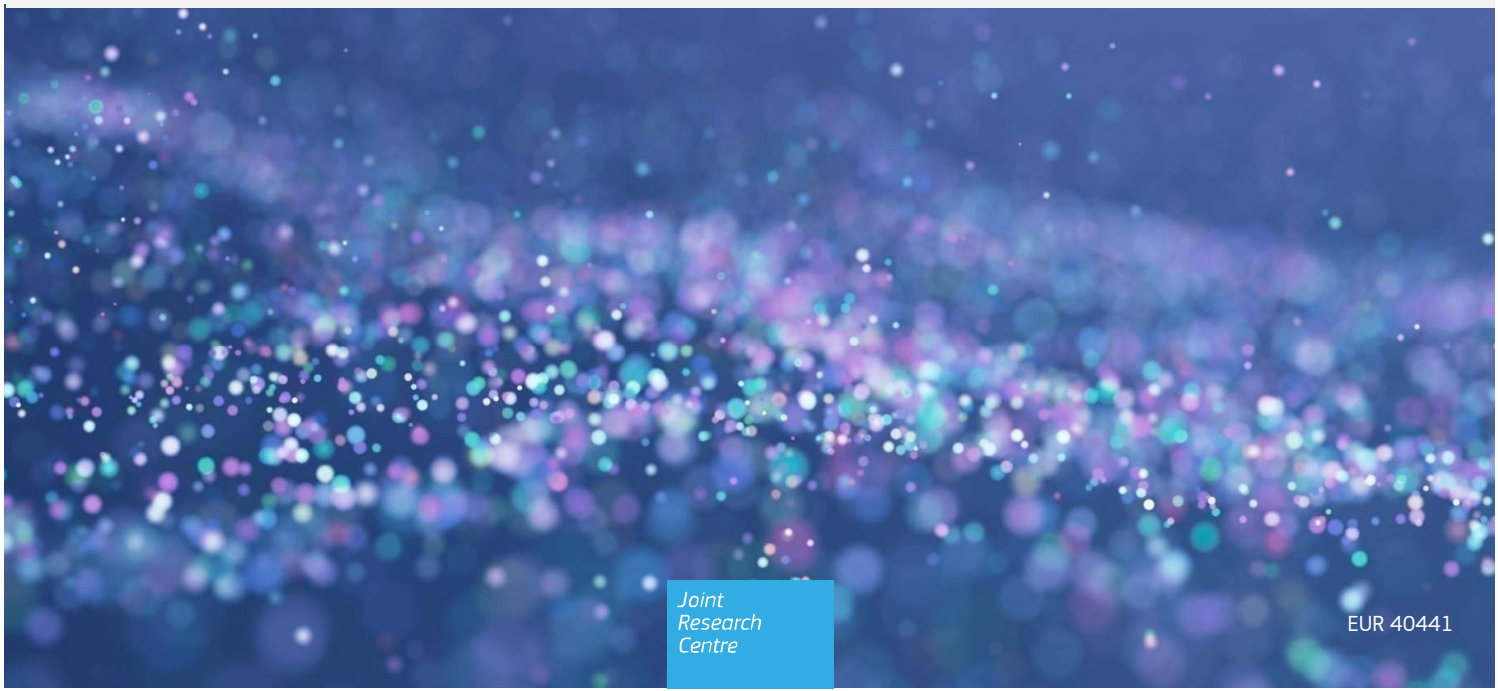
ISSN 1831-9424

Quantitative early warning models for mixed migration towards the EU

Pingsdorf, J., Minora, U., Di Salvo, M., Belmonte, M.

2025



EU Science Hub
https://joint-research-centre.ec.europa.eu

JRC142976

EUR 40441

PDF ISBN 978-92-68-31270-4 ISSN 1831-9424 doi:10.2760/6820414 KJ-01-25-446-EN-N

Luxembourg: Publications Office of the European Union, 2025

How to cite this report: Pingsdorf, J., Minora, U., Di Salvo, M. and Belmonte, M., *Quantitative early warning models for mixed migration towards the EU,* Publications Office of the European Union, Luxembourg, 2025, https://data.europa.eu/doi/10.2760/6820414, JRC142976.

Contents

Abstract

Anticipation and preparedness are key objectives of migration management policies at the EU level. The Migration Preparedness and Crisis Blueprint (Blueprint Network), overseen by DG Migration and Home Affairs, provides an operational framework for monitoring and anticipating migration flows, with the primary goal of supporting the EU's emergency and crisis response. Within this framework, the development of an early warning system is seen as an essential step to support preparedness and effective, coordinated responses to emerging migration situations. The European Commission is in the process of developing its own early warning methodology to monitor mixed migration flows. This report contributes to this development by proposing an early warning methodology grounded in an empirical analysis of existing approaches to anomaly detection in the policy context and the literature.

