events (for discussion, see Bijak (2024[35])), concerns *early warnings* – detection of changes in the trends in migration or displacement flows, typically based on higher-frequency data (daily, weekly or monthly), such as asylum-related flows. Such techniques are aimed at identifying the potential changes, ideally contemporaneously or ahead of time, to flag them for further attention and possibly for operational response. The methods used for early warning applications include simple, descriptive techniques, such as cumulative sum (*cusum*) approaches (Napierała et al., 2021[36]) the momentum approach (Carammia, Iacus and Wilkin, 2022[21]) or the use of alerts when flows exceed a high percentile threshold during the preceding period. They also range from econometric models (Barker and Bijak, 2022[37]), to more advanced machine learning techniques, which at the same time aim to identify the signal indicating a possible change in trends, as well as selecting the most meaningful predictors from a range of options, including large-scale text databased (such as GDELT – Global Database of Events, Language and Tone, http://www.gdeltproject.org/, see Carammia, Iacus and Wilin (2022[21])). Early warning models are typically evaluated based on how well they are able to predict the actual changes in trends, avoiding both false alarms and missed detections.

The current methodological state of the art of migration forecasting is one of continuous innovation, constrained, however, from one side by the unavailability of adequate and appropriate data, and from the other side by the inherent uncertainty of migration processes. With these two constraints in mind, one fundamental methodological question in anticipating migration flows is, how to choose the best bespoke model for a task at hand? This section briefly summarises different classes of models that might be relevant for different types of migration and time horizons. It also examines the relative advantages and drawbacks of different types of models, taking into account the available resources – financial, human skills, equipment and so on – as well as data availability.

When discussing the choice of the right modelling approaches, it can help to start from the features of the process at hand. From a statistical viewpoint, one such key feature is *stationarity* – a property indicating whether the migration process has the same characteristics (such as mean or variability) over time (Bijak, 2011[1]; 2024[35])). Stationary processes are easier to model, and their predictive uncertainty is more straightforward to describe by using probabilities, and to calibrate based on past time series data. Stationarity indicates the presence of an equilibrium, to which migration eventually returns even after events generating much higher or lower flows. On the other hand, for non-stationary processes – such as random walks, processes with trends or level shifts – the equilibrium is constantly changing. As a result, the predictive uncertainty is increasing rapidly, rendering forecasts of little use beyond the horizon of a few years ahead (Bijak and Wiśniowski, 2010[19]).

The distinction between stationary and non-stationary processes is visible in the analysis of different types of migration flows, such as irregular border crossings, asylum, family, study, work, or returns. As argued by de Beer (2008[38]), such flows should ideally be analysed and forecast separately, as they are likely to be responding to different drivers and policies and be characterised by different dynamics. Treating them separately not only makes it possible to formally address the heterogeneity of migration flows, but also to find the best-matching statistical model in each case. This suggestion has been confirmed empirically: Bijak et al. (2019[5]) analysed several flows of migration into the United Kingdom in the early 2000s (returns being less volatile, labour migration more volatile, and asylum migration the most), ending up with concrete recommendations for model choice guided by the nature of specific processes and the availability of data. Crucially from the point of view of the practical uses of forecasts, Bijak et al. (2019[5]) suggested adopting a risk management approach, whereby the uncertainty (and unpredictability) of the processes is juxtaposed with their impact for a given stakeholder, to allow prioritising resources in the high-uncertainty and high-impact areas, such as asylum.

The practical recommendations for the choice of an appropriate method for forecasting are therefore flow-specific: Bijak et al. (2019[5]) suggested proceeding from (a) the analysis of statistical properties of the flows of interest (such as stationarity or uncertainty), through (b) the formal assessment of the available data (administrative registration or censuses vs. novel data and digital traces); to (c) selection of an appropriate model that is capable of reflecting the properties (a) and fits to available data (b) (See Figure 4.1).

Table 4.1 summarises the stylised characteristics of seven main types of migration flows, alongside the assessment of their properties and typically used data, and recommendations for method selection, based on the arguments presented above. How to precisely develop the most efficient model is discussed later (see Chapter 6). Unsurprisingly, volatile processes such as irregular border crossings and asylum exhibit non-stationarity and high uncertainty but given the availability of relatively long series of higher-frequency data, offer the possibility of using advanced machine learning methods, unlike most of the other flows, which are measured only quarterly or annually. The potential role of machine learning approaches for regulated migration streams where high-frequency administrative and linked datasets are becoming increasingly available have been barely investigated so far (see Box 6.3 for the use of a machine learning approach to forecast international student flows). It remains to be seen if those approaches may be useful for less volatile and uncertain migration categories.

### Table 4.1. Stylised characteristics of seven main types of migration flows, typical data, and recommendations for selecting possible forecasting methods

| Migration type | Stationarity | Uncertainty | Typical data sources | Frequency of data | Length of series | Possible methods |
|---|---|---|---|---|---|---|
| Irregular border crossings | Highly unlikely | Very high | Administrative data, novel data & "digital traces" | Typically daily, weekly, or monthly | Long: several years of weekly data | Statistical & machine learning; expert-based |
| Asylum | Highly unlikely | Very high | Administrative data, novel data & "digital traces" | Typically daily, weekly, or monthly | Long: several years of weekly data | Statistical & machine learning; expert-based |
| Family | Likely | Medium | Registers, administrative data, surveys | Typically quarterly or annual | Short: up to 10-20 years | Statistical & econometric time series |
| Study | Unlikely | Medium | Registers, administrative data, surveys | Typically annual | Short: up to 10-20 years | Statistical & econometric time series; expert-based |
| Labour | Unlikely | High | Registers, administrative data, surveys | Typically quarterly or annual | Short: up to 10-20 years | Statistical & econometric time series; expert-based |
| Return | Likely | Low to medium | Registers, censuses, surveys | Typically quarterly or annual | Short: up to 10-20 years | Statistical & econometric time series |
| Transit | Unlikely | Medium | Registers, surveys | Typically quarterly or annual | Short: up to 10-20 years | Statistical & econometric time series; expert-based |

In OECD countries and international organisations involving OECD countries, approximately half of the forecasting exercises combine statistical/econometric time series models with expert opinions (Table 4.2). These models rely on past data, such as asylum applications or residence permits, alongside non-quantitative sources like policy changes, intelligence reports, and scenario planning (see Table 5.1 for more details on data sources). In contrast, over one-third of forecasting approaches rely exclusively on data-driven models, while expert-based models are used by only three countries, notably in Switzerland.

For forecasting forced migration, most countries employ the mixed-method approach, while forecasting regulated migration flows is more varied. For the latter, the number of mixed approaches is roughly equal to those based purely on quantitative methods (four countries each), with two countries using expert elicitation only. Notably, very few countries and EU agencies have explored the use of machine learning models for migration forecasting so far.

### Table 4.2. What models are currently used in OECD countries?

| Institution | Migration type | Model |
|---|---|---|
| Frontex | Border Crossings | Methods including time series and regression analysis, historical analogy, panel consensus and qualitative risk management analysis. |
| Germany | Border Crossings | Methodologies incorporating AI and machine learning technologies (still in its nascent stage) have been used within a 12-month timeframe. |
| United States MAC | Border Crossings | Mixed-methods framework that combines machine learning models with judgmental adjustments from subject matter experts (SMEs). |
| Austria | Asylum Applications | Extrapolations of observed trends using scenario assumptions. |
| Belgium | Asylum Applications | Time series regression leading to 3 scenarios, with a seasonal effect included. Scenarios are then discussed by experts. |
| EUAA | Asylum Applications | Several models (e.g. TSLM, exponential smoothing, ARIMA, Machine Learning, time series regression based on month-to-month percentile changes in asylum applications over the last 6 years) are tried and their results are used to form 4 to 5 scenarios, from the (very) lowest to the (very) highest scenario. The likelihood of the scenarios built from the models are decided with expert elicitation. |
| France | Asylum Applications | DSED/DGEF: Linear, log or exponential regression model based on past asylum applications, adjusted after an expert meeting.<br>OFPRA: Linear or log regression model based on past asylum application changes. |
| Germany | Asylum Applications | Scenario-based forecast adapted from the EUAA methodology (see above). |
| Ireland | Asylum Applications | Scenario-based forecast adapted from the EUAA methodology (see above). This approach includes five scenarios, ranging from a “low” scenario (20th percentile changes) to a “high” scenario (80th percentile changes). Each one of these is based on historical data, supplemented by an additional data-driven scenario based on Seasonal and Trend decomposition. |
| Netherlands | Asylum Applications | Bayesian probabilistic forecasting combination of statistical time series, and harmonised expert intuition. |
| Norway | Asylum Applications | Regression models based on recent data and predictive factors that may affect future flows. |
| Poland | Asylum Applications | Method that accounts for the seasonality of migration phenomena based on the analysis of historical data and the knowledge of employees and collaborators. This includes information about countries of origin, diasporas in Poland and other European countries, and current migration practices of citizens from specific countries of origin. |
| Switzerland | Asylum Applications | Four possible scenarios are decided by experts, in terms of the number of applicants: Low, Medium, High, and Very High. Each scenario considers the previous year’s number of asylum applications, changes in the countries of origin of asylum seekers, migration routes to and within Europe, and significant transit and/or target countries. Additionally, changes in the asylum and return practices of Switzerland and other European countries (including their implementation of the Dublin Convention) and the United States are also taken into account. |
| Australia | Labour migration | Targets are defined through expert elicitation, with consultations (e.g. with federal administration, State and Territories administration, researchers, modellers from the Department of Treasury). |
| Japan | Labour migration | Macro-econometric model (forecasting future GDP etc., reflecting innovation & demographic changes) to estimate Japan’s foreign labour demand and country of origin’s labour supply, applied general equilibrium model to estimate industry-specific foreign labour demand (see Box 8.2). |
| Korea | Labour migration | Occupational labour supply projection and general equilibrium model forecasting labour demand (based on productivity and technological change), under 4 scenarios with different substitutability assumptions. Annual visa quota by sector can be established based on anticipated differential impacts of these labour demand/supply forecasts (see Box 8.2). |
| Poland | Labour migration | Quantitative model based on historical data augmented by qualitative research. |
| Türkiye | Labour migration | Delphi survey method using voluntary representation from public institutions, international organisations, academia and social partners. |

| Institution | Migration type | Model |
|---|---|---|
| Canada | All categories of regulated flows | Canada conducts demand forecasting for Temporary Residents (TR), Biometrics Enrollments, Citizenship Grants and Proofs, Passport Demand, and Permanent Resident Cards (Phase II). The forecasting framework for TR and Biometrics includes key categories such as Temporary Resident Visas (TRV), Work Permits (WP), Study Permits (SP), and their associated extensions. Forecasts are generated monthly using ARIMAX, Exponential Smoothing, and Unobserved Components models, supported by operational insights, domain expertise, and business knowledge. Assumption-based forecasting is also applied during periods of uncertainty, such as pandemics, conflicts, or significant policy changes. Building on this foundation, IRCC plans to explore the use of machine learning and hybrid statistical – machine learning approaches for migration forecasting. |
| France | All categories of regulated flows | ARIMA and SARIMA models (depending of the migration category), using past data from 2011 (excluding 2020). |
| Norway | All categories of regulated flows | Regression models based on recent past data. |
| Finland | All categories of flows (but family migration) | Qualitative and quantitative models using expert opinion, operating environment data and factors driving change. |
| Sweden | All categories of flows | Mixed approach based on quantitative time series analysis of historical patterns of applications in Sweden, including predictive factors, complemented with expert-based qualitative analysis. |
| United Kingdom | All categories of flows | Stable migration scenario based on 2022 visas, stay rates and planned policy changes. |
| United States CIS | All categories of flows | Time-series modelling selecting from a variety of time-series forecasts (ETS, ARIMA, ARFIMA, STL, TBATS) to choose the best model fit, excluding outliers when external forces (e.g. fee changes) inorganically affect receipts. For family migration (Box 6.1), for instance they leverage structural immigration relationships to forecast the time of the application (e.g. filing the family reunification form two years after the initial form is approved).Results are then discussed with subject matter experts from the operational directorates and the office responsible for new regulations. |

**Checklist**:

- Is the model considered fit for purpose from an operational perspective?
  - Model choice depends on the type of migration to be forecast (notably, its volatility) and on the time horizon concerned.
- Are the data adapted for this type of model?
  - Model choice depends on data frequency and quality, as well as whether past time series were stationary.

## What resources do countries need to run forecasting?

Establishing a functional and sustainable migration forecasting system requires not only technical infrastructure and skilled personnel but also strong institutional co-ordination, leadership, and long-term commitment across stakeholders.

Even the simplest forecasting approaches rely on solid IT infrastructure. Handling large and often complex datasets – such as migration flows, drivers, and contextual indicators – requires both sufficient storage and processing power. At a minimum, countries need a secure data server to store harmonised datasets, a computing environment capable of running statistical or machine learning models without performance delays, and a scraping server or API setup to ingest external data sources efficiently for big data or automated input systems. While such infrastructure is common in statistical agencies and research institutions, it is often less available in public administrations not typically focussed on data-intensive work. The initial investment may seem high, but it lays the foundation for all subsequent forecasting work.

### Box 4.4. Staffing Situation for Migration Forecasting in MAP Task Force countries

In the United States, as of April 2025, USCIS (U.S. Citizenship and Immigration Services) employed the equivalent of three full-time equivalents (FTEs) dedicated to forecasting, including two federal employees and one contractor. However, the standard staffing includes one branch chief and three peer contributors (all federal).

Furthermore, the Migration Analysis Center – the research hub located within the Office of Homeland Security Statistics (OHSS), which focusses on Monthly Encounter Projections relies on one FTE internally to manage projections. Additionally, the equivalent of 2 FTEs worth of expertise is contributed by subject matter experts who participate in surveys, monthly forecasting meetings, and other consultations.

In Switzerland, the migration analysis unit is made up of three analysts and a branch chief – although forecasting is only a small part of their activity. Considering the time spent by the staff for observation, analysis and explanation of trends in irregular migration and asylum, as well as the time spent on discussion between them and their country-of-origin expert counterparts, it is estimated that a total of 1.5 FTE is involved in asylum forecasting.

In the Netherlands, asylum forecasting is managed by a senior advisor, half-time employed by the Ministry of Justice. To produce the forecasts, the senior advisor uses migration related developments, which are analysed by a "Migration Radar" unit, comprising two FTEs. The senior advisor also uses expert intuitions gathered from colleagues in international meetings. The time of these colleagues cannot be quantified. About two to three staff in the department then calculate the required staffing and budgets for the asylum process, based on the forecasts. All in all, about 5 FTEs in the Netherlands have their activity linked with asylum forecasting, with roughly half on the frontend.

The Swedish Migration Agency (SMA) has a dedicated function within a unit that works on forecasting. The Unit for Migration Analysis has five full-time Migration Intelligence Analysts and one Senior Analyst. The Migration Intelligence Analysts analyse how domestic and foreign developments will impact the future number of first-time applications (e.g. asylum, work permits, study permits, family reunification) lodged at the SMA. The majority of their monitoring and analytical reports have an element of forecasting, which goes into an integrated model. The SMA produces a weekly situational report. The Unit for Migration Analysis also produces a quarterly prognosis paper – not directly connected to the situational reports – in co-operation with the SMA's Planning Department. Although the unit produces a number of other reports not directly connected to forecasting, the Planning Department co-ordinates the SMA's planning as well as financial and statistical reporting. Forecasts of expenditure and operational forecasts are made by the Planning Department in the quarterly prognosis paper. Three people in the department hold primary responsibility for co-ordinating the prognosis report and developing and maintaining the forecast models: a Financial Controller, a Statistician, and an Operational Analyst. Most other experts and financial controllers who participate in this work do so as a minor part of their duties. In total, 9 staff are involved in migration forecast reporting in Sweden, but it is difficult to quantify exactly their full-time equivalent.

A dedicated core team is also essential, even for basic forecasting models (Box 4.4). This typically includes data analysts, statisticians, economists or coders who can manage data preparation, model development, and interpret results. As forecasting methods evolve, expertise in machine learning, web scraping, and big data analytics is increasingly valuable. However, these profiles are highly sought after in the private sector and often difficult to attract to public administrations due to wage constraints or hiring restrictions – especially in contexts where civil service positions are limited to nationals. In addition to technical staff,

effective forecasting systems require the involvement of wider working groups across relevant ministries, statistical offices and policy units. These groups contribute to framing scenarios, validating assumptions, and discussing timeframes and outputs. Regular co-ordination meetings are crucial for aligning expectations and ensuring the forecasts are actionable and policy-relevant.

Beyond staff and IT tools, several enabling conditions determine whether forecasting efforts will be effective and sustainable:

- **Leadership and governance**: A clear mandate and leadership structure are needed to champion the project and secure its integration into policymaking.
- **Data access and sharing**: Forecasting depends on the willingness of different actors – national statistical offices, ministries, border agencies, and international partners – to share timely and reliable data. Long-term data-sharing agreements, based on trust and mutual benefit, are critical.
- **Capacity to identify and harness relevant data**: Technical expertise must be complemented by the ability to map available data sources, assess quality and integrate them into forecasting models.
- **Creating the right conditions**: Adequate time, cross-institutional engagement, and political support must be in place to move from pilot projects to regularised forecasting systems.

Countries may also consider outsourcing parts of the forecasting work to national statistical offices, academic institutions, think tanks, or private sector firms – particularly when internal resources are limited. Outsourcing can offer access to cutting-edge technical expertise, but requires clear governance, quality control, and sustainability plans. Strategic partnerships can also foster innovation and enable joint access to restricted datasets or proprietary tools.

While most forecasting software tools (such as R or Python) are free and open-source, other costs may arise. Accessing data from commercial providers – such as airlines, social media platforms, or polling institutes – can be expensive. Some datasets produced by research projects or international organisations may only be free for academic use, excluding government users. In such cases, **data access agreements, participation in expert networks, or financial contributions** may be required.

**Checklist**:

- Can I secure medium- to long-term financial and human resources needed to develop and maintain a forecasting model?
  - Investing in IT infrastructure and staff with the relevant skills (internally or outsourced) is the foundation of an efficient forecasting system, which can only be maintained with the right staff mandate and political support.
- Is access to public and private data ensured overtime?
  - Data access agreements, participation in expert networks and financial contributions may be necessary to use data in the long run.

## How complex should the model be?

In deciding between simple and complex models, the question is how close should it be to reality? Existing literature on simple versus complex demographic forecasts, dating back to the 1990s, has been largely inconclusive about which is the better choice (e.g. Ahlburg (1995[39])). Similarly, in the area of general forecasting, a series of “Modelling Competitions” organised by Makridakis and co-authors since the 1980s (e.g. Makridakis, Spiliotis and Assimakopoulos (2022[40])), yielded conclusions that varied and changed over time. Earlier competitions favoured simple models, such as exponential smoothing, while the most recent editions indicated increasing accuracy of complex, and often proprietary, machine learning

algorithms, sometimes coupled with statistical methods and other approaches, such as forecast averaging (Makridakis, Spiliotis and Assimakopoulos (2022[40])). This clearly reflects the rapid evolution of the machine learning algorithms, availability of high-granularity data, and exponential increase in computational power in the last decades.

By design, the competitive advantage of machine learning and hybrid methods grows with the size of the available datasets and the frequency of data. For regular migration flows with quarterly or annual data, such as labour, study or family migration, we can expect simpler models to perform nearly as well as more complex ones, while being easier to implement, maintain and explain to stakeholders. In such situations, it makes sense to choose the lower complexity models: when performance is similar, a simpler model is likely to be a more cost-effective. At the same time, for more volatile migration processes, such as irregular border crossings or asylum, for which several large sources of high-frequency data exist, more sophisticated methods may be justified.

Another dimension of modelling complexity concerns the representation of the mechanisms driving migration. Ideally, models should include a causal description of underlying processes, as it can significantly enhance predictive capabilities (Willekens (2019[11]); see also Chapter 9). In recent years, there have been several attempts to model the causal dynamics of migration using agent-based models. In these models, simulated agents (such as migrants and decision makers) make choices based on their individual situations, knowledge, interactions with others and environment (e.g. Kniveton, Smith and Wood (2011[41]); Naivinit et al. (2010[42]); Klabunde and Willekens (2016[12]); Suleimenova and Groen (2020[43]); Bijak (2022[44]); Hinsch and Bijak (2023[45])). Whether the resulting model uncertainty is small enough to enable meaningful predictions, is another open question, addressed in Chapter 7. Still, for most practical applications, such complex models are too idiosyncratic to enable efficient implementation at a scale needed for most of the challenges of migration policy and practice.

**Checklist**:

- Is the model foreseen fit for purpose from a policy perspective?
  - Choosing the right model requires ensuring that the uncertainty is not too large to estimate the impact of relevant policy changes.
- What would be the pros and cons for using a simpler approach/model?
  - Model complexity depends on the migration category considered, as the level of uncertainty and the availability of high-frequency data varies. Forecasting volatile forced migration may justify use of sophisticated methods, while simpler methods usually work for regulated migration flows.

## How often should model results be updated?

No predictive model of social processes, no matter how complex and how accurate, can indefinitely withstand the test of time. One of the ways in which the forecasting approaches can be “future-proofed” – made more resilient and more reactive to changing migration reality – is to update them regularly. The update mechanism needs to be a part of the “business as usual” operations, built by design into processes aimed at ensuring anticipation and preparedness. Updating a model does not imply that the original model has failed; rather, updates reflect the availability of new data, improved methodologies, or changing contextual factors, which may affect the relevance or reliability of earlier results. Regular updates help ensure that models remain aligned with current realities and continue to support sound decision making.

The frequency of update clearly depends in part on the availability of resources. From the methodological viewpoint, the key variables are the horizon of the predictions and the frequency of data – whether they are daily, weekly, monthly, quarterly, or annual. For predictions with a horizon of 12-18 months, a standard

routine of annual updates seems appropriate: a full new year of data usually allows for recalibrating not only the trends, but also any patterns that may exhibit seasonality. At the same time, early warning models (for asylum or irregular migration forecasts) may need to be updated at least monthly, so that they can serve their intended purposes. Similarly, for long-range migration scenarios, such as those required for population projections, updates every couple of years would suffice, in line with the current practice of many national and international statistical institutions (UN, 2024[46]).

As for data frequency, methods for integrating data of different frequencies into migration forecasts exist, for example relying on multivariate models (such as vector autoregressions – VAR, e.g. Barker and Bijak (2025[47]). Using such models comes at an additional expense of increased predictive uncertainty, and additional resources needed to carry out the updates. Still, the extra effort may be justified if some covariates of the forecasting models are available at higher frequency than the migration variables (for instance, media reporting of events of political violence may be available daily, while asylum applications numbers may only be collated at a lower frequency). There may be some additional signal in such data. Automating data acquisition (for example via dedicated API – Application Programming Interface, or by web scraping) and re-estimation of forecasting models can help implement the update routines in practice.

In addition to regular updates, the forecasting models will also need amending and resetting outside of the standard cycle, whenever there is a major migration event, change in political circumstances, or the policy environment. This would be a simple reflection of the dynamic nature of migration, moving from one social, political and economic equilibrium to the next often in a matter of months, if not days. Examples of such events that would trigger updates can include an eruption of major conflict leading to a large number of asylum seekers moving to other countries, a major shift in a visa regime, either relaxing or tightening the rules of admission and stay, or an economic downturn, likely to drive either more or less labour migration, depending on the nature of a particular event.

**Checklist**:

- Can routine updates of the data in the model be matched with the frequency of the prediction?
  - Updates are necessary but only possible as new data become available or when a major event occurs.