The proposed approach features two different early warning models which have been tested on two case studies, selected based on their potential relevance to policy makers. These are the number of asylum applications lodged by Turkish and Egyptian nationals. Results show that both models are effective in detecting early signals of important changes in the series, with one model limiting the rate of false alarms at the expense of missing a few real anomalies, and the other detecting a higher proportion of real anomalies but generating more false alarms.

The work suggests that these methods can provide a robust foundation for an early warning system to support preparedness at the policy level. Policy experts would likely benefit from such a system, as it enables timely monitoring and proactive responses to events that can impact the migration management system.

Acknowledgements

The authors would like to thank Fabiana Scapolo, Dario Tarchi and Alexander Bossakov for their insightful comments and discussions during the development of this report.

The authors would also like to thank participants to the technical workshop "Early warning methodologies for mixed migration flows", co-organised with DG HOME in February 2025. In particular, they would like to express their gratitute towards: Martin Taschner, Magnus Ovilius, Lorenzo Franceschelli, Mathieu Tartar, Alessio Baccari, Brunoro Calderico Buzzaccarini, Elena Fiore, Francesca Pavano, Ilker Aydin, Jonas Begemann, Luca Barani, Stella Riggi, Viviana Trinchi. Their expertise and feedback have been instrumental in shaping the content of this report.

Authors

Julia Pingsdorf

Umberto Minora

Mattia di Salvo

Martina Belmonte

1 Introduction

Anticipation and preparedness are key objectives of migration management policies at the EU level. The Migration Preparedness and Crisis Blueprint (Blueprint Network), overseen by DG Migration and Home Affairs (DG HOME), provides an operational framework for monitoring and anticipating migration flows, with the primary goal of supporting the EU's emergency and crisis response.[1] Within this framework, **the development of an early warning system is seen as an essential step to support preparedness and effective, coordinated responses to emerging migration situations.** This system should facilitate the timely and comprehensive understanding of events and emerging trends by relevant stakeholders, triggering the monitoring of changing situations and enhancing proactive responses.

To contribute to this objective, the European Commission is developing its own early warning system to monitor mixed migration flows. This report describes an early warning methodology that supports the development of such a system, grounded in an empirical analysis of existing approaches to anomaly detection in the policy context and the literature.

This report focuses on an early warning methodology for quantitative indicators; however before delving into the specifics of the proposed methodology, it is important to first define the key concepts of a more general early warning system that underpin this work.

1.1 Definitions

Early warning systems (EWSs) are processes designed to analyse and disseminate timely and relevant data and information, enabling individuals, communities, and organisations to anticipate risks and take proactive preparedness measures (UNDRR, 2007). There are many areas such as hydrological risks, food security, earthquake and tsunami events, or humanitarian crises in general, that are addressed with early warning systems at various levels of administration.[2] **Their primary goal is to detect changes early enough to allow for the implementation of adjustments or corrective measures, while minimising the number of false alarms** (Napierała et al., 2022).

In the context of the Blueprint Network's activities, an **Early Warning System refers to a mechanism designed to support preparedness efforts by anticipating crises or challenges through the identification of changing patterns and anomalies that point to potential migration**. It involves the collection, analysis, and interpretation of data and information related to migration trends, patterns, and drivers, as well as factors that may exacerbate migration flows, such as conflict, environmental disasters, or economic instability.

Anomalies are patterns in data that do not conform to a well-defined notion of normal behaviour. Anomalies can be **contextual:** this is the case when data is only unusual in a specific situation or context but would be considered normal in other contexts. The **contextual attributes** are used to determine the context (or neighbourhood) for that instance. In time-series data, for

[1] See: Commission Recommendation (EU) 2020/1366 of 23 September 2020 on an EU mechanism for preparedness and management of crises related to migration.

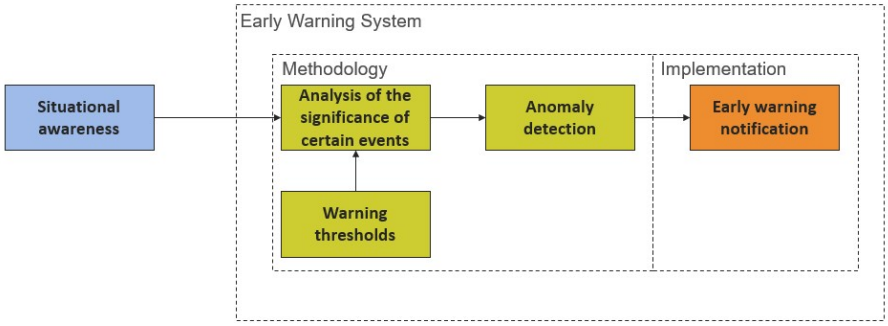
[2] For a review, see: Early Warning System—An overview | ScienceDirect Topics. (n.d.). https://www.sciencedirect.com/topics/social-sciences/early-warning-system

instance, time is a contextual attribute that determines the position of an instance within the larger sequence (Chandola et al., 2009). Examples of contextual attributes include seasonal or monthly patterns, where a value might be normal during one period but unusual during another.

Early warning notifications, in turn, result from a formalised process of applying a predefined methodology to identify anomalies, in order to identify and flag these unusual patterns.

The process of early warning (Figure 1) starts from being aware of what is happening in the EU and in third countries. Building on this first situational awareness step, however, **early warning involves an analytical component informing about the *significance* of some events for migration**. The significance is defined in relationship to warning *thresholds or criticality levels*. These can be pre-defined by analysts based on user needs or dynamically adjusted by the chosen anomaly detection technique. **When the threshold or a criticality level is reached, the event under monitoring becomes an *anomaly*, and an early warning notification is issued to users of the system**.

Figure 1. The early warning process



Source: JRC elaboration

1.2 Scope and categories of events

An Early Warning system for mixed-migration flows[3] in the EU aims to identify *significant changes* in events that are relevant for migration to the EU. Events are of two broad categories:

[3] According to the EMN Asylum and Migration Glossary, mixed migration flows refer to "complex migratory population movement including refugees, asylum seekers, economic migrants and other types of migrants as opposed to migratory population movements that consist entirely of one category of migrants [...] IOM defines mixed movements as a movement in which a number of people are travelling together, generally in an irregular manner, using the same routes and means of transport, but for different reasons". See: EMN. Mixed Migration Flow. https://home-affairs.ec.europa.eu/networks/european-migration-network-emn/emn-asylum-and-migration-glossary/glossary/mixed-migration-flow_en

1. Events in third countries, that are linked to migration in the EU, either because they concern drivers of migration (such as conflict, environmental disasters, or economic instability);[4] or because they indicate mobility outside the EU (e.g. migrants registered in traditional transit countries).
2. Migratory events in the EU – including events related to inflows (e.g., asylum applications or irregular border-crossings [IBC]); migration patterns within the EU (e.g., secondary movements) or the impact on migration management system (e.g., changes in the asylum backlog).

In the second case, early warning is about migration phenomena in the EU that are – to some extent – *already* occurring; in the first case, early warning is about migration in the EU that *might* occur.

There is a trade-off between *how early* anomalies are identified, and the *certainty* with which these are linked to migration events to the EU (Figure 2).

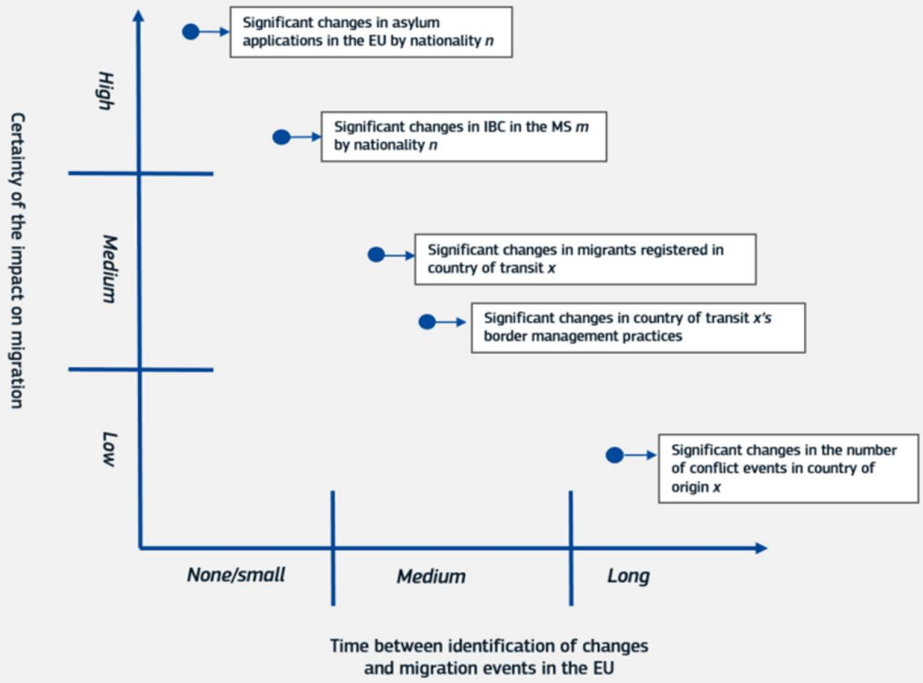
[4] The term "drivers" refers to factors or elements that influence particular outcomes, processes or phenomena. Analysing drivers of migration can provide insights on the complex and multi-dimensional factors that contribute to migration processes. By analysing anomalies in driver-related indicators, early warning systems can identify potential migration risks and facilitate timely interventions. For more information see: Rosinska A., *Drivers of migration: dimensions and indicators. A policy-oriented literature review*. Publications Office of the European Union, Luxembourg, 2024.