## References


Ahlburg, D. (1995), “Simple versus complex models: Evaluation, accuracy, and combining”, *Mathematical Population Studies*, Vol. 5/3, pp. 281-290, https://doi.org/10.1080/08898489509525406. [39]

Alexander, M., K. Polimis and E. Zagheni (2019), “The Impact of Hurricane Maria on Out-migration from Puerto Rico: Evidence from Facebook Data”, *Population and Development Review*, Vol. 45/3, pp. 617-630, https://doi.org/10.1111/padr.12289. [26]

Avramescu, A. and A. Wisniowski (2021), “Now-casting Romanian migration into the United Kingdom by using Google Search engine data”, *Demographic Research*, Vol. 45, pp. 1219-1254, https://doi.org/10.4054/demres.2021.45.40. [31]

Azose, J. and A. Raftery (2015), “Bayesian Probabilistic Projection of International Migration”, *Demography*, Vol. 52/5, pp. 1627-1650, https://doi.org/10.1007/s13524-015-0415-0. [23]

Azose, J., H. Sevcikova and A. Raftery (2023), "bayesMig: Bayesian Projection of Migration", *CRAN: Contributed Packages* (database), https://doi.org/10.32614/cran.package.bayesmig (accessed on 4 September 2025). [24]

Barker, E. and J. Bijak (2025), "Mixed-frequency VAR: a new approach to forecasting migration in Europe using macroeconomic data", *Data & Policy*, Vol. 7, https://doi.org/10.1017/dap.2024.82. [47]

Barker, E. and J. Bijak (2022), *Could we have seen it coming? Towards an early warning system for asylum applications in the EU.*. [37]

Beine, M., S. Bertoli and J. Fernández-Huertas Moraga (2015), "A Practitioners' Guide to Gravity Models of International Migration", *The World Economy*, Vol. 39/4, pp. 496-512, https://doi.org/10.1111/twec.12265. [14]

Beyer, R., J. Schewe and H. Lotze-Campen (2022), "Gravity models do not explain, and cannot predict, international migration dynamics", *Humanities and Social Sciences Communications*, Vol. 9/1, https://doi.org/10.1057/s41599-022-01067-x. [15]

Bijak, J. (ed.) (2024), *From Uncertainty to Policy: A Guide to Migration Scenarios*, Edward Elgar Publishing, https://doi.org/10.4337/9781035319800. [35]

Bijak, J. (2022), *Towards Bayesian Model-Based Demography*, Springer International Publishing, Cham, https://doi.org/10.1007/978-3-030-83039-7. [44]

Bijak, J. (2011), *Forecasting International Migration in Europe: A Bayesian View*, Springer Netherlands, Dordrecht, https://doi.org/10.1007/978-90-481-8897-0. [1]

Bijak, J. (2005), "Bayesian Methods in International Migration Forecasting", *CEFMR Working Paper*. [20]

Bijak, J. et al. (2019), "Assessing time series models for forecasting international migration: Lessons from the United Kingdom", *Journal of Forecasting*, Vol. 38/5, pp. 470-487, https://doi.org/10.1002/for.2576. [5]

Bijak, J. and A. Wiśniowski (2010), "Bayesian Forecasting of Immigration to Selected European Countries by using Expert Knowledge", *Journal of the Royal Statistical Society Series A: Statistics in Society*, Vol. 173/4, pp. 775-796, https://doi.org/10.1111/j.1467-985x.2009.00635.x. [19]

Böhme, M., A. Gröger and T. Stöhr (2020), "Searching for a better life: Predicting international migration with online search keywords", *Journal of Development Economics*, Vol. 142, p. 102347, https://doi.org/10.1016/j.jdeveco.2019.04.002. [30]

Burzynski, M., C. Deuster and F. Docquier (2020), "Geography of skills and global inequality", *Journal of Development Economics*, Vol. 142, p. 102333, https://doi.org/10.1016/j.jdeveco.2019.02.003. [17]

Carammia, M., S. Iacus and T. Wilkin (2022), "Forecasting asylum-related migration flows with machine learning and data at scale", *Scientific Reports*, Vol. 12/1, https://doi.org/10.1038/s41598-022-05241-8. [21]

Culora, A. et al. (2021), *Using social media data to 'nowcast' international migration around the globe.*. [27]

Dao, T. et al. (2021), “Global migration in the twentieth and twenty-first centuries: the unstoppable force of demography”, *Review of World Economics*, Vol. 157/2, pp. 417-449, https://doi.org/10.1007/s10290-020-00402-1. [18]

de Beer, J. (1997), “The effect of uncertainty of migration on national population forecasts: The case of the netherlands”, *Journal of Official Statistics*, Vol. 13/3. [8]

Disney, G. et al. (2015), *Evaluation of existing migration forecasting methods and models*. [4]

Hinsch, M. and J. Bijak (2023), “The Effects of Information on the Formation of Migration Routes and the Dynamics of Migration”, *Artificial Life*, Vol. 29/1, pp. 3-20, https://doi.org/10.1162/artl_a_00388. [45]

IDMC (2022), *Nowcasting internal migration movements by using an AI-enabled news analyses tool. Internal Displacement Monitoring Centre*. [33]

Keilman, N., D. Pham and A. Hetland (2001), *Norway’s Uncertain Demographic Future*, http://hdl.handle.net/11250/181350. [9]

Klabunde, A. and F. Willekens (2016), “Decision-Making in Agent-Based Models of Migration: State of the Art and Challenges”, *European Journal of Population*, Vol. 32/1, pp. 73-97, https://doi.org/10.1007/s10680-015-9362-0. [12]

Kniveton, D., C. Smith and S. Wood (2011), “Agent-based model simulations of future changes in migration flows for Burkina Faso”, *Global Environmental Change*, Vol. 21, pp. S34-S40, https://doi.org/10.1016/j.gloenvcha.2011.09.006. [41]

Lai, S. et al. (2019), “Exploring the use of mobile phone data for national migration statistics”, *Palgrave Communications*, Vol. 5/1, https://doi.org/10.1057/s41599-019-0242-9. [25]

Leasure, D. et al. (2023), “Nowcasting Daily Population Displacement in Ukraine through Social Media Advertising Data”, *Population and Development Review*, Vol. 49/2, pp. 231-254, https://doi.org/10.1111/padr.12558. [28]

Lutz, W. et al. (2000), *New Developments in the Methodology of Expert- and Argument-Based Probabilistic Population Forecasting*. [10]

Lutz, W. and S. Scherbov (1998), , *European Journal of Population*, Vol. 14/1, pp. 1-17, https://doi.org/10.1023/a:1006040321755. [6]

Makridakis, S., E. Spiliotis and V. Assimakopoulos (2022), “M5 accuracy competition: Results, findings, and conclusions”, *International Journal of Forecasting*, Vol. 38/4, pp. 1346-1364, https://doi.org/10.1016/j.ijforecast.2021.11.013. [40]

McAlpine, A. et al. (2020), “Agent-based modeling for migration and modern slavery research: a systematic review”, *Journal of Computational Social Science*, Vol. 4/1, pp. 243-332, https://doi.org/10.1007/s42001-020-00076-7. [13]

Morgenstern, S. and O. Strijbis (2024), “Forecasting migration movements using prediction markets”, *Comparative Migration Studies*, Vol. 12/1, https://doi.org/10.1186/s40878-024-00404-0. [32]

Naivinit, W. et al. (2010), "Participatory agent-based modeling and simulation of rice production and labor migrations in Northeast Thailand", *Environmental Modelling & Software*, Vol. 25/11, pp. 1345-1358, https://doi.org/10.1016/j.envsoft.2010.01.012. [42]

Napierała, J. et al. (2021), "Toward an Early Warning System for Monitoring Asylum-Related Migration Flows in Europe", *International Migration Review*, Vol. 56/1, pp. 33-62, https://doi.org/10.1177/01979183211035736. [36]

Rampazzo, F. et al. (2024), "Assessing Timely Migration Trends Through Digital Traces: A Case Study of the UK Before Brexit", *International Migration Review*, Vol. 59/1, pp. 119-140, https://doi.org/10.1177/01979183241247009. [29]

Raymer, J. and F. Willekens (eds.) (2008), *Forecasting international migration: time series projections vs argument-based*, John Wiley. [38]

Robinson, C. and B. Dilkina (2018), "A Machine Learning Approach to Modeling Human Migration", *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies*, pp. 1-8, https://doi.org/10.1145/3209811.3209868. [22]

Sardoschau, S. (2020), *The Future of Migration to Germany." Assessing Methods in Migration Forecasting*. [2]

Sohst, R., E. Acostamadiedo and J. Tjaden (2023), "Reducing uncertainty in Delphi surveys: A case study on immigration to the EU", *Demographic Research*, Vol. 49, pp. 983-1020, https://doi.org/10.4054/demres.2023.49.36. [7]

Suleimenova, D. and D. Groen (2020), "How Policy Decisions Affect Refugee Journeys in South Sudan: A Study Using Automated Ensemble Simulations", *Journal of Artificial Societies and Social Simulation*, Vol. 23/1, https://doi.org/10.18564/jasss.4193. [43]

UN (2024), *World Population Prospects – 2024 Revision*, https://population.un.org/wpp (accessed on 15 November 2024). [46]

UNHCR (2025), *Nowcasting of refugee and asylum-seeker statistics*. [34]

Welch, N. and A. Raftery (2022), "Probabilistic forecasts of international bilateral migration flows", *Proceedings of the National Academy of Sciences*, Vol. 119/35, https://doi.org/10.1073/pnas.2203822119. [16]

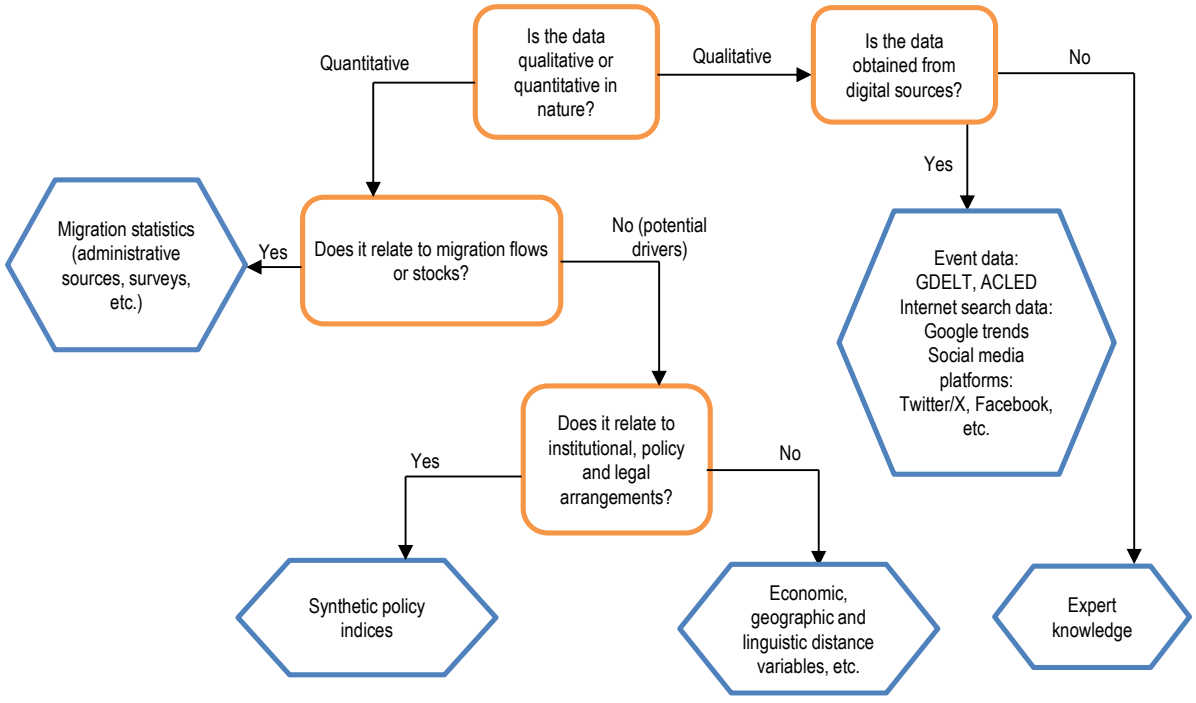
Willekens, F. (2019), "Towards causal forecasting of international migration", *Vienna Yearbook of Population Research*, Vol. 1, pp. 199-218, https://doi.org/10.1553/populationyearbook2018s199. [11]

Wiśniowski, A. (2021), "Migration forecasting using new technology and methods", in *Research Handbook on International Migration and Digital Technology*, Edward Elgar Publishing, https://doi.org/10.4337/9781839100611.00039. [3]

# 5 What data are necessary to conduct the forecasting exercise?

After selecting an appropriate forecasting model (Chapter 4, Figure 4.1), identifying the right data becomes more straightforward. In some cases, the model itself is data-driven – such as for machine learning approaches – or its functional form is determined by data availability. However, beyond considerations such as migration category and time horizon, other critical dimensions must be taken into account when selecting suitable data (Figure 5.1).

### Figure 5.1. Decision tree to identify the best data according to migration forecasting models




## How to assess the existing sources of data?

Reliable data are the foundation of migration forecasting. However, data on migration often suffer from gaps, variability in quality and availability, and sometimes discontinuity. Among the most serious constraints are timeliness and frequency. Most data collections, whatever the migration category, are published several weeks or months after the reporting period, which make them of little use for short-term forecasting. Similarly, many data sources are available at the yearly or monthly level, which again may not match with the granularity of certain forecasting tasks. More generally, as increasingly sophisticated statistical models become available, as well as the necessary software and computational capacity, data can become the main

constraint for forecasting exercises. This subsection provides a framework to assess existing data sources, ensuring they meet the quality and availability requirements necessary for robust forecasting.

The data used for forecasting, of whatever type and provenance, should be subject to a quality assessment, not only to gauge the different issues that can potentially impact the forecasts, but also to assess the associated measurement uncertainty that can be propagated into the predictions. To enable comparisons, and improve replicability, such an assessment can follow a standardised protocol, adapted where needed to suit specific needs and forecasting tasks. Elements of such a protocol would include creating standardised meta information about the data. Assembling and documenting the inventory not only helps provide crucial information on measurement uncertainty but also helps ensure smooth continuity of the forecasting processes, for example following staffing changes, as well as enabling quality assurance.

Assessing data quality also allows to better map gaps in data series. These gaps must not be considered as insurmountable hurdles, but opportunities for improvement. Indeed, gaps can be addressed by improving the data collection, through the use of alternative accurate data sources or the use of statistical techniques to fill them (such as imputations). If data sharing is part of the data collection process, it can also be improved through more regular contacts with data providers, data sharing agreements, with precise discussions about advantages to better share data for both providers and users. Going forward, the assessment of data gaps can be used as arguments to improve data infrastructures. In order to support these actions, financial resources must be invested accurately and organised, starting with the most troublesome data gaps preventing the forecasting model to be fully efficient.

Although quality requirements depend on the objectives of the forecasting task (mainly the type of migration and the forecasting horizon), the approach to quality assurance can be similar, whatever the migration category is. The first step consists of selecting the main criteria, which can be summarised in an assessment matrix. Important criteria that will likely be relevant to most forecasting tasks are definitions, frequency, coverage, accuracy, timeliness, granularity and continuity of available data (for a similar approach, see Bijak, Forster and Hilton (2017[1])).

- Definitions: What is the object of measurement, and how clearly are the indicators defined? Data definitions refer to the descriptions or specifications that explain the meaning, format, and structure of data elements in a dataset.
- Frequency: How often are data points collected or recorded over a given period of time? Data frequency determines the granularity and temporal resolution of a dataset.
- Coverage: To what extent does the dataset capture relevant information across specific dimensions – such as time, geography, population, or subject areas? Coverage determines how comprehensively a dataset represents the phenomenon it aims to describe.
- Accuracy: To what degree does the dataset correctly represent the real-world phenomena or conditions it aims to describe? Data almost always incorporate some degree of uncertainty, but can be more or less transparent about the biases they may have.
- Timeliness: To what extent are data available and up to date in relation to the use case? Timeliness depends on how quickly data are collected, processed, and made accessible after an event or observation. Timely data can be essential for short-term forecast, but not necessarily for longer-term projections.
- Granularity: What level of detail is required? For example, labour migration forecasting may require data disaggregated by industry, occupation and for some countries by region (see Box 8.2 for an example). Forecasting asylum applications may need not only sociodemographic information but the last county of residence (transit) to ensure accurate estimates.
- Continuity: Will we have access to the data in future as well? Sometimes external factors – such as changes in data collection technology, software platforms, institutional priorities or funding can affect the ongoing availability of the data. Furthermore, will data characteristics (with respect to all previous criteria) remain stable overtime?

More detailed metadata may also include their type (do they relate to migration directly or its drivers (contextual variables), collection methods, in addition to a multidimensional quality assessment according to a pre-defined set of criteria (Nurse, Hinsch and Bijak, 2023[2]). Data currently used by OECD countries for forecasting, described in Table 5.1, can be assessed based on these criteria systematically.

Probably the most important trade-off is between timeliness (which always goes along with frequency) and accuracy. Accurate indicators are subject to strict protocols in terms of collection, validation and standardisation, which require a significant amount of time and effort. In turn, highly timely data may be subject to limited quality assessment procedures. Granularity of data may also be inversely related to both accuracy and timeliness, as more detailed data can be harder to obtain quickly and reliably.

### Table 5.1. What data sources are currently used in OECD countries and by international organisations and agencies to forecast migration flows?

| Institution | Migration type | Variables included | |
|---|---|---|---|
| | | Quantitative | Qualitative |
| Frontex (CIRAM) | Border Crossings | | Reports on illegal border crossings by entry points; country reports on illegal stay detections, facilitator detections, irregular migrant apprehensions, refusals of entry, asylum applications, fraudulent document user detections, smuggled goods, return decisions, effective returns, passenger flows, data from VIS (Visa Information System), SIS (Schengen Information System), Eurodac; airport passenger reports on passenger inflows, number of staff at borders, information on passengers profile (country of origin, reason for entering the country, trip details, transportation means), police report on information from police records on cross-border crimes (from EUROSUR and EU Member States (MS) reports: wanted persons, criminal records, intelligence, document thefts), intelligence reports on situation in third countries (push factors, economic crisis, large or small incidents, difficulty of routes, health threats), situation at borders (such as information on the smugglers goals, motives, modus operandi and capabilities; meteorological conditions; EU MS staff skills, equipment, operational practices at border, interoperability) and situation in EU MS (policy changes, pull factors, procedures at border, effectiveness of countermeasures) |
| Germany | | Migration data | Event data, unstructured textual data |
| USA MAC | | Government of Mexico encounters/USA unemployment rate/ Remittances (Mexico, Guatemala, Colombia, Nicaragua /Ratio of removals divided by Title 8 encounters | Fatalities from violence against civilians in origin country/Fatalities from other violence in origin country)/Disasters<br>Judgmental adjustments from subject matter experts (SMEs) |
| Belgium | Asylum Applications | Asylum applications | Expert opinion from countries of origin, other EU MS and analytical experts |
| EUAA | | Asylum applications | Expert opinion from other EU MS |
| France | | Asylum applications, Dublin Statistics | Expert opinion from the Ministry of Interior and the Ministry of Foreign Affairs gathered in quarterly meetings, with inputs from EUAA reports, Frontex reports on Irregular Border Crossings (IBCs), French border police reports on border crossings, Ministry of Foreign Affairs country reports, French Refugee Office reports |

| Institution | Migration type | Variables included | |
|---|---|---|---|
| | | Quantitative | Qualitative |
| Germany | | Asylum applications | |
| Ireland | | Asylum applications | |
| Netherlands | | Asylum applications | Expert opinion (other EU MS) |
| Norway | | Asylum applications, border crossings in EU countries, different indicators in countries of origin and smuggling routes | |
| Poland | | Asylum applications | Expert opinion |
| Switzerland | | | Situation in the countries of origin, situation on the routes, economic situation in the transit countries, measures taken by European destination countries in the area of (asylum) migration, Switzerland's measures in the area of (asylum) migration, implementation of the Dublin Convention and EU Pact on migration and asylum by Switzerland |
| USA CIS | Family Migration | Applications for family-based cases by family type | Judgmental adjustments on recent procedure or policy changes |
| | | | |
| Australia | Labour migration | | Bilateral consultations with e.g. the department of Treasury, other federal administrations, States and Territories administration, researchers |
| Japan | | Target GDP, capital stock, baseline labour force, per capita GDP growth rate by country, total population by country, number of foreign workers coming to Japan (gross and net flows) | Scenario based on further automation. |
| Korea | | Labour supply by sex, age, education, sector, and occupation, productivity growth, technological change | Scenarios based on the substitutability of the labour force across sectors and skill levels |
| Poland | | Work permits | Qualitative research, using in particular migration intentions in the main countries/regions of origin |
| Türkiye | | | Expert opinion from public institutions, international organisations, academia and social partners |
| Canada | All categories of regulated flows | Demand forecasting uses a mix of IRCC operational data which is available daily and can be grouped into weekly or monthly time series and external data from organisations such as Statistics Canada, Transport Canada, the International Air Transport Association (IATA), the Conference Board of Canada, and Oxford Economics. For modelling, monthly time series are usually used. These external datasets offer a wide range of socio-economic indicators accessed through institutional subscriptions. Depending on the source, the information may be released annually or monthly and is updated on a monthly or quarterly basis. The indicators include economic measures that shape travel trends through changes in financial conditions and consumer behaviour, as well as social and geopolitical measures, since conflicts, disruptions, and other unexpected events can strongly influence global travel patterns. Taken together, these inputs provide a better overall picture and help improve forecast accuracy. | Demand forecasting incorporates a wide range of qualitative inputs to complement quantitative data. This includes structured collaboration with subject-matter experts, environmental scanning to identify emerging trends, and the review of outbound travel reports. These insights are drawn from multiple reputable sources – such as Statistics Canada, Transport Canada, the International Air Transport Association (IATA), the Conference Board of Canada, and Oxford Economics – which provide industry intelligence and socio-economic outlooks, influencing travel behaviour. Together, these qualitative elements help build a more comprehensive and forward-looking view of future demand. |
| France | | Residence permits by reason | |
| Norway | | Residence permits by reason | |
| Finland* | | Quantitative operating environment data and immigration driving factors | Qualitative operating environment data and immigration driving factors.<br>Expert opinion from the Ministry of Interior gathered in a foresight team |

| Institution | Migration type | Variables included | |
|---|---|---|---|
| | | Quantitative | Qualitative |
| Sweden | | Asylum applications, residence permits by reason | Scenarios prepared by experts |
| United Kingdom | | Long-term international migration statistics by reason (Visas, British nationals), long-run stay rates | Scenarios on changes in migration policy and situation in countries of origin |

* Family migration not included.

**Checklist**:

- What sources of data do I have to mobilise from other ministries or from private service providers?
- Is access to the data guaranteed over time?
  - Anticipating changes in data collection technology, software platforms, institutional priorities, database structure or funding may be necessary.
- What is the level of quality and granularity of the data? Can it be improved?
  - The level of quality and granularity impacts the model choice. If it prevents the model to be fully efficient, improving data collection and infrastructures may be necessary steps.
- Are data available under similar timeframe?
  - The data frequency and differences in timeframe and data availability (calendar vs. fiscal year, final vs. provisional) determine the forecast frequency.

## Where to find and how to incorporate qualitative data?

Qualitative data have long been understood as non-numerical information describing qualities, characteristics, or attributes of a given phenomenon. As such, they capture the subjective and often contextual aspects of an issue, providing insights into the why and how of behaviours, attitudes and processes. Text analysis models, however, make it possible to quantify traditionally non-numerical data. Moreover, the explosion of online information and the growing availability of computational capacity have made available large volumes of (often real-time) traditionally qualitative data. This is the case, for example, with digital traces left by various types of behaviour, including human movements. As a result, "qualitative data" are increasingly used to describe data that originate from non-numerical sources but may be (and often are) turned into numerical quantities.

Qualitative data can provide critical insights into migration drivers, complementing quantitative datasets. Key sources include digital sources, expert knowledge, and information such as reports and open questions in surveys. Among digital sources potentially relevant to migration forecasting, we find event data, satellite imagery, web-search data, data on air passengers or social media data. OECD countries have been slow to use these for forecasting purposes (Table 5.1). However, each of these types of data may provide insights to forecast all kind migration categories, especially for more complex forced migration phenomena. They can be used as a proxy for motivations or intentions to migrate, actual mobility, displacements, attitudes, financial transactions, and so forth (for recent reviews, see Cesare et al. (2018[3]), Sirbu et al. (2020[4]), Iacus et al. (2022[5])).

Data on online searches have been used either in isolation (Böhme, Gröger and Stöhr, 2020[6]) to predict total international migration or combined with other data (Carammia, Iacus and Wilkin, 2022[7]) to predict asylum applications. Such data are made freely available by Google Trends (https://trends.google.com/) and can be checked via user-friendly dashboards and downloaded in tabular format. The main parameters that can be set include keywords (or topics, pre-aggregated sets of keywords that have the advantage of being language-insensitive), location (country or sub-national levels), and time window. Other digital data relevant to migration modelling are those about events in the world. Two established sources of such data

are the Global Database of Events, Language, and Tone (GDELT) and the Armed Conflict Location & Event Data Project (ACLED). GDELT (https://www.gdeltproject.org/) is a repository of a large number of types of geolocated and content-coded event reported in the world's broadcast, print and web media, in more than 100 languages. ACLED also collects and analyses data on political violence and protest events worldwide (see Box 5.1 for more details). Both sources are particularly useful for forecasting forced migration such as asylum and irregular border crossings.

### Box 5.1. Examples of digital sources collecting qualitative data

#### Google Trends data

Google Trends data do not provide information on absolute levels of searches, but on relative levels (ranging between 0-100). When a single indicator is selected, the peak of the series represents the point in the selected time window with the maximum volume of searches for that topic at that point in time (and space). The rest of the points are relative to that peak. When instead more than one series is selected, they are scaled against each other. This means that for most modelling purposes, single indicators must be downloaded separately so that they can then be analysed individually. Running the analysis on different time windows results in different data, since, if new peaks in the volume of searches happen, the series must be re-scaled. Several statistical packages exist that support downloading (via the API), processing and analysing the data. Among R packages, "gtrendsR" is the most established and is available on CRAN (https://cran.r-project.org/web/packages/gtrendsR/index.html). Newer packages include "trendecon" (https://trendecon.github.io/trendecon/index.html) or "gtrendsAPI" (Correia, 2024[8]).

#### GDELT

GDELT allows for real-time tracking of events, sentiment analysis and geospatial mapping of crises or trends. GDELT's methodology emphasises coverage, automation, and scalability. The sources of data comprise various global news media, including traditional media outlets and blogs.

Data ingestion in GDELT is continuous, with updates every 15 minutes. Data collection is fully automated and based on scraping content from the monitored media sources, with non-English content translated automatically. Events are identified and coded with text-processing algorithms, with relevant codes describing the type of event, its location, the actors involved, and even the tone based on sentiment analysis. The largely automated approach to data collection and coding results in high coverage, frequency, and timeliness. This may be to the detriment of accuracy, although data can be inspected ex-post by the users as each observation provides a link to the original news item (links are however sometimes broken). Another potential source of bias is underreporting in media-poorer regions (and over-reporting in media-rich ones), in addition to ordinary biases in coverage and tone within media sources. GDELT provides access to its API, and R packages are available for querying and analysing both platforms. GDELT also permits direct analysis via Google BigQueries. The EUAA has already aggregated GDELT data relevant to asylum-related migration into a composite indicator, referred to as the Push Factor Index (PFI).