Figure 2. Examples of events plotted on a 'certainty of impact-time between the identification of changes and the migratory outcome' matrix.

Trade-off between timeliness and uncertainty

The plot displays some events along two main axes: the time between the identification of changes and the occurrence of migration outcome in the EU (horizontal axes), and the certainty of the impact of the event on migration flows in the EU (vertical). Below there is a list of some events that show the trade-off between timelines and uncertainty, from the top-left corner to the bottom-right cornet.

- The identification of significant increases in asylum applications in the EU from one particular nationality is a case of a certain 'anomaly' in migratory flows, yet it notifies about a migratory event (i.e. asylum applications) to the EU that is already occurring.
- A significant increase in irregular border crossings along a specific route is a certain migration event that may anticipate an increase in the number of asylum applications in country along the route (another migration event).
- The identification of changes in border management practices in countries of transit neighbouring the EU, or changes in the number of migrants registered in these countries, are less certain to generate migratory flows to the EU but provide more time to prepare early for such possible changes.
- The identification of changes in the security situation (e.g. the increase in the number of conflict events) of a country of origin is even less certain to generate migratory flows into the EU (especially if it is geographically distant), but provides even more time to prepare.



Source: JRC elaboration

1.3 Sources

Early warning systems can benefit from a multi-faceted approach combining both quantitative and qualitative sources of information.

- Algorithms on quantitative indicators can provide a set of early warnings triggered once thresholds are reached. For example, a warning could be automatically triggered once the number of irregular-border crossings within a given period of time reaches a record high.
- At the same time, early warnings can also be issued on the basis of qualitative information as primary sources. These qualitative inputs should be assessed and classified according to their level of criticality (in terms of their potential impact on migration), on the basis of pre-defined criteria. For instance, qualitative information on events in third countries could be classified according to the likelihood that the event will lead to a change in mobility (low to high) and the expected timeframe for the impact to materialise.

A **comprehensive assessment of a situation thus requires the integration of both quantitative and qualitative sources**. Findings from quantitative analyses can be validated and supplemented by qualitative insights, and vice versa. By combining multiple sources and perspectives, a more nuanced understanding of a given situation can be gained, including a deeper understanding of the underlying context and potential drivers of instability. Notifications on anomalous data patterns can point analysts' attention to specific countries or situations, prompting them to monitor these countries or situations by looking at developments from a qualitative angle to gain a deeper understanding of the underlying factors driving the quantitative changes. **Although this report focuses on quantitative analysis, using algorithms and quantitative indicators to identify potential early warnings, it is essential to acknowledge the importance of qualitative sources in providing contextual insights and validating quantitative findings.**

1.4 Objectives

This report defines an early warning methodology for quantitative indicators related to migration based on an empirical analysis of two main algorithms. It features the introduction of **a newly developed algorithm, in conjunction with a well-established anomaly detection technique**. The performance of these methods is validated against public historical data on asylum applications – a key indicator for mixed migration flows, highlighted in exchanges with policy experts as a priority area for early warning. Through the testing against historical data, this report assesses these methods' performance and provides insights into their implementation, paving the way for further testing and refinement of the methodology.

The report is structured as follows:

- Section 2 provides an overview of quantitative early warning methodologies currently in use in the literature and in the policy context.
- Based on this literature review and on consultations with experts, Section 3 outlines the key criteria that an EWS for migration should follow, and presents the two use cases that will be used to evaluate the proposed algorithms' performance.

- Section 4 outlines the proposal of two main methodologies for early warning: the multi-criteria algorithm and the CUSUM model.
- Section 5 evaluates the performance of the proposed algorithms by analysing anomaly detection in two case-studies: Türkiye and Egypt.

2 Methodologies for early warning

The definition of methods for anomaly detection is a crucial step in the development of an Early Warning System. A review of approaches to early warning currently implemented in the literature (Section 2.1) and in the policy context (Section 2.2.) is a logical first step for the development of a tailored solution within the Migration and Crisis Blueprint. On the basis of this review, Section 3 presents the key requirements of an EWS for migration, and Section 4 introduces a proposal for an early warning methodology.

2.1 Anomaly detection in migration studies

Anomaly detection in Early Warning Systems is an issue that has been researched and applied within diverse research areas and application domains (Chandola et al., 2009). There are many formulations of anomaly detection in the literature, each with its own advantages and disadvantages, and different fields of application, including migration.

Anomaly detection techniques can operate in one of three different modes: supervised, semi-supervised, or unsupervised. The choice of one mode over the other depends on the availability of data labels, that is, a tag that classify a data instance as normal or anomalous (Chandola et al., 2009), usually provided by humans. In summary:

- **Supervised Anomaly Detection**: data instances in a training set are labelled as either "anomaly" or "not-anomaly", and an anomaly detection model is trained on these labels to learn how to distinguish between the two classes on new (unlabelled) instances. This approach can be seen as a classification problem, since its aim is to replicate the way the labels were defined (e.g. by human assignment, Hyndman, 2024).
- **Semi-supervised Anomaly Detection**: similar to the Supervised mode, the Semi-supervised mode uses data labels, but only one type of label is available (either the anomalous class or the normal class).
- **Unsupervised Anomaly Detection**: it does not require labelled training data and it is based on the assumption that normal instances are far more frequent than anomalies in the test data.

In general, **unsupervised techniques are most widely applicable,** since obtaining quality labelled data is often very expensive.

According to Gordenker (1984) an EWS on forced migration should aim at both prevention and efficient advance preparations for a possible emergency. Bijak et al. (2017) and later Napierała et al. (2022) applied an unsupervised early warning model based on change-point detection to asylum applications data, inspired by statistical control theory (the "CUSUM" model set out by Page [1954]). Their proposed approach enables a shift of perspective, from a purely reactive towards a proactive approach for increased preparedness against asylum's potentially high societal impacts on receiving countries and the resource implications of asylum processes, providing a tool for policy and operational support, which necessitates exchange between analysists, decision-makers and modelers. A preliminary application of this technique was employed by the European Union Agency for Asylum (EUAA) to monitor the number of asylum applications lodged by different nationalities in EU+ countries (Melachrinos et al., 2023).

Carammia et al. (2022) developed an entirely data-driven EWS on a mix of administrative statistics and non-traditional migration-related data sources to effectively capture early warning signals of asylum-related migration. It is based on unsupervised techniques borrowed from financial econometrics, namely change point analysis,[5] the momentum approach and Lead-Lag correlations. Their system sets dynamic thresholds to trigger alerts, based on a moving average window of the latest data available, and works on a selection of 70 dyads of origin and destination countries (though it can be extended to include more if needed). The signals generated by this early warning model are then fed to a machine learning forecasting model aimed at estimating the future number of asylum applications in European countries of destination, aggregated by applicants' nationalities.

Sfyridis et al., (2017) used a semi-supervised One-Class Support Vector Machine (Scholkopf et al., 1999) trained with labelled instances of anomaly classes to detect potential smuggling vessels carrying migrants. In their work, the authors analysed over 6.7 million points, representing approximately 6,000 vessels of which three were labelled as "migrant vessel". These were used as labels for anomalies, expecting that the three vessels exhibit abnormal behaviour at some part of their trajectory compared to the other vessels, although normality was not defined.

Finally, Macis et al. (2024) developed an EWS to prevent conflict and unrest (which can be related to migration (Crippa et al., 2024), using an unsupervised anomaly detection method based on machine learning, specifically an Autoencoder model. They made use of the Armed Conflict Location and Event Data Project (ACLED) and Global Database of Events, Language, and Tone (GDELT) datasets due to their availability, coverage, frequency and update, and prioritized the identification of sudden outbreaks over precise numerical predictions of conflicts.

To effectively design an EWS for migration, it is also crucial to first understand the policy landscape in which it will operate. This includes examining the existing early warning methods currently informing policy decisions.