#### ACLED

Its methodology emphasises comprehensive coverage, high granularity, and systematic validation to provide reliable, real-time data. The sources of data include media reports, partner networks, academic and research publications, social media, and government and international reports. The data collection process includes more rigorous protocols for event identification, coding, validation and quality control, with involvement of human researchers. Data, in turn, are updated on a weekly basis. Although not

comparable to GDELT with this respect, this is still an impressive timeliness considering the validation efforts, and may be sufficient for most forecasting tasks with respect to forced migration, except of course for early warning and nowcasting.

ACLED provides access to its APIs, and R packages are available for querying and analysing both platforms. GDELT also permits direct analysis via Google BigQueries.

1. See for example GDELTtools (https://cran.r-project.org/web/packages/GDELTtools/index.html).
2. See for example acled.api (https://cran.r-project.org/web/packages/acled.api/index.html) and acledR (https://github.com/dtacled/acledR)

Other important digital sources of qualitative data are social media platforms, including Twitter/X, Facebook, Instagram, or Linkedin (Iacus et al., 2022[5]). Applications can include sentiment analysis to measure migration intentions or attitudes towards migration policies or events; monitoring or nowcasting stocks and flows (Zagheni et al., 2014[9]) or advertising data, which were proven to anticipate migration movements (Minora et al. (2022[10]), Zagheni, Weber and Gummadi (2017[11])).

Challenges and limitations of social media data and digital sources include data privacy issues, bias and representativeness, and noise (Iacus et al., 2022[5]). In particular, data from web-scrapping or APIs may not be fully anonymised, which might lead to breach in personal data privacy. Definitions of population and migration in digital sources usually do not comply with official statistics definition. Moreover, it remains difficult to fully assess the representativeness of these sources, given these tools are not used by the full extent of the covered population (selectivity bias), and this use may change over time due to technological changes or target audience evolution. More broadly, the use of these sources may also vary a lot across countries, which does not always ensure comparability between destination and origin countries.

Another important challenge is the regulatory and policy instability of social media and digital data (Iacus et al., 2022[5]). Data sharing between private companies and public and research institutions vary a lot and is not for granted. Single platforms can change their data access policy at any time. A good example of this risk is X, which was widely popular among researchers for providing free access to its database via API, but suddenly moved to paid subscription plans in May 2023, a few weeks after change in ownership. Last, digital sources may face technical issues, such as power outages of their infrastructures, which might create breach in time series, such as the one faced by GDELT in July 2025. Although the use of these innovative sources help improving migration forecasts, all these caveats regarding their data quality need to be taken into account when choosing to include them in forecasting models.

Alongside innovative data sources, more traditional qualitative data can still play an important role in migration analysis and forecasting. Experts can be a crucial source of information. Besides providing insights into migration processes which can help to design or fine-tune models, expert knowledge and intuition can moderate the forecasting models or feed directly into them. Qualitative expertise on scenarios plausibility, expert attribution of a probability to these forecasting scenarios, Delphi surveys and other expert elicitation methods are covered in the subsequent Chapter 6).

**Checklist**:

- What sort of qualitative data are available and/or accessible?
  - Digital sources such as Google Trends, GDELT or ACLED are increasingly used in migration forecasting as qualitative data, although their use has some data quality caveats to be taken into account.

## How to quantify policy indicators for statistical models?

Migration policies of various types may be relevant predictors of migration. Quantifying policy indicators enables analysts to evaluate their impact and integrate this knowledge into statistical models. This subsection presents some approaches to identify, measure, and model policy indicators. It discusses their pros and limitations and proposes ways for integrating them.

As an analytical problem, developing policy indicators is a challenge similar to other qualitative sources, as non-numerical information must be quantified. However, compared to other qualitative sources, public policies more easily lend themselves to systematic classification. Indeed, several projects exist that generated quantitative indicators of migration policy. Here we briefly review the most established ones (for a comprehensive review, see Scipioni and Urso (2017[12])). However, there is a limitation of all datasets developed within the framework of academic research projects. Such projects are not developed by governmental institutions and typically rely on external financial support. At the end of their lifecycle, they either secure structural support or risk being discontinued – or, at best, updated at irregular intervals.

### *1. IMPIC (Immigration Policies in Comparison)*

IMPIC (Helbling et al., 2017[13]) quantifies immigration policy restrictiveness across OECD countries, focussing on different migration categories (http://www.impic-project.org/). It covers 33 OECD countries from 1980 to 2010 and contains expert-coded assessments of policy restrictiveness. IMPIC categorises policies into four areas (labour migration, family reunification, asylum policies, and co-ethnic migration), using a 0-1 scale where 0 is least restrictive and 1 is most restrictive. Its strengths include the comprehensive coverage across time and countries, a differentiation among policy areas allowing for detailed analysis, and the use of a structured, transparent coding system with expert validation. The key limitation is the limited temporal coverage (1980-2010), which is an important issue in the migration context, as migration policies may change rapidly. Temporal coverage is a recurrent limitation of coding projects developed within academia, which lack continuous external financial support to ensure ongoing updates.

### *2. MIPEX (Migrant Integration Policy Index)*

MIPEX, which evaluates migrant integration policies across multiple domains, is run by an NGO. It covers 56 countries (mostly Europe, but also Canada, the United States, Australia, and others), assessing policies in eight areas: labour market mobility, family reunion, education, political participation, permanent residence, access to nationality, anti-discrimination and health (http://www.mipex.eu/). It uses a 100-point scale, where 100 indicates the best integration conditions. Key variables include the rights granted to migrants in each area, the legal framework and implementation, as well as some specific policy measures (such as language training or voting rights). MIPEX datasets also include a synthetic indicator measuring the overall policy approach to integration. MIPEX's strengths include the large coverage, which make it the most comprehensive integration index available, and the longitudinal comparisons available for many countries. Limitations include the focus on integration policies only, with no measures of restrictiveness of immigration policy outside of those that have an impact on integration. MIPEX may therefore be particularly poorly adapted to quantify asylum policies or irregular migration enforcement law. Another limitation is a subjective scoring system based on expert judgement.

### *3. DEMIG Policy Database (Determinants of International Migration)*

DEMIG tracks migration policy changes over time, differentiating between policies that become more restrictive or open (http://www.migrationinstitute.org/data/demig-data/demig-policy-1). The version originally developed at Oxford's International Migration Institute (de Haas, Natter and Vezzoli, 2015[14]) covered 45 countries for the period 1945-2013 (and even further back for some countries). Recently, the

dataset has been updated and extended within the framework of the QuantMig project (https://quantmig.eu/data_and_estimates/policy_database/), which released a version covering 31 European countries between 1990-2020 (Schreier et al. (2023[15]); for an application with an analysis of the "European migration policy mix", Czaika et al. (2023[16])).

While for most policy indicators the level of analysis (single observation) is a country at a point in time (typically years), in the DEMIG dataset each observation is an occurrence of policy change.[1] Each policy change is coded as a restricting (e.g. introducing visa restrictions, tightening deportation policies) or liberalizing policy (e.g. expanding labour migration quotas). Key variables are the policy area (e.g. labour migration, asylum), policy change type (expansion vs. restriction) and implementation mechanism (laws, executive orders, agreements). Clearly, coverage (the number of countries and the historical depth) is an important strength of DEMIG, as well as the quality of definition of policy categories. The main limitation is the lack of measures of absolute levels of restrictiveness/openness, and the lack of a complete matrix of country/year scores. Moreover, a future update of the dataset is not guaranteed, like other academic projects dependent on external financial support.

### *4. Dataset of World Refugee and Asylum Policies (DWRAP)*

The DWRAP, developed by Blair, Grossman and Weinstein (2021[17]) is the first global resource that compiles de jure asylum and refugee policies (https://datanalytics.worldbank.org/dwrap/).

This dataset codes national laws into a scored index relevant to forcibly displaced populations across 193 countries, from 1951 (the year of the UN Refugee Convention) to 2022. It encompasses 54 indicators on five core dimensions: Access to asylum rights (the ease of entrance and security of status, such as non-admission policy, rights exclusion, stay, penalty for irregular entry, accompanying family rights, family reunification, appeal procedure, subsequent applications); access to services (provision of public services and welfare, such as education, vocational training, language courses, affirmative action, care, health cost, sick foreigner procedure, subsidies, welfare); access to livelihoods (the ability to work and own property, such as access to jobs, self-employment, regulated professions, tax policy, real estate access, financial goods seizure, intellectual property, rent rights); movement (free-mobility, encampment policies, access to documentation and their costs); and participation in society (citizenship and political rights, such as naturalisation procedure, duration of residence to access rights, unaccompanied minor procedure, voter turnout, right to association). All indicators are compiled into five core dimension scores and then into a "comprehensive" score. Unlike IMPIC or DEMIG, the DWRAP project is intended to be transferred to JDC (the Joint Data Center on Forced Displacement from the World Bank and UNHCR), enabling potential updates on a regular basis.

### *5. UN-DESA Immigration Indicators within the World Population Policies Database*

The United Nations Department of Economic and Social Affairs (UN-DESA) provides migration policy indicators through its World Population Policies Database (WPPD) (https://www.un.org/development/desa/pd/content/world-population-policies). These indicators focus on governments' views and policies on international migration, covering aspects such as immigration levels, emigration, integration, and specific migration categories (e.g. skilled workers, refugees). This dataset has comprehensive spatial and temporal coverage, covering 195 countries and territories since 1976. However, it is updated only approximately every five years.

UN-DESA data focus on government perceptions and policy stances on migration, meaning it does not directly measure migration policy outputs. Key aspects include immigration, emigration, naturalisation, integration, border control, and refugees. The methodology is based on government-reported data collected via UN surveys and reports. Measures of governments' policy objectives include immigration levels (increase, maintain, decrease), emigration policies (facilitating or restricting outflows), skilled

migration (attracting high-skilled workers), refugee and asylum policies, and integration policies (citizenship, residency requirements). The dataset uses categorical rather than continuous quantitative measures, classifying policy types into three categories: no intervention (neutral stance), restrictive measures (e.g. tightening immigration rules) and liberal measures (e.g. encouraging immigration). In this sense, the approach resembles DEMIG. Limitations include the lack of direct measures of policy (since the data provide self-reported government positions, which may lead to potential bias), categorical (not numerical) classifications, and limited granularity compared to datasets like IMPIC, MIPEX, or DEMIG.

### *6. The OECD Indicators of Talent Attractiveness (ITA)*

The OECD ITA is the first comprehensive tool designed to evaluate the strengths and weaknesses of OECD countries in attracting and retaining different types of talented migrants. It assesses countries across seven core dimensions (Quality of Opportunity, Income and Tax, Future Prospects, Family Environment, Skills Environment, Inclusiveness, Quality of Life) and Visa and Admission Policy. Initially developed in 2019, the tool also provides post-simulation results by introducing the most favourable migration policies for each country. The 2023 edition of the ITA (http://www.oecd.org/en/data/tools/talent-attractiveness-2023.html) expands on the previous version by including four categories of talented migrants: highly educated workers, foreign entrepreneurs, university students, and start-up founders. Additionally, the new edition includes an expanded set of dimensions (e.g. health) for assessing a country's overall attractiveness. ITA is planned to be updated every three to five years.

In addition to the six indicators discussed above, Carammia and Iacus (2025[18]) are working on a method to extract the underlying shared variation from all these datasets, resulting in synthetic policy indicators. The approach is based on an algorithm developed for extracting continuous measures from heterogeneous survey data (Stimson, 2018[19]), something that standard factor analysis or principal component analysis cannot do because of the sparse nature of the data. The result are continuous measures of policy indicators, which can be estimated at different frequencies. The algorithm can ingest data at different frequencies and can work with matrices of varying density in terms of temporal coverage. In this way, the resulting indicator can fill temporal gaps among the various indicators. Moreover, because the resulting indicators describe the shared variation among the underlying data sources, they should also incorporate a smaller error.

To summarise, available policy indicators are quite different. Datasets vary in their objective, with some specialised in particular policy areas (e.g. MIPEX on integration policy); in their spatial and temporal coverage, covering different countries and time periods; and in their frequency. A general limitation is timeliness, as these datasets are not necessarily updated or have even been discontinued (as in the case of IMPIC), or their future update is unknown (DEMIG).

Timeliness is the most serious limitation of the indicators in their usefulness to feed in forecasting models. Indeed, political measures such as pull factors can rapidly impact variations in migration flow. Recent and frequently updated indicators are needed as input variables to forecasting models and exercises. In the end, most of these indicators are better suited for analyses of the past relationship between policy changes and migration flows. Supplementing these indicators with recent and high-frequently data such as social networks or web search engine queries could help solve the timeliness issue.

**Checklist**:

- Is it possible to use some existing database on policy indicators to improve forecasting models? Is that dataset accurate and up to date, and does it focus on the right issues?
- Should a specific parametrisation of the migration policies in my country be developed?
  - To be useful such parametrisation needs to ensure that the policy impact is neutrally assessed.
- How to maintain policy indicators when most have unreliable funding and are infrequently updated to incorporate policy changes?
  - As policy indicators datasets typically rely on external financial support, regular updates require renewed funding and knowledge transfers.

## Which software is the best fits for forecasting, causal effect estimation and projection?

Forecasting migration flows requires robust tools capable of handling diverse data types, forecasting horizons, and methodological approaches. Selecting the right software depends on the specific use case, the migration category covered, data requirements, and level of expertise among practitioners. Below is an overview of practical software packages in the R and Python language suited for time series analysis, forecasting, and causal effect estimation, tailored to the context of migration forecasting. R and Python are among the most widely used platforms for forecasting due to their extensive libraries, flexibility, and active user communities. These tools are well-suited for both traditional statistical models and modern machine learning approaches. We consider three large families of problems that can be addressed in R or Python: *i)* time series analysis and forecasting; *ii)* causal effect estimation for policy evaluation and *iii)* long-term projections. This list cannot be considered exhaustive, but its elements contain important building blocks. Most R packages and Python libraries come with on-line resources. R packages usually come with a so-called *vignette* or even journal papers, that contain many examples and reproducible codes. Python libraries often come with a notebook that can be executed on-line step by step. Note that both R and Python have low-level libraries and object types to deal with time series data, which are used across all the libraries and packages mentioned in this section.

### *1. Software for time series analysis in the context of migration forecasting*

The R package **forecast** (Hyndman and Khandakar (2008[20]); Hyndman et al. (2024[21])) is the ideal starting piece of software for traditional time series models like ARIMA and it handles exponential smoothing, seasonal decomposition and automated model selection and parameter tuning. It is suitable for practitioners aiming for efficient workflows. ARIMAX is not a function of the R package forecast, but is covered in the **TSA** package. On the Python side, **statsmodels** (Seabold and Perktold, 2010[22]) supports ARIMA, SARIMA, ARIMAX and other classical econometric models. It includes diagnostic tools to validate assumptions and assess model performance. As the standard forecasting approach to time series extrapolation is to apply ARIMA models, typically within the frequentist statistical paradigm for long enough series (as a rule of thumb, at least 20 observations), most of formal forecasting models for all categories as well as net migration rely on time series analysis (Bijak et al., 2019[23]). These models however appear to perform either relatively well or relatively poorly, depending on the type of migration and data availability. For example, non-stable flows such as asylum, can be forecast using models that assume non-stationarity (e.g. random walk model), while more stable labour migration have more orderly features than a non-stationary model would predict.

Short data series may require additional expert input regarding the future migration or the features of the processes. A direct R implementation of Bayesian Structural Time Series (BSTS) models is available in the **bsts** package (Scott, 2024[24]).

For Vector Auto-regressive (VAR) and Structural VAR models, the entry R package is **vars** (Pfaff, 2008[25]). VAR models can be useful, for example, for modelling interdependencies in migration drivers like macroeconomic indicators (GDP, wages and salaries, unemployment rates, and employment indicators) and systemic shocks (policy shifts, lowering or raising of migration and trade barriers, or political crises). The methodology is flexible enough to allow the inclusion of other drivers of migration in the models as long as the relevant data series are available. When it comes to Bayesian Panel VAR forecasting, a MATLAB-based BEAR toolbox (Bayesian Estimation, Analysis and Regression) is a powerful tool for academics, central bankers and policymakers (Dieppe, van Roye and Legrand, 2016[26]). For example Barker and Bijak (2025[27]) attempts to forecast immigration, emigration and net migration rates for both long- and short-term horizons through macroeconomic models using mixed-frequency.

### *2. Software for causal effects estimation in the context of migration forecasting*

R package ***CausalImpact*** (Brodersen et al., 2015[28]) uses Bayesian Structural Time Series (BSTS) to estimate causal effects of policy changes or interventions. It is ideal for evaluating the impact of migration policies such as visa reforms or border controls. The library ***causalimpact*** exists for Python too.

The package ***Synth*** (Abadie, Diamond and Hainmueller, 2011[29]) in R implements the Synthetic Control Methods (see Chapter 9) for policy evaluation. It allows practitioners to construct counterfactual scenarios to estimate causal impacts in non-experimental situations like in the migration policies context. An equivalent Python library is ***pysincon*** (Fordham, 2022[30]). These packages implement the basic version of the synthetic control methods. An approach that incorporates machine learning into the modelling is called Augmented Synthetic Control Method (Ben-Michael, Feller and Rothstein, 2021[31]). The corresponding software implementations are ***augsynth*** (Ben-Michael, 2025[32]) for R and again ***pysincon*** in Python.

### *3. Software for long-term demographic and migration projections*

Projections require tools that can simulate complex systems, capture individual behaviours, and aggregate them into macro-level insights. Methods like agent-based modelling (ABM), micro-simulation, and macro-simulation provide powerful frameworks for migration projections. These models can be informed by both previously mentioned forecasting models and experts' evaluations in different ways. In the class of ABM, among many software options, ***NetLogo*** (Wilensky, 1999[33]) can be used to create population projection models by simulating the behaviour of individual agents within a population. This allows you to explore how factors like birth rates, death rates, migration, and age structure can impact future population sizes over time, essentially creating a "virtual population" to study demographic trends under different scenarios. A well-known counterpart to NetLogo in Python for ABM is the ***Mesa*** (Kazil, Masad and Crooks, 2020[34]).

For micro-simulation projections, ***MicSim*** (Zinn, 2024[35]) is an R package suitable for analysing the impact of policies on individual-level migration decisions over time. Another well-known project is ***Modgen*** or Model Generator (Bélanger and Sabourin, 2017[36]) that enables the modelling of population dynamics and migration flows with detailed individual attributes. A sort of open-source counterpart is ***OpenM++*** (OpenM++, 2024[37]) that has both R and Python bindings.

One open-source macro-simulation software is ***DAPPS*** (Demographic Analysis and Population Projection System) developed by the U.S. Census Bureau for which an R package is being developed.[2] ***DAPPS*** is a comprehensive tool for demographic analysis and population projections. It utilises the cohort-component method and offers a modern graphical user interface for efficient workflow and enhanced data visualisation. For R packages used for UN long-term population projections (including migration), see Box 4.2.

### *4. Software for machine learning in the context of migration forecasting*

On the machine learning side, the ***prophet*** package (Taylor and Letham, 2021[38]) developed by Facebook, handles irregularities (like strong seasonalities, changepoints, event data, etc) or missing observations. This approach has also been proposed in long-term migration trend projections. This model can accept external regression variables. The prophet library also exists for Python.

The package ***glmnet*** (Friedman, Hastie and Tibshirani, 2010[39]) that implements Elastic Net regularisation, forms the backbone of the DynENet or Dynamic Elastic Net model (Carammia, Iacus and Wilkin, 2022[7]) which allows for scalable modelling with high-dimensional data, ideal for irregular migration flows such as asylum applications. The Python library ***scikit-learn*** (Pedregosa et al., 2011[40]) can be used for a custom implementation of the DynENet model.

A popular Python library that implements modern deep-learning techniques, is ***darts*** (Herzen et al., 2022[41]). The library offers a unified interface for traditional and deep learning-based forecasting models. It includes support for Recursive Neural Networks, LSTM (Long Short Term Memory), and ensemble models for forecasting. The package ***TSLSTM*** (Paul and Yeasin, 2022[42]), implements LSTM in R. More recently, Golenvaux et al. (2020[43]) included in their forecast a yearly incoming migratory flow to members of the OECD using a short-term memory (LSTM) approach combined with Google Trends data.

**Checklist**:

- Can the most adapted software for the foreseen model be purchased and uploaded?
  - Forecasters usually use R or Python to develop their models. Both are free-of-charge software.
- Can my old model be migrated to more efficient and flexible software?
  - Model software migration needs to avoid losing continuity of existing forecasts.

## References


Abadie, A., A. Diamond and J. Hainmueller (2011), "Synth: An R Package for Synthetic Control Methods in Comparative Case Studies", *Journal of Statistical Software*, Vol. 42/13, https://doi.org/10.18637/jss.v042.i13. [29]

Barker, E. and J. Bijak (2025), "Mixed-frequency VAR: a new approach to forecasting migration in Europe using macroeconomic data", *Data & Policy*, Vol. 7, https://doi.org/10.1017/dap.2024.82. [27]

Bélanger, A. and P. Sabourin (2017), *Microsimulation and Population Dynamics*, Springer International Publishing, Cham, https://doi.org/10.1007/978-3-319-44663-9. [36]

Ben-Michael, E. (2025), *augsynth: The Augmented Synthetic Control Method, R package version 0.2.0.*. [32]

Ben-Michael, E., A. Feller and J. Rothstein (2021), "The Augmented Synthetic Control Method", *Journal of the American Statistical Association*, Vol. 116/536, pp. 1789-1803, https://doi.org/10.1080/01621459.2021.1929245. [31]

Bijak, J. et al. (2019), "Assessing time series models for forecasting international migration: Lessons from the United Kingdom", *Journal of Forecasting*, Vol. 38/5, pp. 470-487, https://doi.org/10.1002/for.2576. [23]

Bijak, J., J. Forster and J. Hilton (2017), *Quantitative assessment of asylum-related migration: A survey of methodology*. [1]

Blair, C., G. Grossman and J. Weinstein (2021), "Liberal Displacement Policies Attract Forced Migrants in the Global South", *American Political Science Review*, Vol. 116/1, pp. 351-358, https://doi.org/10.1017/s0003055421000848. [17]

Böhme, M., A. Gröger and T. Stöhr (2020), "Searching for a better life: Predicting international migration with online search keywords", *Journal of Development Economics*, Vol. 142, p. 102347, https://doi.org/10.1016/j.jdeveco.2019.04.002. [6]

Brodersen, K. et al. (2015), "Inferring causal impact using Bayesian structural time-series models", *The Annals of Applied Statistics*, Vol. 9/1, https://doi.org/10.1214/14-aoas788. [28]

Carammia, M. and S. Iacus (2025), "Migration mood and policy responsiveness: a structural analysis of public opinion, policy, and migration flows in Italy (1990–2020)", *Journal of European Public Policy*, Vol. 33/1, pp. 74-104, https://doi.org/10.1080/13501763.2025.2584564. [18]

Carammia, M., S. Iacus and T. Wilkin (2022), "Forecasting asylum-related migration flows with machine learning and data at scale", *Scientific Reports*, Vol. 12/1, https://doi.org/10.1038/s41598-022-05241-8. [7]

Cesare, N. et al. (2018), "Promises and Pitfalls of Using Digital Traces for Demographic Research", *Demography*, Vol. 55/5, pp. 1979-1999, https://doi.org/10.1007/s13524-018-0715-2. [3]

Correia, R. (2024), "gtrendsAPI: An R wrapper for the Google Trends API", *Software Impacts*, Vol. 20, p. 100634, https://doi.org/10.1016/j.simpa.2024.100634. [8]

Czaika, M., H. Bohnet and F. Zardo (2023), "Categorical and spatial interlinkages within the European migration policy mix", *European Union Politics*, Vol. 25/1, pp. 173-196, https://doi.org/10.1177/14651165231209941. [16]

de Haas, H., K. Natter and S. Vezzoli (2015), "Conceptualizing and measuring migration policy change", *Comparative Migration Studies*, Vol. 3/1, https://doi.org/10.1186/s40878-015-0016-5. [14]

Dieppe, A., B. van Roye and R. Legrand (2016), "The BEAR toolbox", *European Central Bank Working Paper Series*, Vol. 1934. [26]

Fordham, S. (2022), *pysyncon: a Python package for the Synthetic Control Method*. [30]

Friedman, J., T. Hastie and R. Tibshirani (2010), "Regularization Paths for Generalized Linear Models via Coordinate Descent", *Journal of Statistical Software*, Vol. 33/1, https://doi.org/10.18637/jss.v033.i01. [39]

Golenvaux, N. et al. (2020), "An LSTM approach to Forecast Migration using Google Trends. :", *ArXiv*, https://arxiv.org/abs/2005.09902. [43]

Helbling, M. et al. (2017), "measuring immigration policies: the IMPIC database", *European Political Science*, Vol. 16/1, pp. 79-98, https://doi.org/10.1057/eps.2016.4. [13]

Herzen, J. et al. (2022), “Darts: User-Friendly Modern Machine Learning for Time Series,”, *Journal of Machine Learning Research*, Vol. 23/124, pp. 1-6. [41]

Hyndman, R. et al. (2024), *forecast: Forecasting functions for time series and linear models. R package version 8.23.0.9000*. [21]