2.2 Early warning algorithms in the policy context

While the European Commission is still in the process of developing its own early warning methodology, it currently utilises the expertise and resources available through the Migration Preparedness and Crisis Blueprint (Blueprint Network).[6] Within this framework, the European Commission relies on insights and analysis provided by members of the Network, including the European External Action Service, selected Justice and Home Affairs (JHA) Agencies such the European Border and Coast Guard Agency (Frontex), the European Union Agency for Asylum (EUAA), the European Agency for the operational management of large-scale IT systems in the area of freedom, security and justice (eu-LISA), the European Union Agency for Fundamental Rights (FRA), and the European Police Office (Europol), in addition to contributions from Member States. The European Commission also takes into consideration information from key international organisations,

[5] Change point analysis aims at detecting where the probability distribution of a series or a process change based on some criteria (e.g. mean variance); for a description of the momentum approach and the Lead-Lag analysis refer to Murphy (1999) and Hoffmann et al. (2013) respectively.

[6] Commission Recommendation (EU) 2020/1366, of 23 September 2020 on an EU mechanism for preparedness and management of crises related to migration (Migration and Preparedness Blueprint).

such as the International Organisation for Migration (IOM) and the United Nations High Commissioner for Refugees (UNHCR).

In order to identify how data related to migration trends and patterns is currently processed, the focus was placed on publicly available outputs from key stakeholders.

The examples explored in the following sections are intended to provide an overview of different approaches employed or explored in the policy context, and are not meant to represent a comprehensive list of all early warning algorithms in use.

2.2.1 Migratory events in the EU

This section provides an **overview of publicly available analyses** from key stakeholders, highlighting **methodologies, algorithms, and thresholds used to monitor changes in asylum and migration trends within the EU** for early warning purposes. It focuses on the outputs of Frontex and the EUAA, the primary entities responsible for analysing illegal border crossings and asylum applications (respectively).

Most of the reviewed analyses found on official and public reports rely on a single criterion (typically percentage increases) to detect anomalies. Their reference periods can be month-on-month, quarter-on-quarter or focus on comparisons against a given period of time in the previous year. The level of disaggregation also varies depending on the indicator used, with some algorithms breaking down data by route, nationality, or Member State. Notably, these algorithms are static, meaning they do not adapt to changing patterns or trends over time. Table 1 summarises the analysis identified for two of the main indicators used to report migratory events in the EU: illegal border crossings (IBC) and asylum applications.

Table 1 Early warning algorithms for migratory events in the EU

Indicator	**Irregular border crossings (IBC)**	**Total asylum applications**
Source	Frontex[7]	EUAA[8]
General description	Percentage changes in the total number of IBC during a period in the current year, compared to the same period in the previous year.	Percentage changes in monthly/yearly asylum applications, by nationality and Member State, compared to the previous month/year.
Reference period	Same period (cumulated) in the previous year	Previous month or year
Disaggregation	Route	Nationality, Member State.
Condition for warning	Positive/negative changes could generate a warning	Positive/negative changes could generate a warning
Formula	Condition for ‘warning’ on IBC:	Condition for ‘warning’ on asylum applications:

[7] European Border and Coast Guard Agency (Frontex). (n.d.). Risk analysis for 2023/2024. Retrieved December 14, 2024, from https://www.frontex.europa.eu/assets/Publications/General/ARA_2023.pdf

[8] European Union Agency for Asylum (EUAA). Latest Asylum Trends. Retrieved December 14, 2024, from https://euaa.europa.eu/latest-asylum-trends-asylum; Asylum Report 2024. Retrieved December 14, 2024, from https://euaa.europa.eu/publications/asylum-report-2024

$$W = \frac{(IBC_{N,r,T_y} - IBC_{N,r,T_{y-1}})}{IBC_{N,r,T_{n-1}}} > 1\%$$

Where:
W = warning
N= All nationalities
R = route;
Ty=period of time in the reporting year, e.g. from January to March 2024
Ty-1= same period of time in the previous year, e.g. from January to March 2023.

$$W = \frac{(Asy_{n,MS,t} - Asy_{n,MS,t-1})}{Asy_{n,MS,t-1}} > 1\%$$

Where:
W=warning
n=nationality
MS= Member State
t= month or year of reporting
t-1= previous month or year

Source: JRC elaboration based on Frontex (2024) and EUAA (2024).

In addition, the EUAA also developed the Push Factor Index (PFI), which is based on the elaboration of event-based data from the GDELT Project[9] and provides important insights into migration processes in origin countries, bearing the potential for early warning purposes. This is part of a broader early warning and forecasting system (EPS–Forecasting)[10] that integrates administrative and non-traditional data at scale to effectively capture early warning signals of asylum-related migration, and delivers short-term forecasts of asylum applications from any country of origin to any European Union Member State (Carammia et al., 2022).

Finally, the EUAA also explored other approaches for monitoring time series such as combining the cumulative sum (or "CUSUM") method with the analysis of short-term and medium-term moving averages in the series borrowed from financial econometrics (EUAA, 2023). The CUSUM method is one of the two approaches for automatic early warning generation proposed in this report and it is detailed in Section 4.2.

2.2.2 Events in third countries

When it comes to information on the external dimension, the European Commission primarily relies on intelligence and analysis provided by staff seconded in EU Delegations (such as European Migration Liaison Officers [EMLO]), [11] the EEAS, and JHA Agencies, but also from IOM and UNHCR. Information related to developments in third countries frequently consists in analytical considerations provided directly by stakeholders in the field, which may concern displacement in third countries, but also other events linked to drivers of migration. This type of analysis encompasses a broad range of topics, spanning political, security, economic, and human development dimensions and may include information retrieved in local news reports or direct inputs from local counterparts and interlocutors in third countries.

To complement these insights, **quantitative measures or indexes focused on potential drivers of migration are sometimes used**. While changes in indicators related to migration drivers may

[9] https://www.gdeltproject.org/

[10] 'EPS' stands for Early warning and Preparedness System, a database containing provisional data exchanged between EU Member States and associate countries with which arrangements are in place pursuant to Article 34 of the EUAA Regulation and the EUAA, intended to give an estimation of the latest asylum trends across EU+ countries in as near to real time as possible (https://euaa.europa.eu/sites/default/files/2024-10/MBD_153_2023_EPS_databases_access_policy_FINAL.pdf).

[11] EMN Glossary, European Migration Liaison Officer, https://home-affairs.ec.europa.eu/networks/european-migration-network-emn/emn-asylum-and-migration-glossary/glossary/european-migration-liaison-officer-emlo_en

not be immediately reflected in migratory flows (Figure 1), **identifying anomalies can support preparedness by pointing towards changing situations in countries of origin or transit**.

The sections that follow draw on a review of publicly available information, to illustrate the variety of indicators and methodologies for early warning that could be integrated into an Early Warning System for mixed migration flows. These methodologies focus primarily on drivers of international migration, including conflict and political instability, humanitarian crisis, internal displacement.

2.2.2.1 The Science4Peace Portal: early warning on conflict risks

Science4Peace is a restricted resource managed by the JRC, which provides four analytical tools for early warning and crisis management:

- The Global Conflict Risk Index (GCRI) and the Dynamic Conflict Risk Model (DCRM) are quantitative models which evaluate on a yearly or monthly basis the risk of internal conflict in third countries and feed into the EU's conflict Early Warning System[12],
- Conflict Event Trends (CET) offers a monthly analysis on conflict and protest events,
- The Global Crisis Atlas (GCA) is a catalogue of conflict and crisis maps, which supports situational awareness and preparedness.
- The platform also features country profiles which integrate data from these analytical tools for 140 countries (Schvitz, G. et al, 2023).

Due to their relevance for early warning, methodologies for the GCRI and the DCRM are summarised in Table 4.

Table 2 The Global Conflict Risk Index & the Dynamic Conflict Risk Model

Tool	GCRI	DCRM
General description	Estimates the risk of violent conflict in 140 countries over the next four years.[13]	Estimates the monthly risk of violent conflict in sub-national regions over the next six months
Conflict risk definition	Probability and potential intensity of violent conflict (expressed in battle-related deaths).	Likelihood of either the presence or absence of conflict in each sub-national administrative unit.
Output	Total conflict risk score (derived by multiplying probability and intensity estimates).	Risk probability score.
Variables in the model	22 variables representing structural risk factors of conflict, divided into 6 categories: political, security, social,	**1. Event data**, including past conflict events, protests, and related incidents. **2. Structural variables**, which include local demographic and geographical data, as

[12] The EU's conflict Early Warning System is a tool managed by the European External Action Service (EEAS) to identify and assess the risk of violent conflict in third countries, enabling timely preventive measures before situations deteriorate into crises. For more information see: EEAS (2022). Early Warning System. Fact sheet. https://www.eeas.europa.eu/sites/default/files/documents/Factsheet%20-%20EWS.pdf