Hyndman, R. and Y. Khandakar (2008), “Automatic Time Series Forecasting: The forecast Package for R”, *Journal of Statistical Software*, Vol. 27/3, https://doi.org/10.18637/jss.v027.i03. [20]

Iacus, S. et al. (2022), *Data innovation in demography, migration and human mobility*. [5]

Kazil, J., D. Masad and A. Crooks (2020), “Utilizing Python for Agent-Based Modeling: The Mesa Framework”, in *Lecture Notes in Computer Science, Social, Cultural, and Behavioral Modeling*, Springer International Publishing, Cham, https://doi.org/10.1007/978-3-030-61255-9_30. [34]

Minora, U. et al. (2022), “The potential of Facebook advertising data for understanding flows of people from Ukraine to the European Union”, *EPJ Data Science*, Vol. 11/1, https://doi.org/10.1140/epjds/s13688-022-00370-6. [10]

Nurse, S., M. Hinsch and J. Bijak (2023), “Mapping secondary data gaps for social simulation modelling: A case study of Syrian asylum migration to Europe”, *Open Research Europe*, Vol. 3, p. 216, https://doi.org/10.12688/openreseurope.15583.1. [2]

OpenM++ (2024), *Open-source microsimulation modeling platform*. [37]

Paul, D. and D. Yeasin (2022), *TSLSTM: Long Short Term Memory (LSTM) Model for Time Series Forecasting. R package version 0.1.0.*. [42]

Pedregosa, F. et al. (2011), “Scikit-learn: Machine Learning in Python”, *Journal of Machine Learning Research*, Vol. 12, pp. 2825-2830. [40]

Pfaff, B. (2008), “VAR, SVAR and SVEC Models: Implementation WithinRPackagevars”, *Journal of Statistical Software*, Vol. 27/4, https://doi.org/10.18637/jss.v027.i04. [25]

Schreier, S., L. Skrabal and M. Czaika (2023), *DEMIG-QuantMig Migration Policy Database*. [15]

Scipioni, M. and G. Urso (2017), *Migration Policy Indexes*. [12]

Scott, S. (2024), *bsts: Bayesian Structural Time Series. R package version 0.9.10.*. [24]

Seabold, S. and J. Perktold (2010), “Statsmodels: Econometric and Statistical Modeling with Python”, *Proceedings of the Python in Science Conference, Proceedings of the 9th Python in Science Conference*, pp. 92-96, https://doi.org/10.25080/majora-92bf1922-011. [22]

Sîrbu, A. et al. (2020), “Human migration: the big data perspective”, *International Journal of Data Science and Analytics*, Vol. 11/4, pp. 341-360, https://doi.org/10.1007/s41060-020-00213-5. [4]

Stimson, J. (2018), “The Dyad Ratios Algorithm for Estimating Latent Public Opinion”, *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, Vol. 137-138/1, pp. 201-218, https://doi.org/10.1177/0759106318761614. [19]

Taylor, S. and B. Letham (2021), *prophet: Automatic Forecasting Procedure. R package version 1.0*, https://CRAN.R-project.org/package=prophet. [38]

Wilensky, U. (1999), *NetLogo.*. [33]

Zagheni, E. et al. (2014), "Inferring international and internal migration patterns from Twitter data", *Proceedings of the 23rd International Conference on World Wide Web*, pp. 439-444, https://doi.org/10.1145/2567948.2576930. [9]

Zagheni, E., I. Weber and K. Gummadi (2017), "Leveraging Facebook's Advertising Platform to Monitor Stocks of Migrants", *Population and Development Review*, Vol. 43/4, pp. 721-734, https://doi.org/10.1111/padr.12102. [11]

Zinn, S. (2024), *MicSim: Performing Continuous-Time Microsimulation. R package version 2.0.1. h*. [35]

## Notes

[1] Note that, as a result, the dataset is not a balanced and complete matrix of country/year points, with information on migration policy for each country/year. For a country/year there may be multiple rows, if multiple policy changes happened, followed by no data for the next year if no policy change happened. This means that the data need to be manipulated (possibly computing cumulative measures of change against a baseline) to obtain complete matrices.

[2] https://www.census.gov/data/software/dapps.html.

# 6 How should the forecasting model be developed?

## Is it still accurate to use traditional statistical and econometric time series analysis?

The use of econometric models to forecast international migration took off in the 1990s, when discussions about the enlargement of the European Union to include Central European countries began. Forecasting potential labour mobility from new Member States (NMS) was high on the policy agenda. Forecasts had to take into account the transitional arrangements (i.e. restrictions on mobility), some of which remained in place after accession in 2004. Many academic researchers ran econometric models before the anticipated enlargement (Boeri and Brucker, 2001[1]; Dustmann et al., 2003[2]) or after the enlargement (Brücker, Damelang and Wolf, 2009[3]). These regression-based models (see Bijak (2011[4]) for details on various approaches) estimated different dependent variables (e.g. migration flows, migrant stocks, immigration rates) between destination and selected EU candidate countries, conditional on a set of traditional predictors such as labour market outcomes or income differences, GDP per capita and dummy variables denoting geographic or cultural distance. These forecasts failed to correctly anticipate what was going to happen, giving results far lower than actual flows. However, unemployment rates have been confirmed since then as a powerful explanatory factor (Bijak et al., 2019[5]) for all categories of migration, and not only to forecast labour migration. The models employed at the time have nevertheless been criticised for shortcomings of model specification, especially with respect to demographic variables and country-specific effects, which were missing in many studies (Bijak, 2011[4]). In particular, most assumed that past times series were stationary, an assumption that cannot hold in uncertain time (such as the global economic crisis which occurred in the late 2000s) and with such a volatile phenomenon as migration.

An additional problem with using covariates of migration in forecasting models, or in scenarios, is that they also need forecasting, either separate from or with migration. This means that the predictive uncertainty of migration is compounded by the predictive uncertainty of the covariates, and by the uncertain nature of the relationships between migration and its drivers (Bijak, 2011[4]; Barker and Bijak, 2025[6]). Still, this approach continues to be used in forecasts, as well as in migration scenarios, which chart several *possible* rather than likely future trajectories of migration (e.g. Acostamadiedo et al. (2020[7]) or Wiśniowski et al. (2023[8])). Such scenarios can be based on pre-selected trajectories of drivers, perhaps described qualitatively rather than necessarily being quantified. An alternative approach could involve a driver-less statistical approach to scenario-setting, based on the frequency and magnitudes of rare migration events (Bijak, 2024[9]).

In addition to the aforementioned econometric models involving explanatory variables, another important group of migration forecasting approaches can be based on the traditional analysis and extrapolation of time series. It includes standard approaches to time series extrapolation such as Auto-Regressive Integrated Moving Average (ARIMA) models, including Generalised Autoregressive Conditional Heteroskedasticity (GARCH) and Stochastic Volatility (SV) extensions, as well as Vector Auto-Regressive (VAR) models (Barker and Bijak, 2025[6]). Bijak, et al. (2019[10]) assessed migration forecasting approaches using the United Kingdom data with various extrapolation methods and econometric models. Econometric

models evaluated yield poor or only reasonable calibration with middle-sized or even high measurement errors. Traditional extrapolation of time series are also miscalibrated or biased in most cases. This is true when forecasts are based on shorter series of data, or/and when series are non-stationary.

Some migration phenomena are nevertheless sufficiently stationary to be forecasted through traditional time series analysis based on past data only (e.g. family migration, see Box 6.1 and Table 4.1). Therefore, except for these few migration categories that are likely stationary, forecasts of other migration flows increasingly rely on Bayesian or Machine learning models.

### Box 6.1. The use of traditional statistical analysis as a relevant forecasting tool: The example of time series analysis for family migration and survival analysis for naturalisations in the United States

The United States receives more family migration than any other OECD country. To predict the number of family visas issued, USCIS forecasts starting from the number of application forms completed by principal sponsors. Historical data shows the average time between the submission of the application and the actual issue of the visa, depending on the migratory category and the profile of the sponsor.

As time series data for family migration is very stable, forecasts require much less complex models than for “forced” migration. Forecasts are run for each of the specific family migration categories, which are governed by different rules according to sponsor citizenship. In almost all family migration categories, the United States uses standard forecasting models with no predictive variables included, except the numerical limits applying to each category. Possible models, such as exponential smoothing or ARIMA, are generated through R packages. The plausibility of candidate models is then assessed by MAPE (Mean Absolute Percent Error). The model chosen is usually the one with the lowest MAPE level, unless another model is deemed more reasonable by the forecasting unit’s experts. The forecasting unit does not use other predictive variables.

When forecasting family migration in the United States, it is particularly relevant to estimate the number of family members of US citizens. The latter can apply for family reunification for a large set of family members, often without a cap. As a result, naturalisations leading to family immigration have to be forecasted as well.

By using survival analysis, USCIS can follow cohorts of foreign nationals until they are naturalised and thus establish profiles of future Americans and the time they take to obtain nationality. These profiles by country of origin, age, migratory category and length of permanent residence are then assigned a rate of chance of being naturalised, with the rate varying according to procedural “shocks” (more applications before the elections or before the implementation of a higher procedure fee, for instance). Based on historical data for each profile of new American citizens, it is eventually possible to predict the number of family members joining them, as well as the time between naturalisation and application (the vast majority being made within the first six months).

If the use of survival analysis is relevant to better predict family migration from newly naturalised US citizens, it is worth noting that, for family members of US-born citizens, the traditional times series analysis described above are still used, given the stability of that inflow.

**Checklist**:

- Is the phenomenon being forecasted stationary?
  - *Yes*: Consider using basic time series models such as ARIMA or VAR, which can perform well for stationary processes (e.g. family migration).
  - *No*: Traditional statistical and econometric models may have limited predictive power. Alternative methods (e.g. machine learning or mixed approach with expert opinions) should be considered.

## What are Bayesian Models, their principles, and key features?

The contemporary methodological mainstream of statistical migration forecasting largely relies on Bayesian statistical methods as a natural way of describing the predictive uncertainty (e.g. Bijak (2011[4]); (Bijak and Wiśniowski (2010[11]); Azose and Raftery (2015[12]); Welch and Raftery (2022[13])). Two approaches dominate: time series models – univariate or multivariate – and hierarchical models, which also usually include time series components. In either case, past data are used as the primary source of information about the likely future developments of migration processes, augmented by expert opinion. The models can include additional, theory-based covariates (e.g. BSTS) or be atheoretical and entirely data-driven, as in the case of autoregressive models. The main arguments for the latter are that migration theories are not very predictive anyway and any covariates would need to be forecast themselves (Bijak, 2011[4]). Typical practical applications usually involve forecasting total migration inflows and outflows, with increasing literature looking into origin-destination flows (Welch and Raftery, 2022[13])) or flows additionally disaggregated by age (Raymer and Wiśniowski, 2018[14]) or type (Bijak et al., 2019[10]). The use of Bayesian statistical and econometric time series models augmented by expert-based knowledge is indeed particularly useful for non-stationary migration categories with medium volatility, such as labour migration and international student mobility (Table 4.1). This is also a relatively good alternative for highly volatile categories such as asylum and border crossings when machine learning models cannot be implemented easily (Figure 4.1).The key principles and features of Bayesian statistics are discussed elsewhere – an excellent resource for applied work is for example *Bayesian Data Analysis* (Gelman et al., 2013[15]) – but in summary, the main elements of Bayesian inference are as follows. As in other statistical approaches, the point of departure is a set of data, $\mathbf{x}$, which are assumed to be generated by a statistical model with parameters $\theta$. The likelihood of the particular data being generated by a model with given parameters is described by a probability distribution $p(\mathbf{x}|\theta)$, where the symbol '|' denotes that the distribution is conditional on $\theta$. In the Bayesian analysis, the aim is to estimate the *posterior distribution* of the parameters given the observed data, $p(\theta|\mathbf{x})$. Based on the Bayes Theorem, a mathematical relationship dating back to the 18th century, the posterior distribution can be obtained by multiplying the likelihood by the *prior distribution* of the parameters, $p(\theta)$. The posterior distribution is then proportional to the prior distribution multiplied by the likelihood. The prior reflects additional knowledge about the parameters beyond what is contained in the data, and can be subjective. More generally, Bayesian statistical inference is based on a subjective notion of probability as a measure of subjective belief, the initial assessment of which (*prior*) gets continually updated as new data become available, producing new (*posterior*) knowledge.

The main advantage of using Bayesian methods in the context of migration forecasting are sixfold (see Bijak (2011[4])).

- First, for such uncertain processes as migration, Bayesian methods by design integrate different sources of uncertainty and error.
- Second, the approach allows for incorporating extraneous information, for example, information elicited from subject-matter experts, in a formal and coherent way.

- Third, the use of prior information, for example expert-based, allows to correct for some of the common drawbacks of migration data, such as short series, possibly resulting in better-calibrated forecasts.
- Fourth, the underlying subjective definition of probability as a continuously updated measure of belief is more accurate for unique and non-repeatable phenomena than approaches based on probability seen purely as a frequency of events.
- Fifth, Bayesian methods, by producing full probability distributions of the estimates and forecasts, offer a natural departure point for formal statistical decision analysis.
- Sixth, the Bayesian approach offers a natural mechanism for continually updating the forecasts in the light of new data, in line with the best forecasting practice (e.g. Tetlock and Gardner (2015[16])).

**Checklist**:

- Do I need to incorporate expert knowledge into the model?
  - Bayesian methods allow integration of expert-elicited information, especially when data is scarce or incomplete.
- Do I need full probability distributions rather than single-point estimates?
  - Bayesian approaches provide a full predictive distribution, offering a richer understanding of uncertainty.
- What is the best model for regular forecasting updates?
  - Bayesian forecasting is well-suited for iterative updates, making it ideal for dynamic policy environments.
- Am I forecasting labour migration, student migration, asylum flows, or border crossings?
  - Bayesian methods have shown particular effectiveness in forecasting these types of flows due to their flexibility and ability to handle uncertainty.

## How to include contextual knowledge in forecasting: Expert opinion and Delphi surveys

Country experts may be able to offer valuable opinions, which incorporate implicit knowledge about the functioning of migration processes from their field of expertise.

Given that the role of expert judgement in informing prior assumptions can be important for forecasting models, especially for short or non-stationary series of data (such as most regulated migration categories, see Table 4.1 for more details), an important question arises about robust ways of eliciting such information from experts, who are not trained statisticians in general. The contemporary literature on expert elicitation is very broad. Specifically in regards to eliciting uncertain information in a probabilistic manner from subject-matter experts, an important reference work is *Uncertain Judgements* (O'Hagan et al., 2006[17]).

Some important guidelines for the forecasting expert running the elicitation as well as critical discussion on the choice of experts for this exercise are available in the literature (e.g. Rowe and Wright (1999[18])), but the general considerations are as follows. First, experts should be knowledgeable in the subject matter (migration processes), but do not need to be statisticians or forecasters – it is for the elicitation team to explain exactly what is sought from the experts and prepare the right tools (see O'Hagan et al. (2006[17]) and the following section for details). Secondly, the group of experts should be neither too small, nor too large: the recommended group size typically involves between five and 20 experts (Rowe and Wright, 1999[18]). Thirdly, the group of experts should ideally be heterogeneous, so that a variety of views is represented in the study. Depending on the task, this can involve, for example, civil servants from migration

administrations or international organisations, academics from various migration-related disciplines, NGOs representatives or experts from neighbouring countries.

It is worth bearing in mind that in many practical applications, the expert views will not be directly used to *forecast* migration, but rather will inform the forecasting models through *prior* distributions, which will be modified and updated by the data. In other words, the expert input can be an important ingredient of forecasts – but is far from being the only one.

**Checklist**:

- Do I need to mobilise a network of experts?
  - Where data in some areas are limited or uncertain, expert elicitation can significantly enhance the model's outputs by incorporating informed judgment.
- Does the expert group include a diverse range of profiles?
  - Ensure representation from national administrations, international organisations, academia, NGOs, and, where relevant, experts from neighbouring countries.
- Have I included participants who are not data experts?
  - A mix of technical and non-technical perspectives strengthens the breadth of insights and avoids over-reliance on quantitative expertise alone.
- Is the group size manageable?
  - Keeping the group within a range between 5-20 experts helps balance diversity of input with the practicalities of co-ordination and decision making.
- Can the expert network be maintained overtime with limited turn over?

## How to run a Delphi survey and how to include expert replies in the model?

The outcome of elicitation should ideally be expressed in probabilistic terms, to feed in the forecasting models for example through prior assumptions about their parameters. Generally, information elicited from experts can be used to inform migration forecasts. For example, following the 2014 Scottish independence referendum, experts were asked about possible ranges of future migration conditional on different referendum results (Wiśniowski, Bijak and Shang, 2014[19]). In such cases, the experts are typically asked about the ranges of their beliefs about future value of migration flows in a certain year, and their responses are subsequently converted into probability distributions describing the parameters in question. Combinations of these distributions for all experts, for example obtained by calculating weighted "averages" (formally called statistical mixture distributions) provide prior distributions for these mode parameters (Wiśniowski, Bijak and Shang, 2014[19]). Information collected from expert opinions can be used for scenarios, with experts providing probabilities of different scenarios or driver trajectories and the expected numbers of total migrants or flow levels on selected migration categories (Acostamadiedo et al., 2020[7]) or information about possible drivers and the composition of migration flows (Wiśniowski, Kim and Campbell, 2023[8]) (see also Box 6.2 for examples of questions asked to experts). Alternatively, the experts can be asked to comment on the statistical features of the models, such as stationarity, and the possible values of model parameters, ideally using visualisations to aid understanding. In such cases, the experts can be asked about the ranges of their beliefs about certain parameters of a model (for example, an autoregression coefficient) or a variance of the error term. They can also be asked about the probability that any particular model from a given set of models is likely to be true (Bijak and Wiśniowski, 2010[11]). The responses can then be averaged following a similar approach as discussed above. The elicitation can be set as a multi-stage process, with interim feedback, for example following the template of Delphi studies, which have been used for over 50 years for decision support (Dalkey, 1969[20]). In forward-looking migration studies, Delphi methods were pioneered by Drbohlav (1997[21]).

### Box 6.2. Examples of Delphi questionnaires towards migration experts

**Questioning expected numbers of migrants under different scenarios**

IOM (Acostamadiedo et al., 2020[7]) questioned 110 experts on the migration inflows of some categories (Total inflows, labour migration, highly skilled inflows, first asylum applications and irregular border crossings) in the EU-28 (excluding free mobility) until 2030, according to four broad migration scenarios validated by a previous two-round Delphi Survey of 178 experts. After an introduction sentence giving the current number of non-EU inflows, questions were asked as follows:

- "What would be the approximate number in the year 2030 in the EU-28 for each of the scenarios described above?": The answers were completely opened
- "How confident are you about your estimation? Please provide a percentage based on the scale below": The answers were built from 5 modalities, from 80-100% confidence i.e. very confident, to 1-19% confidence, i.e. very unsure.
- "What is the probability of each of the scenarios becoming a reality in 2030 measured by a percentage between 0 (very improbable) and 100 (very probable)? The percentages must add up to 100 across all scenarios. If all scenarios are equally probable, each should have 25"

Questions were asked again on wave 2, after having shown median and mean estimates from wave 1 aggregated responses.

A relatively similar exercise was conducted to assess expected changes in the number of various migration flows from Middle East and North Africa (MENA) to Europe, including family, labour, refugee and return migration through a factorial survey (Boissonneault and Costa, 2023[22]). The expert panel's estimates were made using a selection of vignettes that provided assumptions about the evolution of the general context in origin and destination countries. Each expert was randomly presented with four vignettes out of the 128 vignettes created and asked the following:

- "Compared to 2019, the total number of migrants from the Middle East & North Africa to Europe will be in 2030…" The answers were chosen from 11 modalities, from ÷5 to ×5, through e.g. no change or ×1.5.

The survey also asked the level of confidence about the chosen value and include questions about the expert's profile (thought about the future of migration in the job, familiarity with migration drivers, highest level of education, sector of activity, and years of experience on migration issues). Information on expert profiles may help weight each expert's answers differently according to their expertise on the specific question.

**Questioning possible drivers and composition of migration flows**

The FUME project (Future migration scenarios for Europe) conveyed a Delphi survey among experts on migration policies, where they were asked to predict for 2030 key migration drivers, future migration policies and the composition of migration by gender, level of skills and regions of origins and destination (Wiśniowski, Kim and Campbell, 2023[8]). The questions were the following:

- "In the next 10 years, what drivers/motivations of migration from various regions to the EU will be most important?" The answers were selected from a list of 15 drivers/motivations grouped in 5 categories (demography and education, economy, environment, governance, society, other to be specified by the expert) and the relevant drivers had to be put in a table crossing region of origins (rows) and EU destinations (columns).

- "Which regions will dominate in sending immigrants to the different parts of the EU? Please check up to three routes in the table." The same question was asked for skilled and low-skilled migration, as well as for intra-EU migration and migration from non-EU countries (four questions in total). The reply tables crossed the regional disaggregation applied previously.
- During 2009-2018, the proportion of flows of male immigrants to the EU-27 member states was higher than female migrants (Males=54% vs. Females=46%. Source: Eurostat database). Compared with the above statistics, what gender balance do you expect in the number of immigrants in next 10 years?" Answer modalities included decreased, maintained or increased gender gap.
- "How will the COVID-19 pandemic affect the economic growth in the EU in the next 10 years? Please specify a probability for each of the scenarios making sure that they sum up to 100%." Answer modalities included slow recovery, fast recovery, stagnation and persistent negative growth, in an addition to an open answer to be specified by the expert.

Other questions discussed issues that will be considered by policymakers in the future, migration policy priorities and impact of COVID-19 on migration policies and the society at large. Information built from these answers can help understand the drivers and hypothetical scenarios, which influence experts in their previous answers.

In short, Delphi surveys are voluntary and completely anonymous, with no individual opinions cited at any point. Researchers running the survey always insist there are no "right" or "wrong" answers, as they are interested in opinions in the experts' personal capacity, not in opinions from their affiliated institution (Wiśniowski, Kim and Campbell, 2023[8]). A Delphi survey involves at least two rounds among experts, who are asked to independently answer a series of questions intended to inform the forecasting model. In the second (and any subsequent) round, the study additionally includes anonymised information from the first round, allowing the experts to reconsider their views and thus enabling convergence of opinion. Still, in the context of such an uncertain phenomenon as migration, too much convergence and alignment of views could be potentially seen as problematic; it is natural that the uncertainty of the phenomenon should be reflected in the uncertainty of the expert views. At the same time, the second round of a Delphi survey can clarify the method and ensure shared understanding of what is expected from the elicitation exercise (e.g. Wiśniowski, Bijak and Shang (2014[19])). For that reason, and to allow exchange of views between the experts, it is recommended that the Delphi surveys contain two rounds, with the second one aimed at informed fine-tuning of expert answers following a deliberative process, in which new information may be anonymously shared within the expert group.

Reviewing expert input, including probabilities of different scenarios, requires the same general principles applied as in the case of forecasts, discussed at the end of Chapter 4. Indeed, there is a need for a structured updating process, for example every year or two, augmented by *ad hoc* re-elicitation of the required quantities every time the circumstances radically change (Bijak, 2024[9]). In the context of expert-based input into forecast and scenarios, providing a formal institutional framework for the elicitation process could be valuable. For institutions mandated with making forecasts, establishing formal advisory bodies, and re-evaluating expert judgement both periodically and on an *ad hoc* basis whenever required, can greatly improve the supply of necessary expert-based information for amending the forecasts and scenarios.

Alternatively to traditional Delphi surveys, recent advances in expert elicitation also include online software (MATCH, Morris, Oakley and Crowe (2014[23])), which allows for intuitive probabilistic elicitation from multiple participants in real time by allowing them to "bet" imaginary stakes (with no financial element) on certain future outcomes. In the context of migration, the software was applied by Wiśniowski, Bijak and Shang (2014[19]) to predict migration between Scotland and the rest of the United Kingdom.

**Checklist**:

- How many rounds of Delphi survey are efficient?
  - Conducting at least two rounds of surveys is the best practice. It ensures experts independently provide input across multiple rounds to refine estimates and encourage convergence of opinions.
- Have I clearly defined the type of input required from experts?
  - Experts may be asked about: Expected numbers or evolution of specific migration flows, probabilities of different migration scenarios, forecast estimates or assessments of key migration drivers.
- Do I have a plan to update expert input periodically?
  - Consider refreshing expert opinions annually or biannually, or more frequently if there are major shifts in context or drivers.

## How to develop machine learning models?

Machine learning models are best-fit to use both administrative traditional historical data and innovative data such as digital traces (see Chapter 5 and Box 5.1 for further details on those data). Carammia et al. (2022[24]) extended the use of the Elastic Net algorithm to the context of time series forecasting and applied it to asylum-seekers applications. The idea behind the Elastic Net model (Zou and Hastie, 2005[25]) is to constrain a high-dimensional linear model by adding both a LASSO and a Ridge penalty to take into account predictor selection while maintaining robustness against multicollinearity. The DynENet model proposed by Carammia et al. (2022[24]) introduced a dynamic approach that allows for non-stationary time series (such as those described in Table 4.1) and multiple data frequencies (collected in particular for asylum and border crossings). By employing a rolling time window to train the model, DynENet remains adaptive to shifting patterns in the data, ensuring robust and accurate forecasts. Its parameters are fully explainable as in a standard linear model. Additionally, the moving window approach facilitates the use of the Importance-Frequency (IF) space (Carpi et al., (2022[26]); Qi et al., (2024[27])), which offers insights into the temporal and spatial persistence of selected predictors. The IF space is particularly valuable for identifying variables that consistently influence migration flows, enabling a deeper understanding of the underlying mechanisms driving these flows and that can be used to inform causal models.

Golenvaux et al. (2020[28]) applied Long Short-Term Memory (LSTM) neural network models to forecast legal migration inflows to OECD countries. LSTM models are deep-learning models based on recurrent neural networks (RNNs). They are powerful but typically black-box models by nature. While LSTMs excel at capturing complex temporal dependencies, their interpretability is inherently limited. Typical measures of explainability like the SHapley Additive exPlanations (SHAP (Lundberg and Lee, 2017[29]), Saliency Maps (Simonyan, Vedaldi and Zisserman, 2014[30]), Partial Dependence Plots (Friedman, (2001[31])) can be used in the attempt to understand the relative importance of the predictors. Like most deep-learning models, LSTMs are prone to overfitting. It has been particularly more investigated to forecast highly volatile phenomena, such as forced migration categories. A growing demand to use this approach for regulated migration categories might merge going forward, as little academic work has been done so far for these more stable categories, with few exceptions (Box 6.3).

### Box 6.3. Machine learning approach to forecast outflows of international students from OECD countries

Given that international student mobility is not as volatile as forced migration, little academic work has been done to forecast such flows. Using a dataset of outbound student mobility between 1998 and 2018 from Chinese Taipei to ten OECD countries (the United States, the United Kingdom, Australia, Japan, Canada, France, Germany, New Zealand, Spain and Korea), a recent study (Yang et al., 2020[32]) employed a machine learning approach. The aim of the study was to propose a hybrid model, "FSDESVR," which combines feature selection (FS, using a random forest method to select the most important features) with support vector regression (SVR, as developed by Drucker et al. (1997[33])) and differential evolution (DE as developed by Storn and Price (1997[34])). The experimental results indicate that FSDESVR achieved higher forecasting accuracy than other models tested (ETS, ARIMA, VAR, SVR, and DESVR).

**Checklist**:

- Am I forecasting asylum flows or border crossings?
  - Machine learning models are well-suited for non-stationary phenomena that evolve over time, such as forced migration due to their complexity and unpredictability.
- Are multiple data frequencies involved in the forecast?
  - Machine learning can efficiently handle datasets with varying temporal resolutions (e.g. daily, weekly, monthly inputs).
- Have I considered using the Dynamic Elastic Net model?
  - The Dynamic Elastic Net provides a balance between accuracy and robustness, making it a reliable choice for forecasting complex migration patterns.

## How to incorporate migration drivers into the machine learning models?

Several machine learning and statistical models (for example, BSTS, prophet (Chapter 5), DynENet (see section above), etc.) and some multidimensional time series models (e.g. VAR) explicitly include migration drivers (or push/pull factors) among their predictors. Adaptive models like BSTS and DynENet, tend to select those that are particularly relevant for the given migration flow. Other models, like prophet, keep all of them with estimated weights, and VAR keep them all. Models such as Garch or ARIMA are based solely on historical values of the target migration flow. While BSTS and DynENet are very light in terms of assumptions, VAR models require each predictor to be stationary. For time series analysis at high frequency (daily to monthly), which is usually the case for asylum applications or border crossings, it is important to correctly handle event data (like wars, policy enforcements, etc). Those data are usually sparse variables, or static data (like GDP, diaspora size, etc.), which may induce effects of spurious correlation. Not correctly processing such migration drivers may induce effects of spurious correlation. These exogenous variables do not cause much trouble in static models (like regression models, gravity models, etc.) but can be distortive in VAR models.

**Checklist:**

- Have I identified relevant migration drivers to include in the model?
  - Migration drivers may include economic indicators, conflict data, visa policy changes, demographic trends, or climate-related variables.

- Am I using an adaptive model capable of selecting key drivers?
  - Models such as BSTS (Bayesian Structural Time Series) and Elastic Net can automatically identify and weight the most relevant migration drivers.
- Is the model designed to update as new information becomes available?
  - Adaptive models are particularly suited for dynamic environments, where the influence of different drivers may shift over time.

## How to model causal influence in migration forecasting?

Causal influence in migration forecasting involves identifying and analysing the factors that drive migration to improve predictive models, or it implies examining how policies may impact future migration flows learning from previous facts.

In his manifesto arguing for causal migration forecasting, Willekens (2019[35]) postulated 12 ideal characteristics that such causal forecasts should exhibit. Importantly, the underlying forecasting models should be designed at the individual level (examples include agent-based models (ABM) or microsimulations), with modelled individuals described by a range of attributes and their own life histories, of which migration is one important part (see Courgeau (1985[36])). Crucially, the modelled processes need to be stochastic, which constitutes one of several important sources of uncertainty that need to be acknowledged in the forecasts. As with other uncertain migration forecasts (see Bijak (2011[4]) for an overview), Willekens (2019[35]) also argues that in causal models, the Bayesian statistical approach provides a natural language for describing the uncertainty of the underlying micro- and macro-level processes.

There are several examples of successful models of causal mechanisms driving various migration processes, chiefly implemented through agent-based simulations based on individual-level decision rules (e.g. Kniveton, Smith and Wood (2011[37]) for climate-induced migration; Naivinit et al. (2010[38]) for labour migration; Suleimenova and Groen (2020[39]) for asylum migration). While there is agreement that such models can help shed light on the explanations behind the observed trends and patterns of migration, their predictive capabilities are much more contentious. Once the many unknown elements and parameters in such models are identified and quantified, especially in relation to the rules driving human behaviour (see Klabunde and Willekens (2016[40])), the resulting predictive uncertainty can quickly become too large to be useful for meaningful forecasting applications (Bijak, 2022[41]). Therefore, ABMs are generally not used as predictive models, as they are deemed computationally expensive, not easy to accurately implement, difficult to parameterise and dependent on arbitrary assumptions (Hinsch and Bijak, 2023[42]).

On the other side of causal interpretation of effects, there is a growing interest in synthetic control methods, machine learning approaches and more general stochastic processes settings (beyond the traditional time-series approach) (See Chapter 9). A parallel and less explored set of models is the pure Directed Acyclic Graph (DAG) approach (Pearl, 2009[43]) (See Box 6.4).

**Checklist**:

- Am I aiming to identify causal mechanisms behind migration drivers?
  - Agent-based models are effective for exploring causality but are generally not suited for predictive forecasting.
- Are there alternative methods for analysing causal influence?
  - Synthetic control methods and machine learning approaches are increasingly used to assess causal effects in migration research.
- Is it important to balance explanatory power with predictive performance?
  - Consider combining causal modelling approaches with forecasting techniques where appropriate, depending on the policy need.

### Box 6.4. Directed Acyclic Graph

The DAG approach relies on the identification of graph to express causal relationships. In a DAG the causal structure is represented through nodes and edges. The nodes represent variables relevant to, for example, migration, such as economic conditions, political instability, social networks, migration policies, and environmental changes, while the edges represent causal relationships that are determined based on theoretical knowledge, domain expertise, or statistical discovery algorithms.

The theoretical advantages of DAGs are that:

1. they clarify causal mechanisms, improving the interpretability of forecasts;
2. they ensure that only the relevant predictors are included in the forecasting model, avoiding spurious correlations and.
3. confounding is explicitly addressed, enhancing the robustness of predictions.

Still, DAGs require integration with time series models (e.g. ARIMA, BSTS) or other machine learning techniques for prediction as mentioned above, because DAGs do not possess an explicit prediction capability. Best practices involve integrating the discovered DAG with forecasting models like Bayesian networks, dynamic causal models, or machine learning frameworks to propagate causal effects into the future. As far as DAG discovery is concerned, a few wide-spread algorithms exist: those based on testing conditional independence (see, e.g. Spirtes and Glymour, (1991[44]); Spirtes et al., (2001[45]), or Dynamic Bayesian Networks (Murphy, 2002[46]) to name a few.

### Figure 6.1. A hypothetical DAG illustrating a sequence of causal relationships

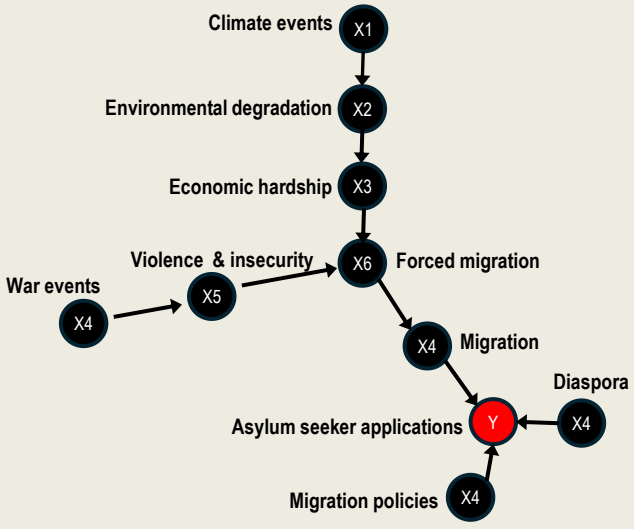



Figure 6.1 shows a made-up example of a DAG describing a sequence of causal relationships among predictor variables (X's: climate/war events, economic situation, displacement, diaspora, migration policies) and the target variable "Asylum Seekers Applications" (Y). These causal relationships are being tested through causal inference techniques. The DAG is the hypothesised state of the system, then causal inference methods will validate the existence of the edges.

## References


Acostamadiedo, E. et al. (2020), *Assessing Immigration Scenarios for the European Union in 2030 – Relevant, Realistic and Reliable?*, International Organization for Migration and the Netherlands Interdisciplinary Demographic Institute. [7]

Azose, J. and A. Raftery (2015), "Bayesian Probabilistic Projection of International Migration", *Demography*, Vol. 52/5, pp. 1627-1650, https://doi.org/10.1007/s13524-015-0415-0. [12]

Barker, E. and J. Bijak (2025), "Mixed-frequency VAR: a new approach to forecasting migration in Europe using macroeconomic data", *Data & Policy*, Vol. 7, https://doi.org/10.1017/dap.2024.82. [6]

Bijak, J. (ed.) (2024), *From Uncertainty to Policy: A Guide to Migration Scenarios*, Edward Elgar Publishing, https://doi.org/10.4337/9781035319800. [9]

Bijak, J. (2022), *Towards Bayesian Model-Based Demography*, Springer International Publishing, Cham, https://doi.org/10.1007/978-3-030-83039-7. [41]

Bijak, J. (2011), *Forecasting International Migration in Europe: A Bayesian View*, Springer Netherlands, Dordrecht, https://doi.org/10.1007/978-90-481-8897-0. [4]

Bijak, J. et al. (2019), "Assessing time series models for forecasting international migration: Lessons from the United Kingdom", *Journal of Forecasting*, Vol. 38/5, pp. 470-487, https://doi.org/10.1002/for.2576. [10]

Bijak, J. et al. (2019), "Assessing time series models for forecasting international migration: Lessons from the United Kingdom", *Journal of Forecasting*, Vol. 38/5, pp. 470-487, https://doi.org/10.1002/for.2576. [5]

Bijak, J. and A. Wiśniowski (2010), "Bayesian Forecasting of Immigration to Selected European Countries by using Expert Knowledge", *Journal of the Royal Statistical Society Series A: Statistics in Society*, Vol. 173/4, pp. 775-796, https://doi.org/10.1111/j.1467-985x.2009.00635.x. [11]

Boeri, T. and H. Brucker (2001), "Eastern Enlargement and EU Labour Markets", *World Economics*, pp. 49-68. [1]

Boissonneault, M. and R. Costa (2023), "Experts' assessments of migration scenarios between the Middle East & North Africa and Europe", *Scientific Data*, Vol. 10/1, https://doi.org/10.1038/s41597-023-02532-1. [22]

Brücker, H., A. Damelang and K. Wolf (2009), *Forecasting potential migration from the new member states into the EU-15. Review of literature, evaluation of forecasting methods and forecast results*. [3]

Carammia, M., S. Iacus and T. Wilkin (2022), "Forecasting asylum-related migration flows with machine learning and data at scale", *Scientific Reports*, Vol. 12/1, https://doi.org/10.1038/s41598-022-05241-8. [24]

Carpi, T. et al. (2022), "The Impact of COVID-19 on Subjective Well-Being: Evidence from Twitter Data", *Journal of Data Science*, pp. 761-780, https://doi.org/10.6339/22-jds1066. [26]

Courgeau, D. (1985), “Interaction between spatial mobility, family and career life-cycle: A French survey”, *European Sociological Review*, Vol. 1/2, pp. 139-162, https://doi.org/10.1093/oxfordjournals.esr.a036382. [36]

Dalkey, N. (1969), “An experimental study of group opinion”, *Futures*, Vol. 1/5, pp. 408-426, https://doi.org/10.1016/s0016-3287(69)80025-x. [20]

Drbohlav, D. (1997), “Migration Policy Objectives for European East-West International Migration”, *International Migration*, Vol. 35/1, pp. 85-108, https://doi.org/10.1111/1468-2435.00005. [21]

Drucker, H. et al. (1997), *Support vector regression machines, advances in neural information processing systems*. [33]

Dustmann, C. et al. (2003), *The impact of EU enlargement on migration flows*, Home Office. [2]

Friedman, J. (2001), “Greedy function approximation: A gradient boosting machine.”, *The Annals of Statistics*, Vol. 29/5, https://doi.org/10.1214/aos/1013203451. [31]

Gelman, A. et al. (2013), *Bayesian Data Analysis*, Chapman and Hall/CRC, https://doi.org/10.1201/b16018. [15]

Golenvaux, N. et al. (2020), “An LSTM approach to Forecast Migration using Google Trends. :”, *ArXiv*, https://arxiv.org/abs/2005.09902. [28]

Hinsch, M. and J. Bijak (2023), “The Effects of Information on the Formation of Migration Routes and the Dynamics of Migration”, *Artificial Life*, Vol. 29/1, pp. 3-20, https://doi.org/10.1162/artl_a_00388. [42]

Klabunde, A. and F. Willekens (2016), “Decision-Making in Agent-Based Models of Migration: State of the Art and Challenges”, *European Journal of Population*, Vol. 32/1, pp. 73-97, https://doi.org/10.1007/s10680-015-9362-0. [40]

Kniveton, D., C. Smith and S. Wood (2011), “Agent-based model simulations of future changes in migration flows for Burkina Faso”, *Global Environmental Change*, Vol. 21, pp. S34-S40, https://doi.org/10.1016/j.gloenvcha.2011.09.006. [37]

Lundberg, S. and S. Lee (2017), *A Unified Approach to Interpreting Model Predictions.*. [29]

Morris, D., J. Oakley and J. Crowe (2014), “A web-based tool for eliciting probability distributions from experts”, *Environmental Modelling & Software*, Vol. 52, pp. 1-4, https://doi.org/10.1016/j.envsoft.2013.10.010. [23]

Murphy, K. (2002), *Dynamic Bayesian Networks: Representation, Inference and Learnin*. [46]

Naivinit, W. et al. (2010), “Participatory agent-based modeling and simulation of rice production and labor migrations in Northeast Thailand”, *Environmental Modelling & Software*, Vol. 25/11, pp. 1345-1358, https://doi.org/10.1016/j.envsoft.2010.01.012. [38]

O’Hagan, A. et al. (2006), *Uncertain Judgements: Eliciting Experts’ Probabilities*, Wiley, https://doi.org/10.1002/0470033312. [17]

Pearl, J. (2009), *Causality: Models, Reasoning and Inference (*, Cambridge University Press, USA. [43]

Qi, H. et al. (2024), *Modeling Climate-Induced Refugee Migration: An Explainable Machine Learning Approach*, Springer Science and Business Media LLC, https://doi.org/10.21203/rs.3.rs-4931065/v1. [27]

Raymer, J. and A. Wiśniowski (2018), "Applying and testing a forecasting model for age and sex patterns of immigration and emigration", *Population Studies*, Vol. 72/3, pp. 339-355, https://doi.org/10.1080/00324728.2018.1469784. [14]

Rowe, G. and G. Wright (1999), "The Delphi technique as a forecasting tool: issues and analysis", *International Journal of Forecasting*, Vol. 15/4, pp. 353-375, https://doi.org/10.1016/s0169-2070(99)00018-7. [18]

Simonyan, K., A. Vedaldi and A. Zisserman (2014), *Deep inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps*. [30]

Spirtes, P. and C. Glymour (1991), "An Algorithm for Fast Recovery of Sparse Causal Graphs", *Social Science Computer Review*, Vol. 9/1, pp. 62-72, https://doi.org/10.1177/089443939100900106. [44]

Spirtes, P., C. Glymour and R. Scheines (2001), *Causation, Prediction, and Search*, The MIT Press, https://doi.org/10.7551/mitpress/1754.001.0001. [45]

Storn, R. and K. Price (1997), , *Journal of Global Optimization*, Vol. 11/4, pp. 341-359, https://doi.org/10.1023/a:1008202821328. [34]

Suleimenova, D. and D. Groen (2020), "How Policy Decisions Affect Refugee Journeys in South Sudan: A Study Using Automated Ensemble Simulations", *Journal of Artificial Societies and Social Simulation*, Vol. 23/1, https://doi.org/10.18564/jasss.4193. [39]

Tetlock, P. and D. Gardner (2015), *Superforecasting: The Art and Science of Prediction*, New York: Crown. [16]

Welch, N. and A. Raftery (2022), "Probabilistic forecasts of international bilateral migration flows", *Proceedings of the National Academy of Sciences*, Vol. 119/35, https://doi.org/10.1073/pnas.2203822119. [13]

Willekens, F. (2019), "Towards causal forecasting of international migration", *Vienna Yearbook of Population Research*, Vol. 1, pp. 199-218, https://doi.org/10.1553/populationyearbook2018s199. [35]

Wiśniowski, A., J. Bijak and H. Shang (2014), "Forecasting Scottish Migration in the Context of the 2014 Constitutional Change Debate", *Population, Space and Place*, Vol. 20/5, pp. 455-464, https://doi.org/10.1002/psp.1856. [19]

Wiśniowski, A., J. Kim and G. Campbell (2023), *Delphi Study – Future Migration Scenarios for Europe*. [8]

Zhang, J. (ed.) (2020), "Forecasting outbound student mobility: A machine learning approach", *PLOS ONE*, Vol. 15/9, p. e0238129, https://doi.org/10.1371/journal.pone.0238129. [32]

Zou, H. and T. Hastie (2005), "Regularization and Variable Selection Via the Elastic Net", *Journal of the Royal Statistical Society Series B: Statistical Methodology*, Vol. 67/2, pp. 301-320, https://doi.org/10.1111/j.1467-9868.2005.00503.x. [25]

# 7 How can the model be evaluated?

To evaluate the performance of migration forecasting models and to adapt them to evaluation results, the following framework can be applied, addressing robustness, metrics, and back-testing approaches. In the literature, the precise definitions of the related terms vary, but in this section, we consider evaluation to encompass a group of processes ensuring the quality of the forecasting models. It includes validation of the model performance – here understood as testing the alignment of model results with observed data – as well as calibration. In this particular context, calibration looks at the error measures (predictive intervals), and how they correspond to the relative frequencies of observations of different magnitudes seen in the data series. Machine Learning models, and hence statistical models, learn their parameters from historical data. While traditional statistical models are usually parsimonious in the number of parameters, machine learning models can encompass many. Both cases have a risk of over-fitting, but the risk can be much higher for some classes of machine learning models, such as those based on deep learning techniques. In light of this risk, performance evaluation and back-testing should be applied to both classes of models.

## How to validate model performance?

The forecasting models, once built and estimated, need to have their performance validated. A number of tools (Box 7.1) and methods can be used to do so. A standard approach is to look at past errors, which can be compared to some benchmark, for example a constant forecast or one obtained by running a simple extrapolation, such as exponential smoothing or a basic ARIMA model, which relate the current magnitude of the processes being forecast to their past. The exercise typically involves running the models and the benchmarks on a shortened time series of data, for example setting aside the last five observations, and checking how well the forecasts were able to "predict" the data points that were set aside. This is known as validating the models *ex post*, or *out of sample*. Indeed, the data that are set aside are not a part of the sample and do not inform the model, so they can be used for independent validation. It differs from the *ex-ante*, or *in-sample* validation, which looks at how well the model fits the specific sample of data, and which can be measured by, for example, the determination coefficient ($R^2$), various information criteria (AIC, BIC), and so on. For a given model, the *ex-ante* approach allows one to approximate the errors (how much the actual data differ from the predictions) based only on the data sample. However, the ex ante validation does not constitute, *per se*, a proof of the good predictive power of the model.

In turn, metrics involved in *ex post* validation include a range of error measures, the lower the error measures, the better.

**Traditional Metrics:** Hyndman and Athanasopoulos (2018[1]) give a basic introduction to Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and other traditional metrics used in forecasting. RMSE and MAE evaluate the quadratic and absolute distance, respectively, between the actual value of the target variable and its forecasts. RMSE or MAE are used for continuous migration flow forecasts. F1 Score, Precision, and Recall are particularly useful for "classification-based" migration forecasting. For example, a model that predicts "high", "medium" and "low" migration flows rather than volumes of migration is a classification-based model. *"Precision"* measures how often the model's forecasts for a class are correct, while *"recall"* checks how well the model captures all actual instances of that class. The F1 score balances

both, ensuring accuracy and completeness. In “multiclass classification”, these metrics are calculated for each class and then averaged for overall performance.

**Time Series-Specific Metrics:** Hyndman (2006[2]) introduces Mean Absolute Scaled Error (MASE) and compares it with other metrics, such as MAPE. The Mean Absolute Percentage Error (MAPE) is commonly used for assessing forecast accuracy relative to observed values; the MASE is particularly useful when comparing across models or scales. These statistics are similar to the MAE but consider the error in relative or standardised terms.

**Causal Validation:** Runge et al., (2023[3]) discusses causal validation methods, particularly for dynamic and time-series data in the context of causal forecasting. Comparing predicted impacts of interventions (e.g. policy changes) against known outcomes tests whether causal assumptions hold. This approach is applicable only when one or more intervention periods are well separated, although in many situations, interventions are often staggered.

*Ex post* errors can provide very useful information on whether there is a systematic bias in the forecasts, by how much do the models miss the subsequent real-life developments, and crucially, could the simple benchmarks have done a better job. In fact, the comparative aspect of forecast errors is one of the foundations of validating a range of models and choosing the best one. Ideally, evaluating forecasts should involve both aspects: present the *ex-ante* errors or associated measures to demonstrate that the model fits the data reasonably well, and *ex post* errors to show its predictive capabilities.

### Box 7.1. An open-source toolkit to validate forecasting models: The SEAVEA project

Research teams from Brunel University London and University College London have developed since 2021 an open source toolkit dedicated to VVUQ (Verification, Validation and Uncertainty Quantification). The SEAVEA toolkit (Software Environment for Scalable & VVUQ-evaluated Exascale Applications) includes model verification (confirmation that the mathematical model and corresponding algorithm have been coded correctly), validation (of results compared to qualitative and quantitative measures which apply a validation metric) and uncertainty quantification (understanding the origins of and assessing the magnitudes of the errors which accompany computer simulations, whether epistemic or aleatoric).

Although originally developed for highly intensive exascale computer-based simulations, the SEAVEA Toolkit supports all fields of work that require event simulations to make predictions for decision making, such as fusion energy, climate science, epidemiology, medicine, aerospace, and migration. It then fits with exascale computing models as well as models using much smaller computing resources and smaller number of data points, such as those usually run for migration forecasts. SEAVEA includes a set of interoperable and advanced components (e.g. Easy VVUQ) in order to support modellers making their simulation more reproducible, reliable, and scientifically credible, eventually minimising the expense and time needed to perform calculations. All tools being open-source and having their own GitHub repositories, SEAVEA allows anyone to propose contributions and modifications to the toolkit.

Source: https://sites.google.com/view/seavea-toolkit/; https://excalibur.ac.uk/projects/seavea/.

**Checklist:**

- Have I assessed the model’s predictive accuracy using standard metrics?
  - Use indicators such as RMSE (Root Mean Squared Error) and MAE (Mean Absolute Error) to evaluate forecast performance.
- Have I conducted an ex-post validation?

  - Run the model and relevant benchmarks on a shortened time series to assess how well it predicts out-of-sample data.
- Have I conducted an ex-ante validation?
  - Use measures such as $R^2$ (coefficient of determination) and AIC (Akaike Information Criterion) to test how well the model fits the available data.
- Have I interpreted both predictive power and model fit appropriately?
  - Both ex post and ex ante assessments are needed to demonstrate the model's overall robustness and reliability.

## How to calibrate the model?

For probabilistic forecasting models, another important aspect of evaluation is the calibration of predictive intervals. The aim here is to check whether the probabilistic measures produced by the model are reasonable, neither too narrow nor too wide. This is done by checking whether the predictive intervals predicted by the model ex ante are broadly in line with those obtained ex post, based on the data that have to be set aside. Typically, the 50-per cent predictive intervals produced by the model should cover the future values roughly 50% of the time, 80-per cent intervals roughly 80% of the time, and so on. One simple way of doing that is by calculating the empirical frequencies of instances, for which the real values fall inside (or outside) various predictive intervals (see Gneiting, Balabdaoui and Raftery (2007[4]) and Czado, Gneiting and Held (2009[5]), who refer to this approach as marginal calibration). For instance, if we set aside ten observations to carry out an out-of-sample forecast assessment, for well-calibrated forecasts, we would expect five out of ten of these observations to fall within the 50-per cent predictive intervals, and eight out of ten within the 80-per cent ones.

Predictive intervals that are too wide compared to their nominal probabilities (for example, if 50-per cent intervals cover 90% of observations) are too conservative and can lead to unnecessary hedging against some extreme possibilities, while intervals that are too narrow (for example, when 80-per cent intervals only cover 25% of observations) are unrealistically optimistic, and can lead to being underprepared. This is in addition to the question of which intervals to choose for calibration. Given the paucity of migration data and the fact that the extremes of the distribution may be very difficult to estimate precisely, there is an argument for calibrating the 50- and 80-per cent intervals (Bijak et al., 2019[6]), rather than, for example, 99-per cent ones, although in practical applications, the 95-per cent intervals tend to be used in calibration exercises as well (see Bijak (2011[7]), or Azose and Raftery (2015[8])). Regardless of this choice, significantly miscalibrated forecasts require attention: they can indicate that either the chosen model is not appropriate for the data, which would necessitate specifying it again, or that the assumptions about the error term are either too narrow or too wide. This would require revisiting the prior assumptions for the related model parameters, such as the variance of the error term, possibly within a broader framework if iterative expert elicitation, as discussed in Chapter 6.

Forecast calibration has an important place in the methodological literature on forecasting. Brocker and Smith (2007[9]) explain the use of calibration curves and reliability assessments for probabilistic forecasting. Evaluating how well predicted probabilities align with observed outcomes requires the use of calibration curves or scores. Calibration curves show whether a model's predicted probabilities align with actual outcomes, indicating whether it is over- or underconfident. Gneiting, Balabdaoui and Raftery (2007[4]) and Czado, Gneiting and Held (2009[5]) propose a set of different *scoring rules* for evaluating probabilistic forecasts with respect to both their sharpness and calibration. As an example, for predicting binary events, a simple scoring rule is the Brier score (Brier (1950[10]); see Tetlock and Gardner (2015[11])). The Brier score is the sum of squared differences between the predicted probabilities and the actual binary outcome (1 = the event happened, 0 = it did not): the lower the score, the better, and values below 0.25 ($0.5^2$, indicating a random guess) show the advantage of a certain model or approach. Gneiting and Raftery

(2007[12]) also discusses the use of Continuous Ranked Probability Score (CRPS) as a proper scoring rule to help analyse, for instance, how well the model predicts not just point estimates but entire distributions. The CRPS measures the accuracy of probabilistic forecasts by comparing the predicted cumulative distribution to the actual outcome, with lower scores indicating better predictions. In applied migration forecasting, assessing calibration is becoming a part of the standard evaluation toolkit (Bijak et al., 2019[6]; Welch and Raftery, 2022[13]).

**Checklist:**

- Have I included predictive intervals as part of the model evaluation?
  - It is recommended to calibrate both 50% and 80% predictive intervals to assess the model's reliability.
- Do the observed outcomes fall within the expected ranges while conducting out-of-sample forecast assessments?
  - For example, out of ten observations, approximately five should fall within the 50% interval and eight within the 80% interval.
- Have I adjusted the model if the intervals are consistently too narrow or too wide?
  - Poor calibration may indicate that the model underestimates or overestimates uncertainty and may require refinement.

### *How to adapt the model based on evaluation results or as data come by?*

**Error Diagnostics:** Chatfield (2000[14]) discusses residual diagnostics and error decomposition for time-series models. Examining residuals (e.g. the difference between predicted and actual values) is key to identifying systematic biases (e.g. underestimation during economic shocks or overfitting to seasonal migration patterns). This technique is inherited from standard linear regression analysis but in the context of time-series may reveal additional patterns that an updated version of the model can take into account. Error decomposition techniques are particularly valuable in this context as they help to isolate issues like trend misalignment or spurious seasonality.

**Model Refinement:** Zou and Hastie (2005[15]) and Carammia et al. (2022[16]) apply regularised methods in machine learning and importance extraction. Model refinement may require introducing new features (e.g. policy indices, environmental indicators) or reweighting features based on importance analysis (e.g. SHAP values). Models may also be regularised more aggressively (e.g. elastic net method) or by simplifying overcomplex architectures to reduce overfitting. This refinement can be part of the model building in adaptive models, or be a separate ex-post step.

**Dynamic Adaptation:** Bontempi et al. (2013[17]) introduce online learning techniques for evolving datasets. O*nline learning* techniques are used to update models continuously as new data arrives. Other adaptation techniques incorporate scenario-specific tuning based on historical analogues.

While model refinement is used *after* a model has been fitted to the whole historical data, i.e. when the assumption is that the model does not change with time, online learning and scenario-specific tuning are applied dynamically in adaptive models during the fitting. While Bontempi et al. (2013[17]) propose to update the model at each new observation, Carammia et al. (2022[16]) apply the idea to moving windows assuming that migration processes change over time but have some persistency over short periods.

**Checklist:**

- Has the model failed validation or shown signs of poor performance?
  - Examine residuals to detect patterns or structure that the model may have missed. For Machine learning models, they can be improved by introducing new features or adjusting the weights of existing ones.

- Have I considered techniques for continuous model improvement?
  - Online learning methods allow models to be updated in real time as new data become available, even before formal re-evaluation.
- Have I documented changes to the model following refinement?
  - Keeping track of updates ensures transparency and supports reproducibility in the forecasting process.

## How to assess robustness?

**Robustness Testing:** Barredo Arrieta et al. (2020[18]) review several robustness testing techniques, including adversarial testing, for AI models. These techniques include testing sensitivity of predictions to noise, missing data or out-of-distribution inputs. Performing adversarial tests requires introducing controlled perturbations to inputs.

**Stress Testing:** Kilian and Lütkepohl (2017[19]) discuss stress testing in economic and migration forecasting through extreme scenario simulations. Simulating extreme scenarios (e.g. sudden political crises or environmental disasters) is important to evaluate model stability.

**Cross-Temporal Validation:** Tashman (2000[20]) reviews cross-temporal and rolling origin validation methods. Performing back-testing across multiple historical periods helps to verify consistency over time, especially during disruptive events. This approach is a by-product of the workflow described in Carammia et al. (2022[16]).

**Scenario Analysis:** Predictions under hypothetical scenarios must be evaluated to assess the model's adaptability to structural changes. This involves testing the model's performance under various "what-if" conditions to understand its robustness and reliability. By simulating different hypothetical scenarios, such as economic shifts, policy changes, or unexpected events, it is possible to evaluate how well the model adapts to structural changes. This helps identify potential weaknesses, refine assumptions and improve decision making under uncertainty. Literature on this topic is still limited, as scenario analysis is very ad hoc and also requires the ability to generate scenarios through expert elicitation.

In all the above cases, the performance metrics summarised in the previous section are handy tools to assess the robustness of the forecasts.

**Checklist:**

- Am I using the same metrics for robustness assessment as for performance validation?
  - Metrics such as RMSE, MAE, $R^2$, and AIC can be applied to evaluate both model accuracy and robustness.
- Have I tested the model under different conditions or scenarios?
  - Robustness checks may involve varying input data, timeframes, or assumptions to ensure consistent performance.
- Does the model perform reliably across different subgroups or data segments?
  - Consistent results across contexts increase confidence in the model's generalisability.
- Have I documented how the model responds to changes in assumptions or input data?
  - This helps determine the extent to which the model's results are sensitive to external factors.

## How to back-test ML/Stat models?

The list of traditional and alternative approaches to model performance validation discussed above are all also, in practice, back-testing strategies, as they can be implemented only looking at historical data. Model performance validation and back-testing indeed use the same metrics. More precisely, *model performance evaluation* is an assessment of how well a model predicts outcomes based on historical data only. *Back-testing*, on the other hand, analyses how the model would have performed in a real-world past scenario and can also be performed on fictitious historical data. Compared to evaluation, back-testing is often used in finance or risk modelling, but much less in migration studies. It merely applies the model to historical data without retraining, mimicking real-time decision making and checking if it would have led to effective outcomes.

### *Classic back-testing for forecast models*

**Rolling Forecast Origin:** Bergmeir et al. (2016[21]) discusses rolling origin validation for time-series models. It takes two forms: splitting data into rolling windows to evaluate forecasts iteratively over different periods, or testing models' ability to generalise across time and adjust for seasonality or structural breaks. It is a mix of both the cross-temporal validation of Tashman (2000[20]) and the adaptive method described in Carammia et al. (2022[16]).

### *Model diagnostics in ML/Stat models*

**Residual Analysis:** Durbin and Koopman (2012[22]) make a general introduction to residual diagnostics for statistical time-series models. Those diagnostics evaluate whether residuals are white noise (uncorrelated, normally distributed, with constant variance). They can also detect overfitting by analysing residual patterns over the training and validation datasets. This can be seen as an extension of the error diagnostic (Chatfield, 2000[14]).

**Counterfactual Validation:** Pearl (2009[23]) provides foundational methods for counterfactual validation in causal modelling. Causal models are used to evaluate counterfactual scenarios (e.g. simulating no-policy-change baselines). Counterfactual validation is not suitable for models that apply many predictors.

**Checklist:**

- Is back-testing applied in migration forecasting studies?
  - While underused, back-testing can provide valuable insights into model reliability and should be more widely adopted.
- Have I used back-testing to evaluate model performance over historical periods?
  - Back-testing applies the model to past data to assess how well it would have predicted known outcomes.

## How can expert knowledge and scenarios be back-tested?

Expert knowledge is not usually back-tested, nor is it generally expected to be back-tested, unlike what is usual for statistical or machine learning models. Expert knowledge often involves subjective judgment, qualitative insights, or hypothetical conditions that may not have occurred historically and may not ever occur in the future. Although there is no definitive literature on this point, expert knowledge can be stress-tested or validated using historical analogies, simulations, or comparative analyses. More precisely, a validity framework can be discussed in the following terms:

i. **Historical alignment:** if a past event closely resembles an expert-defined scenario, it would be possible to check whether the expert's reasoning would have led to accurate predictions based on historical outcomes;

ii. **Counterfactual analysis:** using past data, it would be possible to test "what-if" scenarios by modifying certain variables to see if expert-driven insights would have correctly anticipated alternative outcomes;

iii. **Simulation-based back-testing:** if historical data isn't available for a scenario, Monte Carlo simulations or agent-based models can be used to generate possible outcomes and test expert assumptions under varying conditions;

iv. **Benchmarking against model predictions:** comparing expert-driven insights with statistical models applied to past data to highlight inconsistencies and/or refine assumptions.

A notable exception in the literature is Imbens and Rubin (2015[24]), who provide the basic causal simulations for intervention evaluation. Validating against historical scenarios curated by experts ensures that model behaviour aligns with expected outcomes.

**Synthetic Scenario Testing:** Fraccascia et al. (2018[25]) explore scenario testing under synthetic conditions in complex system modelling. Similar ideas can be applied to the context of migration forecasting by constructing, for instance, synthetic scenarios reflecting plausible future conditions (e.g. economic downturns, climate shocks) and assessing model responses.

**Simulated Interventions:** Model robustness can be tested by introducing artificial interventions (e.g. introducing hypothetical policy changes) and comparing outcomes against expert expectations. See again Imbens and Rubin (2015[24]).

**Expert-Elicited Feedback:** Tetlock and Gardner (2015[11]) is a classic reference and it highlights how expert knowledge can enhance model evaluation through structured feedback and scenario development. Running models with and without expert-suggested features or constraints allows one to evaluate the impact on accuracy and explainability.

**Checklist:**

- Have I considered testing the influence of expert knowledge on the model?
  - While expert input is rarely back-tested, it is important to evaluate its impact where possible. By running the model in both configurations to assess how expert knowledge affects accuracy and predictive power.

### Box 7.2. How the United States evaluates forecasted numbers

The United States Office of Homeland Security Statistics (OHSS) Migration Analysis Center (MAC) conducts various types of evaluations to forecast unauthorised migration at the Southwest US-Mexico border.

First, historical evaluations are performed to compare forecasts to actual outcomes on a rolling basis by lookahead month. These evaluations use statistics to assess both accuracy (Median Absolute Percent Error) and bias (Median Percent Error), as well as overall performance for up to a six-month lookahead period. Generally, the one-month lookahead is the most accurate compared to the three-month or six-month lookaheads. However, in some cases, the median absolute per cent error may be lower for the six-month lookahead if there has been substantial encounter volatility, which creates additional challenges in forecasting accurately. Diagnostics are used to identify sub-groups with the greatest potential for improvement and to test model adjustments before implementing

changes. For example, the OHSS MAC identified Cuba, Haiti, Nicaragua, and Venezuela (CHNV) as countries where models required improvement. In January 2023, OHSS MAC removed the trend component from the models for these four countries, which resulted in improved forecast accuracy. While one-off fluctuations – such as Venezuelan migration surges in early 2023 – may still be missed, the models adjust to new realities after 2-3 months of a shift in migration patterns.

Additionally, mid-month evaluations are conducted to compare the one-month lookahead forecast to the encounter actuals for that month, extrapolated from the month-to-date daily average. Lastly, model comparison evaluations assess the historical performance of multiple model specifications. Confidence intervals are calculated using the formula for forecast intervals from the forecast package in R. However, the assumption of growing uncertainty over time, typical of a true forecast interval, is not applied. Since these interval bounds are primarily used for projections or planning lines, maximising the precision of the intervals has not been a priority.

## References


Azose, J. and A. Raftery (2015), "Bayesian Probabilistic Projection of International Migration", *Demography*, Vol. 52/5, pp. 1627-1650, https://doi.org/10.1007/s13524-015-0415-0. [8]

Barredo Arrieta, A. et al. (2020), "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI", *Information Fusion*, Vol. 58, pp. 82-115, https://doi.org/10.1016/j.inffus.2019.12.012. [18]

Bergmeir, C., R. Hyndman and J. Benítez (2016), "Bagging exponential smoothing methods using STL decomposition and Box–Cox transformation", *International Journal of Forecasting*, Vol. 32/2, pp. 303-312, https://doi.org/10.1016/j.ijforecast.2015.07.002. [21]

Bijak, J. (2011), *Forecasting International Migration in Europe: A Bayesian View*, Springer Netherlands, Dordrecht, https://doi.org/10.1007/978-90-481-8897-0. [7]

Bijak, J. et al. (2019), "Assessing time series models for forecasting international migration: Lessons from the United Kingdom", *Journal of Forecasting*, Vol. 38/5, pp. 470-487, https://doi.org/10.1002/for.2576. [6]

Bontempi, G., S. Ben Taieb and Y. Le Borgne (2013), "Machine Learning Strategies for Time Series Forecasting", in *Lecture Notes in Business Information Processing, Business Intelligence*, Springer Berlin Heidelberg, Berlin, Heidelberg, https://doi.org/10.1007/978-3-642-36318-4_3. [17]

Brier, G. (1950), "Verification of Forecasts Expressed in Terms of Probability", *Monthly Weather Review*, Vol. 78/1, pp. 1-3. [10]

Bröcker, J. and L. Smith (2007), "Increasing the Reliability of Reliability Diagrams", *Weather and Forecasting*, Vol. 22/3, pp. 651-661, https://doi.org/10.1175/waf993.1. [9]

Carammia, M., S. Iacus and T. Wilkin (2022), "Forecasting asylum-related migration flows with machine learning and data at scale", *Scientific Reports*, Vol. 12/1, https://doi.org/10.1038/s41598-022-05241-8. [16]

Chatfield, C. (2000), *Time-Series Forecasting*, Chapman and Hall/CRC, https://doi.org/10.1201/9781420036206. [14]

Czado, C., T. Gneiting and L. Held (2009), "Predictive Model Assessment for Count Data", *Biometrics*, Vol. 65/4, pp. 1254-1261, https://doi.org/10.1111/j.1541-0420.2009.01191.x. [5]

De Domenico, M. (ed.) (2018), "Resilience of Complex Systems: State of the Art and Directions for Future Research", *Complexity*, Vol. 2018/1, https://doi.org/10.1155/2018/3421529. [25]

Durbin, J. and S. Koopman (2012), *Time Series Analysis by State Space Methods*, Oxford University Press, https://doi.org/10.1093/acprof:oso/9780199641178.001.0001. [22]

Gneiting, T., F. Balabdaoui and A. Raftery (2007), "Probabilistic Forecasts, Calibration and Sharpness", *Journal of the Royal Statistical Society Series B: Statistical Methodology*, Vol. 69/2, pp. 243-268, https://doi.org/10.1111/j.1467-9868.2007.00587.x. [4]

Gneiting, T. and A. Raftery (2007), "Strictly Proper Scoring Rules, Prediction, and Estimation", *Journal of the American Statistical Association*, Vol. 102/477, pp. 359-378, https://doi.org/10.1198/016214506000001437. [12]

Hyndman, R. and G. Athanasopoulos (2018), *Forecasting: Principles and Practice*, OTexts. [1]

Hyndman, R. and A. Koehler (2006), "Another look at measures of forecast accuracy", *International Journal of Forecasting*, Vol. 22/4, pp. 679-688, https://doi.org/10.1016/j.ijforecast.2006.03.001. [2]

Imbens, G. and D. Rubin (2015), *Causal Inference for Statistics, Social, and Biomedical Sciences*, Cambridge University Press, https://doi.org/10.1017/cbo9781139025751. [24]

Kilian, L. and H. Lütkepohl (2017), *Structural Vector Autoregressive Analysis*, Cambridge University Press, https://doi.org/10.1017/9781108164818. [19]

Pearl, J. (2009), *Causality: Models, Reasoning and Inference (*, Cambridge University Press, USA. [23]

Runge, J. et al. (2023), "Causal inference for time series", *Nature Reviews Earth & Environment*, Vol. 4/7, pp. 487-505, https://doi.org/10.1038/s43017-023-00431-y. [3]

Tashman, L. (2000), "Out-of-sample tests of forecasting accuracy: an analysis and review", *International Journal of Forecasting*, Vol. 16/4, pp. 437-450, https://doi.org/10.1016/s0169-2070(00)00065-0. [20]

Tetlock, P. and D. Gardner (2015), *Superforecasting: The Art and Science of Prediction*, New York: Crown. [11]

Welch, N. and A. Raftery (2022), "Probabilistic forecasts of international bilateral migration flows", *Proceedings of the National Academy of Sciences*, Vol. 119/35, https://doi.org/10.1073/pnas.2203822119. [13]

Zou, H. and T. Hastie (2005), "Regularization and Variable Selection Via the Elastic Net", *Journal of the Royal Statistical Society Series B: Statistical Methodology*, Vol. 67/2, pp. 301-320, https://doi.org/10.1111/j.1467-9868.2005.00503.x. [15]

# 8 How should forecasts be presented to policymakers and ensure effective interaction between policymaking and forecasting?

## How to present forecasts and their inherent uncertainty to policymakers

The previous sections have presented the use case for forecasts and explored technical issues in their design and implementation. However, forecasts are only useful if policymakers understand their constraints and how to make appropriate use of the forecasts. This places an onus on modellers and those responsible for communicating results internally. This section provides indications on how to present forecasts and explain uncertainty to policymakers.

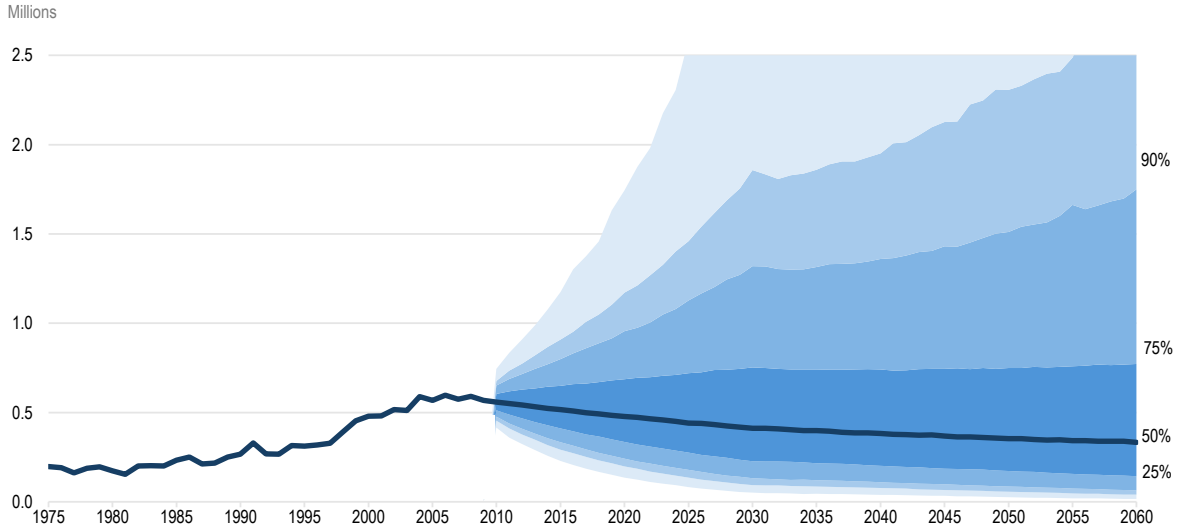
Communicating migration forecasts to policymakers effectively requires clear articulation of uncertainties and practical insights that can inform decision making. Given the complex and unpredictable nature of migration patterns, whatever the migration category under review, forecasts should be conveyed in a way that helps policymakers plan ahead while acknowledging their inherent limitations and the fact that uncertainty is a central element of the message.

One of the main challenges in presenting forecasts is indeed explaining uncertainty. Policymakers often request and expect definitive and precise numbers. Migration forecasts inherently involve uncertainty due to dynamic variables such as economic shifts, geopolitical events, and evolving policy landscapes. The most obvious solution is to present ranges of possible outcomes and scenarios rather than single-point estimates, which better capture the uncertainty associated with migration drivers. Several options are available, including the traditionally-used variant scenarios, either based on some combinations of the underlying migration drivers (e.g. four migration scenarios combining economic convergence/divergence across countries and unilateral/multilateral international co-operation in Acostamadiedo et al. (2020[1])), or full probabilistic outcomes, summarised through probability distributions (see Box 8.1) or a selection of predictive quantiles (Bijak and Wiśniowski, 2010[2]; Azose and Raftery, 2015[3]; Welch and Raftery, 2022[4]).

Presenting quantiles from predictive distributions to policymakers has one important advantage (see Bijak (2011[5]) and Raftery (2016[6])). Quantiles have a natural interpretation related to the frequency and magnitude of events: a median means that the occurrences below and above its value are equally likely, and each can be expected to occur on average half of the time. Quantile of order 0.95 means that higher values can be expected to occur only 1/20th of the time – so for annual forecasts, once every 20 years. Relative frequency of rare events (once in a decade, twice in a century, and so on, see Bijak (2024[7])), forms a base for analysing a range of migration scenarios.[1] These are relatively simple concepts to communicate. For this interpretation to be correct, however, the predictive intervals must be well calibrated (see Chapter 7).

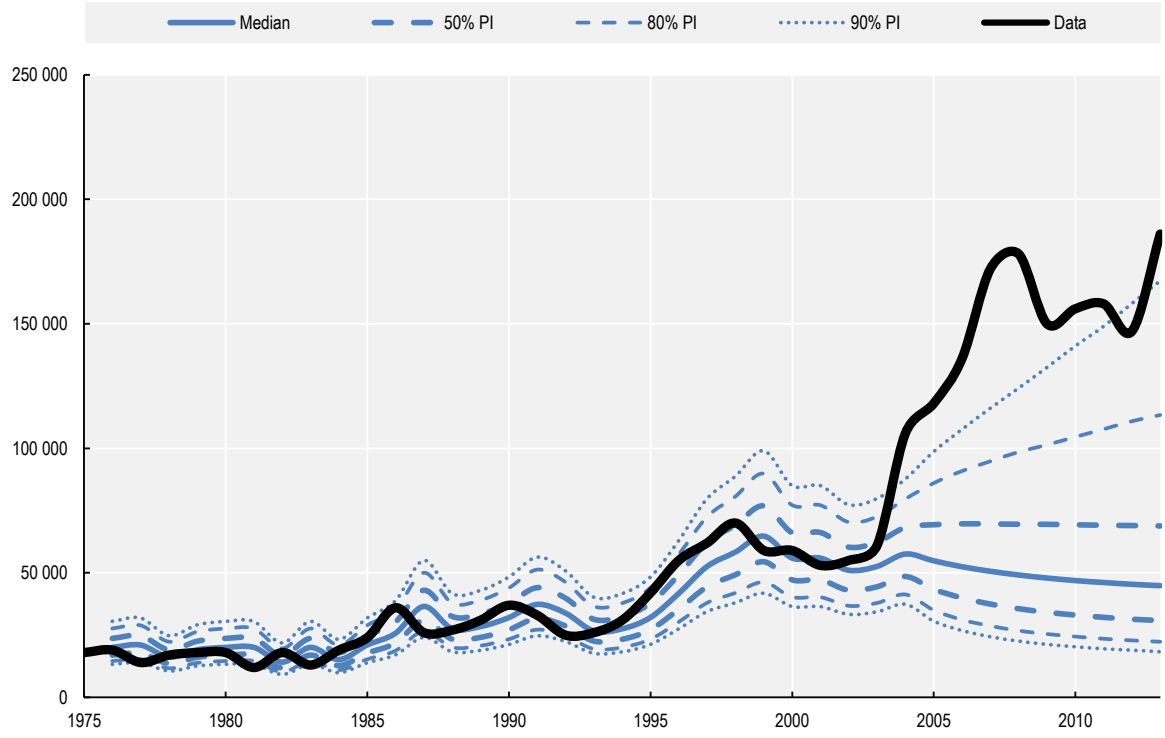
The graphic presentation can determine how policymakers understand uncertainty. One effective method for presenting uncertainty is through risk thresholds (e.g. the likelihood of exceeding a certain migration level) especially with visual tools such as fan charts (Figure 8.1), probability intervals (Figure 8.2), and density plots (Figure 8.3). An appropriate graphic representation can help policymakers grasp the range of potential migration outcomes.

### Figure 8.1. Forecasts of total immigration to the United Kingdom, averaged univariate models (in thousands) 1975-2060




Source: Abel et al. (2013[8]), "Forecasting environmental migration to the United Kingdom: an exploration using Bayesian models", https://doi.org/10.1007/s11111-013-0186-8.

## Figure 8.2. Immigration of EU Citizens to the United Kingdom with different probability intervals, 1975-2013




Source: Bijak et al. (2019[9]), "Assessing time series models for forecasting international migration: Lessons from the United Kingdom", https://doi.org/10.1002/for.2576.

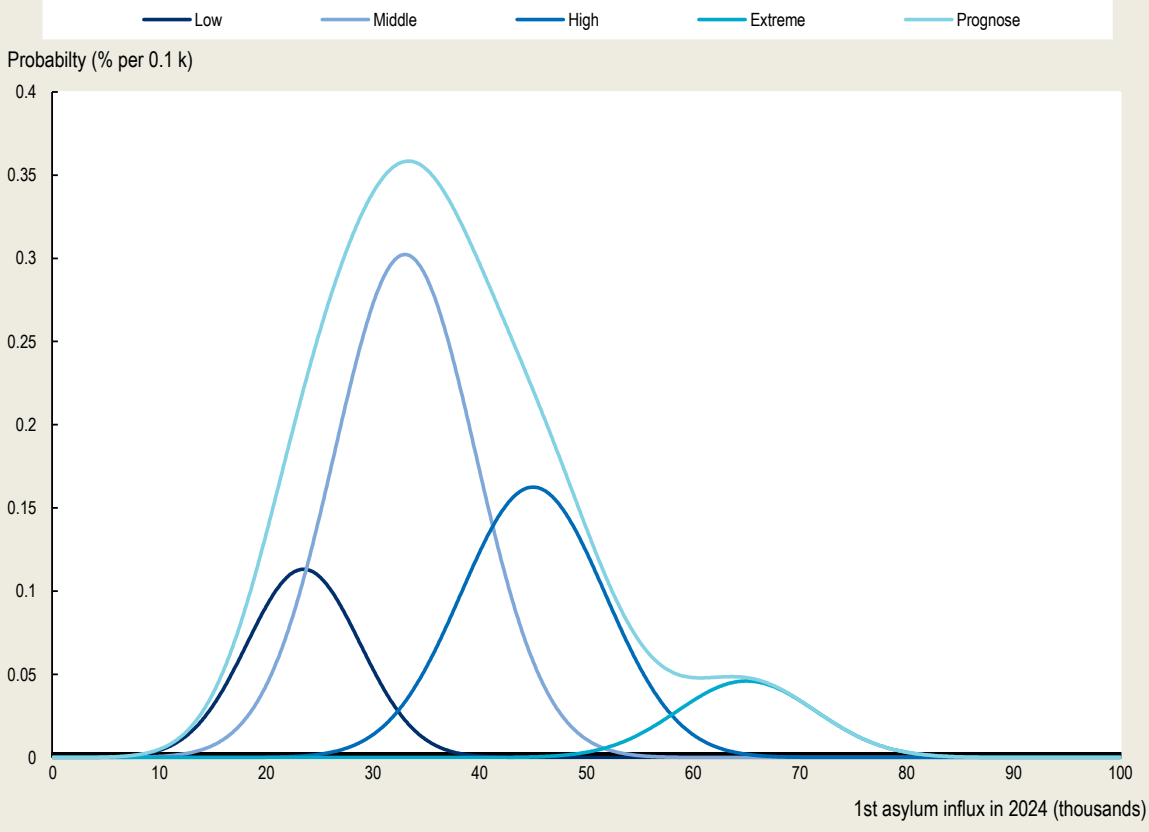
Another option involves preparing a risk management matrix, juxtaposing the uncertainty and user-specific impact of migration flows (Bijak et al., 2019[9]). Such a matrix typically involves a variant of the highly-intuitive traffic-lights colour scheme, with ratings ranging for green for low-uncertainty or low-impact flows and forecasts, through amber as an intermediary category, to red for processes that are both highly uncertain and can have high impact, as is the case with asylum-related migration.

Ideally, the users of forecasts – policymakers – should receive some minimal training in understanding probabilistic forecasts, in order for them to effectively integrate migration projections into strategic planning and not to draw mistaken conclusions from misreading probabilities. One way to equip decision makers to make data-driven choices would be (if relevant to the institutional set up) to offer workshops and training sessions on forecast interpretation to help align insights with policy objectives. This is forecasters' responsibility to teach policymakers how to read a model and that forecasting results are only possibilities of the future. A strong data culture helps policymakers better understand that forecasting models are the most efficient way to anticipate future migration flows, despite the inevitable and frustrating level of uncertainty. Some policymakers may remain reluctant to believe in forecasting outcomes; in this case, it's important to note that a *lack* of prediction may be more costly. Collaboration and regular interaction between forecasters and policymakers are also key to improving the practical application of forecasts, ensuring they are both interpreted correctly and effectively used (Bijak, 2024[7]). The forecasters' role should go beyond providing data and disseminating forecasting results to policymakers. They should regularly engage in discussions to ensure correct interpretation of the findings and to highlight emerging trends that require particular attention moving forward.

## Box 8.1. How the Netherlands presents its forecasts to policymakers

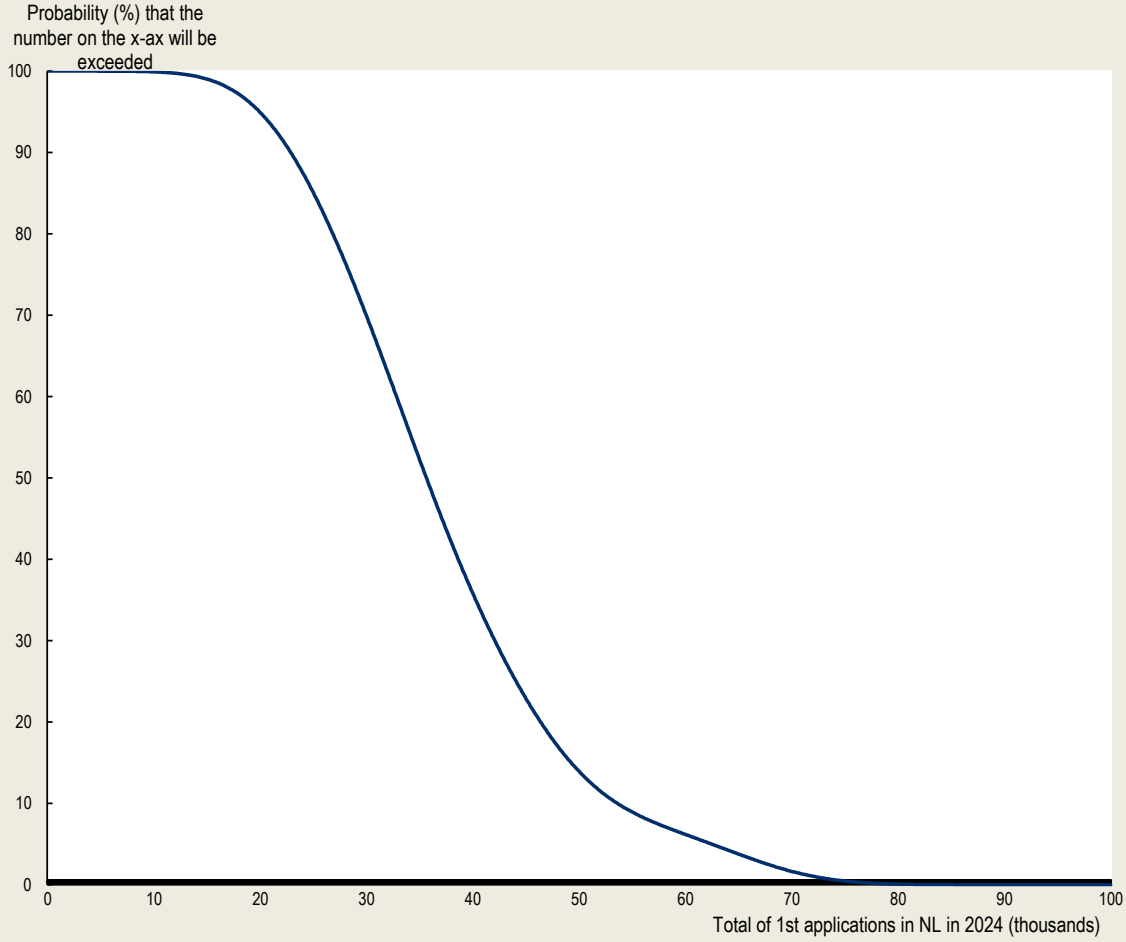
The Netherlands has been conducting asylum application forecasts in a new, systematic way since 2018. The Dutch mixed approach combines Bayesian probabilistic models with expert intuition inputs, eventually providing four or more different forecast scenarios. Each scenario is then presented in a Probability Density Function (PDF) figure, which shows each possible forecasting value with its probability of occurring, according to this specific scenario. A combined PDF, summing all scenarios, weighed according to their estimated total probability of occurring, is also presented (Figure 8.3) and integrated into a cumulative PDF (Figure 8.4). In the end, this approach provides a continuous function of predicted values, rather than one value within a confidence interval. Nonetheless, what interests most policymakers is a unique value. Such a unique value can be derived from the combination-PDF by also providing a so-called "loss-function", i.e. a function that quantifies the costs (both in money and in e.g. political unrest) of aiming "too low" versus aiming "too high". The combination of these two functions yields an optimum number to consider in the preparations. The forecast is used in budget planning: the Dutch budget for asylum reception has to match with a bandwidth between the 30th and the 70th percentiles of the cumulative probability density function.

### Figure 8.3. Probability Density Function for Prognosis – Forecasted first asylum applications 2024 in the Netherlands




Source: Dutch Ministry of Justice (2024).

### Figure 8.4. Cumulative probability for the Dutch number of first asylum applications in 2024




Source: Dutch Ministry of Justice (2024).

Against that background, forecasters in the Netherlands regularly present their results by trying to combat typical probabilistic misunderstandings. Using the PDF representation, they emphasise the wide bandwidth of probability that might occur: a 10% probability does not mean that this outcome will not happen, but rather indicates than once every ten years, something unexpected might happen. The COVID-19 crisis was a recent example of such a situation. Indeed, the number of actual 2020 asylum applications was under 15 000, which had only a 10% probability to occur according to forecasts, but still fell well within the PDF-indicated range predicted in 2019.

One important part of the Dutch approach is the central role of expert intuitions. One of the main arguments for the use of expert elicitation is that migration is such a complex phenomenon that predicting the consequences of various developments will never be fully objectifiable. As expert elicitation gathers many different insights and arguments, it provides a formal and informal exchange of ideas that are continuously challenged by one another. Even hard-to-conceive scenarios (extremely high or low migration flow levels) must be discussed by experts to take into account all possibilities.

The Dutch experience has also served to highlight the kind of misunderstandings policymakers may have, which can lead to wrong anticipations. It has been necessary for forecasters to e.g. insist that monthly forecasts must not be derived from annual forecasts by simply dividing the predicted value into

12 similar pieces. Explaining the seasonality of migration flows can help avoid this risk. Forecasters have also had to reign in the expectation that asylum influxes can be predicted with any precision or reliability beyond the horizon of two years in the future.

More importantly, forecasters have had to be pedagogical in repeating the aim of forecasting. An important task of the forecasters is to keep communicating clearly and transparently about the purpose and use of forecasts. Regular discussions are necessary to make sure policymakers are aware that forecasters are attempting foremost to provide a basis for sensible policy actions, rather than produce a headline figure. Predictions should be shared with all involved services, to make them aware of all hypothetical changes and possibilities. In that sense the forecasts can also be seen as a way to quantify ignorance. In the Dutch forecasts, the "tail" values indeed help policymakers to be better prepared and flexible and must therefore be taken into consideration to adapt the capacity of the asylum system in the extreme-case scenarios.

Even if a wide range of possibilities are forecast, it's still quite possible for reality to fall outside any forecast. In this case, forecasters have the difficult task of explaining how their model failed to policymakers.

Scenario-based forecasting presents migration projections as a range of possible futures rather than a fixed outcome, allowing for greater policy flexibility. It is important to incorporate scenario analysis into migration forecasting, helping policymakers anticipate various potential outcomes and prepare for uncertainty, rather than relying on a single definitive prediction.

However, when actual migration figures deviate from forecasts, transparent communication is crucial. In most cases, using a great range of possible estimates, such as the approach recommended in the previous section, helps to make sure that forecasts encompass the full range of possible outcomes. It's easier for forecasters to explain deviation when the latter still falls within the boundaries (even extreme) of the interval (see Box 8.1 for the interesting example of a forecast that falls within the COVID-19 lockdown). When forecasts fall outside the predicted values, forecasters have to explain to policymakers the underlying reasons for forecast errors, whether they stem from data limitations, unexpected policy changes, or external shocks such as conflicts or natural disasters.

Regularly updating forecasts and engaging in continuous dialogue about their limitations enable more adaptive decision making. Forecasters should clearly explain how external drivers influence migration projections, ensuring that policymakers interpret forecasts as dynamic and evolving rather than fixed predictions.

**Checklist:**

- Have I presented to policymakers the uncertainty of my forecasts in a clear manner?
  - Use charts and density plots showing quantiles from predictive intervals to clearly interpret the frequency of events, or alternatively traffic-light colour schemes in results tables.
- Do policymakers fully understand that forecasts are dynamic and not fixed predictions?
  - Regular interactions between forecasters and policymakers are necessary at each forecast updates. Training session explaining probabilistic forecasts may also help.
- What should I do if forecasts fell outside the range of estimates provided?
  - Precisely explain forecast errors, whether due to data limitations, unexpected policy changes or external shocks.

## How to use the forecast properly and effectively

Migration forecasts are generally commissioned in order to inform policy action, so their conclusions should be policy actionable. In some cases, action needs to be taken quickly: when early warning models identify large additional influxes, the policy response needs to be quick and efficient. An effective way to increase the impact of forecasting is to include forecasters – or at least forecasting results – in emergency budget allocation meetings, which define how to allocate financial resources (e.g. staff, infrastructure, see Chapter 3). A dissemination policy developed in advance can identify which national practitioners to prioritise to receive forecasting results. When warned in advance, these actors and their budget experts can use their knowledge to better anticipate the financial consequences of the migration forecast and better allocate the funds for future needs. Although on a less urgent basis, forecasts made with a longer time horizon (a few years) can support an efficient budget allocation as well with the right dissemination policy followed by a budget allocation meeting in due time.

Forecasts may identify future needs that go beyond budget allocation issues. Forecasts help anticipate the implementation of new policies or rules that enable the country to adapt to possible or expected future flows. For instance, if a mass influx is predicted that could exceed budgetary constraints, forecasting results can provide the impetus for policies that either ease legal frameworks and administrative procedures to better manage larger flows, or alternatively, prompt governments to adopt stricter measures to deter arrivals. Similarly, if governments aim to forecast future labour shortages, policymakers will also need to anticipate corresponding labour migration flows and consider facilitating migration policies accordingly. The examples of Japan and Korea (Box 8.2) are labour migration projections rather than forecasts, but they illustrate how such scenario-based exercises can still inform migration policymaking – for example, by helping to set visa quotas or adapt labour migration schemes. In practice, both forecasts and projections become policy-relevant when their results are clearly communicated and effectively linked to decision making processes.

To make sure forecasts are taken into account for policy action, the forecast results have to be disseminated to a clearly identified network of national practitioners, following a whole-of-society approach. Forecasting results can be shared with all stakeholders responsible for policy action on the ground (from national to local level), supporting their efforts to efficiently anticipate the necessary future action. However, not all forecasts should be made public. As the dissemination of forecasting results is a critical component of an effective policy response, it must adhere to established confidentiality and security protocols when shared with relevant stakeholders.

### Box 8.2. Labour migration projections in Asian OECD countries

In Japan, the JICA Ogata Institute projected foreign labour needs for 2030 and 2040. The study began with GDP growth targets set by the Ministry of Health, Labour and Welfare and calculated the achievable GDP based on demographic projections and capital stock scenarios, including automation. The gap between achievable and target GDP was attributed to foreign labour requirements, which were further disaggregated by industry and region. On the supply side, the model estimated potential migrants from 13 origin countries, using UN population projections, IMF growth forecasts (assuming emigration declines as incomes rise), Japan's historical migrant share, and retention rates. The study concluded that the stock of foreign workers would need to increase by 236% by 2040, yet only around 420 000 workers could realistically be supplied, leaving shortages of 0.77 million in 2030 and 0.97 million in 2040.

In Korea, labour demand – supply projections are used to guide the Ministry of Justice's employment visa quota pre-announcement system. In 2024, the quota for skilled worker visas (E-7-4) was raised to

30 000, up from 2 000 in 2023. A study commissioned to Seoul National University projected labour supply by sex, age, education, sector, and occupation, while labour demand was derived from models incorporating productivity and technological change. By comparing projected supply and demand for 2026-2031 under four scenarios with different assumptions about substitutability across sectors and skill levels, the government estimated varying levels of labour shortages. These results serve as the basis for setting annual visa quotas by sector, with higher caps considered where shortages persist and lower caps where substitution is feasible.

Taken together, the Japanese and Korean cases illustrate how **projections** can inform migration planning. They do not predict actual inflows, but provide structured scenarios linking labour demand, domestic supply, and long-term migration policy needs.

One use case for forecasts expressed as full probability distributions is to input them into formal statistical decision analysis to make informed decisions based on costs incurred to address the different scenarios. The information provided by the whole distributions of results can be combined with national practitioners' knowledge of the possible consequences of the decisions that are based on the forecasts. Such knowledge is best expressed in terms of loss functions, which describe the costs (losses) incurred when the reality departs from the forecasts – in other words, whenever the forecast either underpredicts or overpredicts the actual migration flows. The tacit assumption is that these costs are minimal if the forecast comes true, which however formally speaking has near-zero (or even zero) probability. The shape of the loss function and the relative magnitude of the costs of overprediction and underprediction, together with the probability distribution describing the forecast, jointly determine which predicted value should form the basis for decisions.

To illustrate this approach, for the numbers of asylum seekers, forecasts may serve as a useful tool for planning the resources needed to cover processing costs, as well as provide adequate reception services, registration support and other integration services if any. In this case, overpredicting asylum applications will likely lead to incurring additional costs through providing redundant reception capacity which ends up unused. On the other hand, having insufficient capacity will result in last-minute additional costs to scale up the reception and registration services, likely done not in the most efficient manner, as all emergency costs. In this case, and in most other practical applications, these costs are typically not symmetric: depending on the exact context, overprediction may end up being either more or less costly (per migrant) than underprediction by the same magnitude (see Box 8.3 for a practical example).

### Box 8.3. A Practical Rule from Decision Theory

The methodology and technical aspects of the statistical decision theory is discussed more fully elsewhere (e.g. Robert (2007[10])), but there are some simple and approximate rules, which can be easily applied in practice. One such rule refers to the situation when the costs (losses) scale up approximately linearly with the magnitude of overprediction and underprediction, and that the costs (losses) of underprediction are k times higher than for overprediction of the same size. In such cases, it can be mathematically shown that the optimal predictions, minimising the expected loss under a given predictive distribution, are quantiles of order k/(k+1) from the predictive distributions. For example, if k = 4, the optimal prediction is the fourth quintile (quantile of order 4/(4+1) = 0.8), whereas for k = 0.25, indicating that in fact overprediction is four times costlier, the first quintile (0.25/(1+0.25) = 0.2) is optimal. A special case occurs when the costs of overprediction and underprediction are linear and symmetric (k = 1): the optimal decision is then the median (quantile of order 1/(1+1) = 0.5) of the predictive distribution. The value of k needs to be also elicited, this time from the users of the forecasts (the national practitioners), who will be best positioned to judge the relative costs of over- and underprediction.

One important point to make is that if forecasts serve as input for policy decisions, then these decisions can, by their very nature, influence the migration flows and invalidate the forecasts, at least to some extent (see the *Lucas critique*, Lucas (1976[11])). Still, the policy effects on migration are uncertain and difficult to quantify – and so are the effects of migration on policy (Czaika et al., 2024[12]). One possible way of dealing with such challenges is to create several *conditional forecasts*, which can be in fact seen as scenarios, created under the assumptions of different decisions, or different political or policy developments. One example of implementing such an approach is a set of conditional forecasts of total migration to and from Scotland depending on the results of the 2014 referendum on Scottish (Wiśniowski, Bijak and Shang, 2014[13]). In this example, the calculated expert-based forecasts were conditional on the outcome of the referendum and subsequently averaged over both possible outcomes (Yes and No), based on a prediction of the referendum outcome at the time of preparing the forecast.

These caveats notwithstanding, successfully presenting migration forecasts to policymakers requires a careful balance between clarity and complexity. Ensuring transparent communication about forecast uncertainty, offering explanations for deviations, and integrating forecasts into policy discussions are all crucial for enhancing the role of migration forecasting in decision making. By utilising probabilistic models, scenario-building techniques, and continuous updates, forecasts can become a valuable tool in anticipating and effectively managing migration flows.

**Checklist:**

- How to make sure that forecasts are translated in budget planning and policy action?
  - Disseminate forecasting results to the right network of practitioners and integrate forecasters in budget allocation meetings and policy discussions is key to take the most informed decision.
- Has the magnitude of costs' underprediction versus overprediction been clearly defined?
  - To minimise the loss, the optimal forecast value has to be taken within the predictive distribution bandwidth, after an expert discussion defining the cost of underprediction versus overprediction, For asylum reception for instance, the cost of urgent reception capacity versus unused reception capacity.
- Have I built conditional forecasts depending on different scenarios?
  - Where some policy decisions or development are ongoing, create scenarios based on their different outcomes may help grasp the full extent of migration predictions.

## References


Abel, G. et al. (2013), "Forecasting environmental migration to the United Kingdom: an exploration using Bayesian models", *Population and Environment*, Vol. 35/2, pp. 183-203, https://doi.org/10.1007/s11111-013-0186-8. [8]

Acostamadiedo, E. et al. (2020), *Assessing Immigration Scenarios for the European Union in 2030 – Relevant, Realistic and Reliable?*, International Organization for Migration and the Netherlands Interdisciplinary Demographic Institute. [1]

Azose, J. and A. Raftery (2015), "Bayesian Probabilistic Projection of International Migration", *Demography*, Vol. 52/5, pp. 1627-1650, https://doi.org/10.1007/s13524-015-0415-0. [3]

Bijak, J. (ed.) (2024), *From Uncertainty to Policy: A Guide to Migration Scenarios*, Edward Elgar Publishing, https://doi.org/10.4337/9781035319800. [7]

Bijak, J. (2011), *Forecasting International Migration in Europe: A Bayesian View*, Springer Netherlands, Dordrecht, https://doi.org/10.1007/978-90-481-8897-0. [5]

Bijak, J. et al. (2019), "Assessing time series models for forecasting international migration: Lessons from the United Kingdom", *Journal of Forecasting*, Vol. 38/5, pp. 470-487, https://doi.org/10.1002/for.2576. [9]

Bijak, J. and A. Wiśniowski (2010), "Bayesian Forecasting of Immigration to Selected European Countries by using Expert Knowledge", *Journal of the Royal Statistical Society Series A: Statistics in Society*, Vol. 173/4, pp. 775-796, https://doi.org/10.1111/j.1467-985x.2009.00635.x. [2]

Czaika, M. et al. (2024), "European migration governance in the context of uncertainty", *International Migration*, Vol. 62/6, pp. 93-110, https://doi.org/10.1111/imig.13308. [12]

Lucas, R. (1976), "Econometric policy evaluation: A critique", *Carnegie-Rochester Conference Series on Public Policy*, Vol. 1, pp. 19-46, https://doi.org/10.1016/s0167-2231(76)80003-6. [11]

Raftery, A. (2016), "Use and communication of probabilistic forecasts", *Statistical Analysis and Data Mining: The ASA Data Science Journal*, Vol. 9/6, pp. 397-410, https://doi.org/10.1002/sam.11302. [6]

Robert, C. (2007), *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation*, Springer New York, New York, NY, https://doi.org/10.1007/0-387-71599-1. [10]

Welch, N. and A. Raftery (2022), "Probabilistic forecasts of international bilateral migration flows", *Proceedings of the National Academy of Sciences*, Vol. 119/35, https://doi.org/10.1073/pnas.2203822119. [4]

Wiśniowski, A., J. Bijak and H. Shang (2014), "Forecasting Scottish Migration in the Context of the 2014 Constitutional Change Debate", *Population, Space and Place*, Vol. 20/5, pp. 455-464, https://doi.org/10.1002/psp.1856. [13]

## Notes

[1] For an example, see the QuantMig Migration Scenarios Explorer: https://www.quantmig.eu/data_and_estimates/scenarios_explorer/ (as of 15 November 2024).

# 9 What future developments can be expected in migration forecasting?

## How to measure the potential effect of policies in countries of destination?

There has been a growing interest in approaches to causal interpretation of potential policy effect employing synthetic control methods, machine learning approaches and more general stochastic processes settings.

Since its introduction (Abadie, Diamond and Hainmueller, 2010[1]), the Synthetic Control Method (SCM) has become a powerful statistical technique used to evaluate the effects of interventions or treatments in comparative case studies, particularly when a single unit (e.g. a country, region, or city) undergoes an intervention, and suitable control units are available for comparison. SCM constructs a weighted combination of control units to create a "synthetic" version of the treated unit, which approximates what would have happened in the absence of the intervention. This method has been widely applied across various fields, including economics, political science, and public health. Klößner and Pfeifer (2017[2]), for instance, explored the use of SCM as a pure forecasting tool. By applying SCM to the United States GDP growth, the authors demonstrate that it performs competitively compared to alternative forecasting methods. This idea has now been embraced by several scholars to study the impact of external events (e.g. economic crisis, war, climate events) on migration flows.

SCM has increasingly attracted the attention of migration experts as well particularly for policy impact analysis and natural experiments, like macro shocks (e.g. climate change, conflict). Rodríguez Sánchez et al. (2023[3]), for instance, explored the ability of Bayesian Structural Time Series (BSTS) models (Scott and Varian (2014[4])) in the context of irregular border crossings forecast. BSTS is a robust statistical approach for time series analysis, particularly suited for causal inference in observational data with complex temporal patterns. In that particular application, the model constructs a synthetic counterfactual by forecasting what the number of arrivals would have been in the absence of policy, based on historical trends and covariates such as climate, political events, and economic factors. The BSTS framework uses a Bayesian approach to incorporate prior knowledge and quantify uncertainty around the estimates, making it highly reliable for causal inference in dynamic systems.

Many variant and innovative approaches of causal inference for time series have been developed in recent years (see Box 9.1). In all these variants, two considerations are very important. The first consideration is the accurate selection of control units. The choice of the donor pool is crucial to ensure that the synthetic control closely matches the treated unit's pre-intervention characteristics. The accurate selection is important because these methods always produce "technical" counterfactuals, as an output of a weighting method. If within the pool of control units there are members who are structurally different from the treated unit, the final estimated impact will include bias due to confounding factors. The second consideration concerns the underlying assumptions. Researchers must be mindful of the assumptions underlying SCM, such as the stability of relationships over time (meaning that all the predictors used and the outcome variable have a static relationship, an assumption that is rarely true in migration flows as shown by, e.g.

(Carammia, Iacus and Wilkin, 2022[5])), and the absence of spillover effects between units which is quite likely in migration flows because of the complexity of the phenomena, the networking effects, etc.

### Box 9.1. Other approaches to causal inference

Goldin et al. (2024[6]) introduced an innovative approach to causal inference in panel data settings by leveraging advanced time series forecasting techniques. The focus is on the ability of different deep-neural network algorithms to produce counterfactual evolutions for treated units by incorporating both historical data for the treated units, and contemporaneous data from control units. Using a holistic approach, the proposed model captures both temporal dynamics and cross-sectional relationships.

In a recent contribution, Cinquini et al. (2025[7]) address the challenge of estimating causal effects of interventions in dynamical systems over time. The authors introduce a formal framework to define and analyse causal interventions within the framework of Discrete-Time Stochastic Processes (DSPs), which are models representing systems that evolve in discrete time steps with inherent randomness. In this context it is possible to show, under certain conditions, the equivalence between the equilibrium state of a DSP and a traditional structural causal model (SCM). This equivalence allows for the application of causal inference techniques traditionally used in SCMs to time-series data. In particular, it is possible to explicitly map VAR models to linear SCMs. This mapping accommodates potential cyclic relationships and unmeasured confounders, enhancing the interpretability of VAR models in causal terms. Authors' experiments demonstrate that the framework achieves strong performance in observational forecasting and accurately estimates the causal effects of interventions on dynamical systems.

For a general review of causal inference in time-series forecasting see also Runge et al. (2023[8]).

## Is conformal prediction the final frontier in time series forecasting?

Conformal prediction is a statistical framework that provides valid prediction intervals for machine learning or pure statistical models, ensuring that the true value of a prediction lies within a specified interval with a certain confidence level. This method has been applied across various domains typical of the i.i.d. (independent and identically distributed) setting and has been recently extended to time series analysis. The basic assumption of conformal prediction is exchangeability of the data. Exchangeability means that the order of the data points does not matter, and any permutation of the data points has the same joint probability distribution. This assumption is weaker than the often-used i.i.d. assumption but still remains crucial for the validity of conformal prediction.

The basic idea behind conformal time-series forecasting (Stankeviciute, Alaa and van der Schaar, 2021[9]) is to adapt the conformal prediction framework to time-series data, allowing for the creation of prediction intervals that provide a predefined, valid coverage probability regardless of the underlying model or data distribution. The approach extends the conformal prediction methodology, traditionally used for i.i.d. data, to handle dependent time-series data where observations are typically autocorrelated. The adaptation consists of addressing the temporal dependencies by calibrating residuals from a forecasting model in a way that accounts for the autocorrelations between observations. The advantage of conformal time-series forecasting is that it is model agnostic, i.e. it can in principle be applied to any base forecasting model, including neural networks, ARIMA, or other machine learning-based models, as long as the model provides point predictions. The algorithm proposed by Stankeviciute, Alaa and van der Schaar (2021[9]) is designed to handle multi-step forecasts, constructing prediction intervals for multiple future time steps simultaneously. The intervals reflect uncertainty at each forecast horizon. Despite its advantages, conformal prediction is still not easy to apply and remains under explored in the context of migration forecasting.

## How can forecasting results in one migration category affect forecasting in other categories?

Migration categories – such as international students, labour migrants, asylum seekers, and family reunification migrants – do not function in isolation. Changes in one category can trigger cascading effects on others. For policymakers and forecasters, understanding these interdependencies is crucial for producing accurate and actionable migration forecasts.

One illustrative example of cross-category influence is the transition of international students into the labour market. For example, former international students who have graduated in the host country are an important feeder for labour migration in many countries and the share of educational permits changed to a work permit accounted for a large share of total admissions for work (for OECD countries, OECD (2022[10]); for the United States, Beine, Peri and Raux (2023[11]), for the United Kingdom, Brindle and Sumption (2024[12]), for the Netherlands, Didier and Merve Nezihe (2014[13])). Therefore, in order to accurately forecast labour migration, it is essential to first estimate the number of incoming international students, and then to account for student retention rates following graduation in the workforce or even transition to family permits.

Another example of interconnection between migration categories is family migration. Family migration can be considered a typical example of chain migration. Chain migration is due to earlier arrivals, including settled migrants who bring family members (Böcker, 1994[14]). In all OECD countries, the family reunification process is legally bound by the presence of a sponsor (principal applicant) already in the country for a certain duration of time, usually labour migrants, beneficiaries of international protection, as well as host-country nationals (OECD, 2017[15]). Future family migration is therefore a function of past migration flows of the aforementioned categories. Forecasting family migration of foreign nationals requires four key pieces of information: i) the current and expected future flows of migrants, categorised by types of migration (e.g. labour migration, highly skilled labour, international protection); ii) the proportion of migrants in each category who return to their home country early, before any family reunification process begins (see for instance Viné (2021[16]); iii) the number or share of family members joining the principal applicants who remain in the host country (adjusted by the proportion returning early from point ii); iv) the average time lag between the arrival of the principal applicant in the host country and the arrival of their family members.

Using longitudinal data from the population register in the Netherlands, Wijkhuijs and Jennissen (2010[17]) estimated that the average number of family members who joined third-country national labour migrants was 0.75, while it was 0.66 for Polish labour migrants and 0.45 for EU15 nationals. Exploiting a regularisation act in the United States (IRCA 1986), Cascio et al. (2024[18]) estimated that each regularised Mexican migrant was responsible for the subsequent immigration and legal admission of one relative, in total, through 2019.

Given these interdependencies, migration forecasts must incorporate feedback loops between categories. Rigid, siloed models may overlook how trends in one group influence another. To improve forecasting accuracy and policy preparedness, dynamic, multi-category forecasting models should be employed to integrate multiple migration pathways and capture transition probabilities between categories (e.g. student-to-labour force retention rates).

Policymakers should be aware of these interconnections to avoid unintended consequences when designing migration policies. For instance, changes in student visa policies may have long-term labour market effects that should be factored into workforce planning. By acknowledging and incorporating cross-category influences, migration forecasting can provide a more comprehensive foundation for policymaking, reducing uncertainty and enhancing preparedness for future migration trends.

## How can forecasts in one country affect forecasts in other countries (i.e. interaction across countries)?

Migration is a complex phenomenon that involves not only one mobility from a country A to a country B. Migration is not necessarily a finite process; it can also be circular (back and forth from a country A to a country B) or involve multiple transit countries. Therefore, migration flows in one country can affect flows in other countries, sometimes to a great extent. This is particularly true in the European context, where immigrants, after crossing the external borders of the EU, enjoy a degree of unconstrained mobility within the Schengen area. They can therefore move more easily from one Member State to another. This may lead to asylum claims being lodged not in countries other than the first country of arrival: in 2023 and 2024, EU countries received 185 258 and 167 561 "Dublin" requests, respectively, i.e. requests to reaccept third-country nationals who have applied for asylum or crossed into a country different from the responsible country (usually their country of first entry within the EU).

Against that background, OECD countries, whenever possible, increasingly use data from other countries to anticipate their own migration flows. The United States use Mexico's encounters of irregular migrants to forecast their own irregular migration (Table 5.1). The DynENet model proposed by Carammia et al. (2022[5]) (see section on machine learning models) analyses interactions across countries for forecasting purposes as well. For instance, it showed that events in Türkiye may have a different effect on asylum applications by Syrians from one EU country to another, depending on migration drivers between Syria and each EU country. As a consequence, forecasts run in transit countries, and in particular in neighbouring countries to an OECD country, may significantly affect the forecast for this OECD country. That interaction is stronger for forced migration forecasting but should not be overlooked for other migration categories. Indeed, an economic crisis in a country A may encourage labour migration flows or international students to go further in a country B, especially among the most highly skilled. Similarly, a policy decision in one country can divert migration flows (forced or not) to another country B.

In order to produce the most reliable results, forecasts in one country must therefore take into account forecasts in other countries, where routes exist between them. While this seems obvious from a statistical point of view, its practical implementation is far from simple. As forecasts are usually sensitive, their dissemination may be limited outside national boundaries. Even national migration data are not necessarily disseminated abroad, especially the most sensitive ones, which are often used in forecasting. Irregular border crossing data at the EU external border are available through EU agencies such as Frontex (https://www.frontex.europa.eu/what-we-do/monitoring-and-risk-analysis/migratory-map/) or Eurostat (https://ec.europa.eu/eurostat/databrowser/view/migr_eipre/default/table?lang=fr&category=migr.migr_man.migr_eil), but border crossings within the EU, when available, are seldom disseminated even with the countries sharing a common border. Border crossing data within the EU are implicitly limited, and do not fully capture actual crossings, as the data lack coverage in the context of open borders, and depend on varying levels of border control effectiveness and the size of the border force mobilised over time.

All in all, OECD countries could develop more accurate forecasting models if they had access to other geographically connected countries' forecasts and data. That requires more co-ordination between them and regular contacts between forecasting data providers. This also implicitly requires a certain level of data comparability across OECD countries. In order to do so, a greater alignment of definitions and frequencies across OECD countries is needed. Data harmonisation also requires regular contacts between data providers and an international co-ordination body to ensure a unique definition is clearly defined and applicable. Although full harmonisation may be challenging, this is indeed an important condition under which cross-country forecasting can realistically improve. In the end, all these exchanges would not only improve the quality of migration forecasts but also foster stronger collaboration between forecasting units across OECD countries.

## References


Abadie, A., A. Diamond and J. Hainmueller (2010), "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program", *Journal of the American Statistical Association*, Vol. 105/490, pp. 493-505, https://doi.org/10.1198/jasa.2009.ap08746. [1]

Beine, M., G. Peri and M. Raux (2023), "International college students' impact on the US skilled labor supply", *Journal of Public Economics*, Vol. 223, p. 104917, https://doi.org/10.1016/j.jpubeco.2023.104917. [11]

Böcker, A. (1994), "Chain migration over legally closed borders: Settled immigrants as bridgeheads and gatekeepers", *The Netherlands' Journal of Social Sciences*, Vol. 30/2, pp. 87-106. [14]

Brindle, B. and M. Sumption (2024), *International students entering the UK labour market*, https://migrationobservatory.ox.ac.uk/resources/commentaries/international-students-entering-the-uk-labour-market/. [12]

Carammia, M., S. Iacus and T. Wilkin (2022), "Forecasting asylum-related migration flows with machine learning and data at scale", *Scientific Reports*, Vol. 12/1, https://doi.org/10.1038/s41598-022-05241-8. [5]

Cascio, E., E. Lewis and C. Zhang (2024), *How Good are Proxies for Legal Status? Evidence from the Legalization of Two Million Mexicans*, National Bureau of Economic Research, Cambridge, MA, https://doi.org/10.3386/w32632. [18]

Cinquini, M. et al. (2025), "A Practical Approach to Causal Inference over Time", *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 39/14, pp. 14832-14839, https://doi.org/10.1609/aaai.v39i14.33626. [7]

D., F. and Ö. M.N. (2014), *International mobility of students - Its impact on labour market forecasts and its contribution to the Dutch economy*, University of Maastricht, https://doi.org/10.26481/umarot.2014006. [13]

Goldin, J., J. Nyarko and J. Young (2024), "Forecasting Algorithms for Causal Inference with Panel Data", *Proceedings of the 25th ACM Conference on Economics and Computation*, pp. 1205-1205, https://doi.org/10.1145/3670865.3673641. [6]

Klößner, S. and G. Pfeifer (2017), "Outside the box: using synthetic control methods as a forecasting technique", *Applied Economics Letters*, Vol. 25/9, pp. 615-618, https://doi.org/10.1080/13504851.2017.1352071. [2]

OECD (2022), *International Migration Outlook 2022*, OECD Publishing, Paris, https://doi.org/10.1787/30fe16d2-en. [10]

OECD (2017), "A portrait of family migration in OECD countries", in *International Migration Outlook 2017*, OECD Publishing, Paris, https://doi.org/10.1787/migr_outlook-2017-6-en. [15]

Ranzato, M. et al. (eds.) (2021), *Conformal time-series forecasting*. [9]

Rodríguez Sánchez, A. et al. (2023), "Search-and-rescue in the Central Mediterranean Route does not induce migration: Predictive modeling to answer causal queries in migration research", *Scientific Reports*, Vol. 13/1, https://doi.org/10.1038/s41598-023-38119-4. [3]

Runge, J. et al. (2023), "Causal inference for time series", *Nature Reviews Earth & Environment*, Vol. 4/7, pp. 487-505, https://doi.org/10.1038/s43017-023-00431-y. [8]

Scott, S. and H. Varian (2014), "Predicting the present with Bayesian structural time series", *International Journal of Mathematical Modelling and Numerical Optimisation*, Vol. 5/1/2, p. 4, https://doi.org/10.1504/ijmmno.2014.059942. [4]

Viné, R. (2021), "Economic Drivers of Family Reunification in the Context of International Migration", *SSRN Electronic Journal*, https://doi.org/10.2139/ssrn.3761361. [16]

Wijkhuijs, L. and R. Jennissen (2010), *Arbeidsmigratie naar Nederland*. [17]

# Annex A. List of Migration Anticipation and Preparedness (MAP) task force members

| | | |
|---|---|---|
| Canada | Ministère de l'Immigration, de la Francisation et de l'Intégration | Olivier RANCOURT |
| Canada | Immigration, Refugees and Citizenship Canada | Stéphane ARABACKYJ |
| Canada | Immigration, Refugees and Citizenship Canada | Francis CRADDOCK |
| Canada | Immigration, Refugees and Citizenship Canada | Olga ANGLINOVA |
| Denmark | Office for International Co-ordination on Migration | Nicolai TROCK |
| France | Direction générale des Étrangers Ministère de l'Intérieur | Guillaume MORDANT |
| France | Direction générale des Étrangers en France | Yves BREEM |
| Germany | Federal Office for Migration and Refugees (BAMF) | Tim BREITKREUZ |
| Germany | Federal Office for Migration and Refugees (BAMF) | Mark BERGES |
| Greece | Ministry of Migration & Asylum | Angeliki GRAMMATIKOPOULOU |
| Greece | Ministry of Citizen Protection | Petros KONSTANTOPOULOS |
| Ireland | Ministry of Justice | Gurchand SINGH |
| Ireland | Ministry of Finance | Neil GANNON |
| Ireland | Ministry of Finance | Kevin DALY |
| Japan | National Institute of Population and Social Security Research | Yu KOREKAWA |
| Korea | Immigration Policy Reform Division Ministry of Justice | Myunghoon KIM |
| Korea | Migration Research & Training Center | Changwon LEE |
| Korea | Immigration Policy Division Ministry of Justice | Hye Sun KIM |
| Korea | Immigration Big Data Division Ministry of Justice | Kyungmin JEON |
| Luxembourg | European Migration Network (EMN) Luxembourg | Adolfo SOMMARRIBAS |
| Netherlands | Ministry of Justice | Frans ALKEMADE |
| Netherlands | Ministry of Justice | Arline VAN MAANEN |
| Poland | Office for Foreigners | Magdalena KOZLOWSKA |
| Portugal | Agency for Integration, Migration and Asylum | Eduardo da Fonseca QUA |
| Portugal | Agency for Integration, Migration and Asylum | Angella GRAÇA |
| Sweden | Swedish Migration Agency | Pär WERNER |
| Sweden | Swedish Migration Agency | Sally Persson UPADHYAYA |
| Switzerland | Secrétariat d'État aux migrations (SEM) | Clovis VOISARD |
| Switzerland | Secrétariat d'État aux migrations (SEM) | Christoph CURCHOD |
| United States | Migration Analysis Center | Nadwa MOSSAAD |
| United States | Citizenship and Immigration Services | Delancey GUSTIN |
| Turkiye | Department of International Labour Force Policies and Researches | Aydin ERDOGAN |
| Turkiye | Department of International Labour Force Policies and Researches | Fikret DUMAN |
| | Joint Research Center | Stefano IACUS |
| | University of Catania | Marcello CARAMMIA |
| | University of Southampton | Jakub BIJAK |
| OECD | International Migration Division | Jean-Christophe DUMONT |
| OECD | International Migration Division | Jonathan CHALOFF |
| OECD | International Migration Division | Taehoon LEE |

# Glossary

**Adaptive models**: Models that adjust their parameters or structure dynamically over time in response to new data or changing conditions. For example, the variables selected as well as their relative importance in Dynamic Elastic Net Model (DynENet) may vary from week to week.

**Agent-based modelling (ABM)**: A simulation approach that models the actions and interactions of individual agents (e.g. individuals or households) to explore how individual decisions shape overall dynamics.

**Auto-Regressive Integrated Moving Average (ARIMA)**: ARIMA model is the most general class of models for forecasting a time series which can be made to be "stationary" by differencing in conjunction with nonlinear transformations such as logging or deflating. A non-seasonal ARIMA model is classified as an "ARIMA(p,d,q)" model, p is the number of autoregressive terms, d is the number of non-seasonal differences needed for stationarity, and q is the number of lagged forecast errors in the prediction equation.

**Auto-Regressive Integrated Moving Average with eXogenous inputs (ARIMAX)**: The ARIMAX model extends the ARIMA framework by integrating exogenous variables, which are external predictors of factors that can improve the forecast accuracy.

**Bayesian Network Models**: Probabilistic graphical models representing variables and their conditional dependencies via a Directed Acyclic Graph (DAG), useful for modelling complex causal structures.

**Bayesian Structural Time Series (BSTS)**: A Bayesian time series model that decomposes data into trend, seasonal, and regression components. Useful for forecasting and causal inference under uncertainty.

**Calibration curves/scores**: Tools to assess how well predicted probabilities match observed outcomes. A well-calibrated model outputs probabilities that correspond closely to actual event frequencies.

**Causal forecasting models**: Models that incorporate the mechanisms and motivations behind migration decisions, often using individual-level data to distinguish between types of migrants (e.g. economic vs. forced).

**Classification-based migration forecasting**: Forecasting approaches that predict the *category* (e.g. low/medium/high) of future migration volumes rather than precise numerical counts.

**Directed Acyclic Graph (DAG)**: A finite graph with directed edges and no cycles, often used to represent causal relationships in Bayesian networks or structural models.

**Dynamic causal models**: Time-sensitive causal models that account for changes in variable relationships over time, often used to simulate the impact of shocks or policies on migration behaviour.

**Exponential smoothing models**: Special cases of ARIMA models (ARIMA (0,1,1)) which use an average of the last few observations rather than the most recent observation to filter out the noise and more accurately estimate the local mean. Often used for short-term forecasting of series without strong trends or seasonality.

**Importance-Frequency (IF) space**: A two-dimensional space used to visualise and assess model features based on Importance (How strongly a feature influences the model's predictions and Frequency (How often the feature appears or is used across different models, forecasts, or simulations.).

**Least Absolute Shrinkage and Selection Operator (LASSO)**: A regression technique that includes a penalty for large coefficients, encouraging simpler, more interpretable models and reducing overfitting.

**Long Short-Term Memory (LSTM)**: A type of Recurrent Neural Network (RNN) that captures long-term dependencies in sequential data, commonly used for time series forecasting.

**Machine Learning**: A class of algorithms that identify patterns in data and improve their performance with experience, widely used in predictive migration models.

**Macro-Simulation**: A modelling technique that uses aggregate data and equations to forecast large-scale migration flows.

**Mean Absolute Error (MAE)**: A metric that measures the average magnitude of forecast errors, regardless of direction.

**Mean Absolute Percentage Error (MAPE)**: A metric that expresses forecast errors as a percentage of the actual values, useful for comparing errors across datasets.

**Mean Absolute Scaled Error (MASE)**: A scale-independent metric for forecast accuracy, useful for comparing models across different scales and time series.

**Mean Percentage Error (MPE)**: A metric that captures the average percentage bias of forecasts; it indicates whether a model tends to over- or under-predict.

**Micro-simulation**: A modelling approach simulating individual-level behaviours based on detailed characteristics and decision making rules.

**Modern deep-learning techniques**: Advanced neural network models (e.g. LSTM, RNN, transformers) capable of learning complex, non-linear relationships in high-dimensional data.

**Overfitting**: When a model fits training data too closely, capturing noise rather than generalizable patterns, resulting in poor performance on new data.

**Partial Dependence Plots (PDPs)**: Visualization tools showing the marginal effect of one or two input features on the model's predictions.

**Probability Density Function (PDFs):** Function describing the likelihood of a continuous random variable taking on certain values. The total area under the curve of a PDF is equal to one, and the probability that the variable falls within a specific interval is given by the integral of the PDF over that interval.

**Random Walk model**: A time series model where the current value is equal to the previous value plus a random error term. Random walks are *non-stationary* processes, and the model often serves as naïve benchmarks in forecasting, representing the idea that the best prediction for tomorrow is today's value.
**Recursive Neural Networks**: Neural networks that apply the same weights recursively across structured data (e.g. syntactic trees), not to be confused with Recurrent Neural Networks (RNNs).

**Recurrent Neural Network (RNN)**: A type of neural network particularly well-suited for sequential data, such as time series, due to its memory of previous inputs.

**Ridge penalty**: A regularisation technique that adds a penalty term to regression coefficients to reduce overfitting and improve model generalisation.

**Root Mean Square Error (RMSE)**: A commonly used forecast error metric that penalises large errors more heavily due to squaring the error terms.

**Saliency Maps**: Tools used in neural networks to visualise which input features most influence the model's predictions.

**SARIMA (Seasonal ARIMA)**: An extension of ARIMA that incorporates seasonal components, useful for forecasting time series with repeating seasonal patterns.

**SHAP (SHapley Additive exPlanations)**: A model-agnostic interpretability method that assigns each input feature an importance value for individual predictions based on co-operative game theory.

**Stationary time series:** Time series whose statistical properties such as mean, variance, and autocorrelation are constant over time, i.e. which does not exhibit trends or seasonal effects, and its properties do not change when shifted in time. Stationarity is an important assumption for many time series models.

**Time series-specific metrics**: Metrics designed for evaluating time series forecasts, including MASE, Theil's U-statistic, and others.

**Traditional metrics**: Commonly used forecast error metrics such as MAE, MAPE, RMSE, and MPE.

**Vector Auto-Regressive (VAR) model**: A multivariate time series model where each variable is modelled as a function of past values of itself and other variables in the system.

# Migration Anticipation and Preparedness

## Making Migration Management Work

Governments face increasing pressure to respond swiftly and effectively to rapidly evolving international migration flows – whether forced or planned. Anticipation is key to this response, but remains a challenge, and some important categories of migration are still not the object of forecast. This practical guide offers concrete measures to introduce, run and strengthen migration forecasting systems and preparedness strategies. Drawing on country practices, research and the lessons of the OECD's multi-country Migration Anticipation and Preparedness (MAP) Task Force, it addresses the following questions: for what should forecasting tools be used, why migration needs to be forecast by category rather than only in aggregate, how to select models and data that fit specific purposes, how to credibly evaluate performance, and how to communicate results to policymakers so that forecasts can be used effectively and appropriately.

PRINT ISBN 978-92-64-64270-6
PDF ISBN 978-92-64-55962-2