[13] For example, the 2024 update of the GCRI estimates conflict risks from 2024 to 2027.

	economy, geographical/environmental and demographic.	well as political and economic variables at the country level.
Conflict categories	**State-based conflict (SBC):** Armed conflict between a state government and at least one organized group. **Non-state conflict (NSC):** Armed conflict between two groups, neither of which is directly affiliated with a state government. **One-sided violence (OSV):** Direct and deliberate killing of civilians, carried out either by a state government or a non-state armed group. **Any conflict (ANY):** includes all instances of internal conflict that belong to the three categories above.	
Data source	Uppsala Conflict Data Program (UCDP)	
Forecasting window	4 years	6 months
Update frequency	Yearly	Monthly
Disaggregation	National level	Sub-national level

Source: JRC elaboration

2.2.2.2 The Integrated Food Security Phase Classification: early warning on food insecurity

The Integrated Food Security Phase Classification (IPC) is a tool developed by various partners including national administrations, UN agencies and civil society organisations.[14] The classification proposes common global scales to analyse food security and nutrition conditions, considering factors such as food availability, access and utilisation. By providing a common language (i.e., a standard classification system for the food security and nutrition situation), the IPC allows to compare situations across countries and over time. The methodology is applied to three different scales[15]: IPC Acute Food Insecurity, IPC Chronic Food Insecurity, and IPC Acute Malnutrition and proposes severity phases for each of them, with specific assessment criteria or thresholds applied (See Annex 1).

IPC data is publicly available and regularly updated (typically every 6 months). It aims to provide timely pictures of the food security situation across countries to inform policy-making and prompt both early responses and long-term programming (IPC, 2024).

[14] See: IPC Overview and Classification System | IPC - Integrated Food Security Phase Classification. (n.d.). Retrieved June 17, 2024, from https://www.ipcinfo.org/ipcinfo-website/ipc-overview-and-classification-system/en/

[15] Acute Food Insecurity refers to a sudden or short-term crisis that puts lives or livelihoods at risk, regardless of the cause. Chronic Food Insecurity is a persistent or recurring problem that prevents people from accessing adequate diets, often due to underlying structural issues. Acute Malnutrition that affects children, characterised by a high prevalence of malnutrition, illness, and inadequate food consumption. The distinction between these three categories is based on the recognition that each situation demands a unique response and affects distinct populations, necessitating tailored interventions to address their specific needs. For example, while actions targeting acute food insecurity should aim at urgently reducing food gaps, measures targeting chronic food insecurity should focus on the underlying factors. For more information see: Understanding the IPC Scales. https://reliefweb.int/report/world/understanding-ipc-scales

2.2.2.3 HDX Signals: automatic notifications on early warning

In 2024, the United Nations Office for the Coordination of Humanitarian Affairs (OCHA)'s Centre for Humanitarian Data launched an early warning service: the Humanitarian Data Exchange (HDX) Signals. The service monitors a selection of humanitarian datasets and sends automated e-mails when specific criteria or thresholds are reached.

Table 3 HDX Signals. Datasets and early warning criteria[16]

Dataset	Condition for warning	Frequency
Anomaly Hotspots of Agricultural Production (ASAP) - European Commission – Joint Research Centre	Signals are triggered when new agricultural anomaly hotspots[17] are detected.	Monthly
Armed Conflict - Armed Conflict Location & Event Data Project (ACLED)	Signals are triggered when monthly fatalities reach levels not seen in the past one or three years, and are higher than 100.	Monthly[18]
Acute food insecurity, from two sources: • **Integrated Food Security Phase Classification (IPC)** • **Cadre Harmonisé (CH)[19]**	Signals are triggered when: • There is an increase in populations in Phase 3 and at least 20% of the population are in phase 3 or above, or at least 5% of the population are in Phase 4. • Any populations are estimated to be in catastrophic/famine conditions (Phase 5).	Every six months
Conflict and disaster-driven population displacement,	Signals are triggered when monthly displacements reach levels not seen in the past	Monthly[20]

[16] United Nations Office for the Coordination of Humanitarian Affairs (OCHA), Centre for Humanitarian Data. (2024, June 12). Datasets | HDX Signals. https://un-ocha-centre-for-humanitarian.gitbook.io/hdx-signals/datasets

[17] Anomaly hotspots refer to areas that experience anomalous agricultural production patterns (both in terms of crops and livestock). Through monitoring timely information on crop growing conditions, ASAP allows for the early detection of production deficits. For more information see: Meroni M. et al., (2019). The warning classification scheme of ASAP – Anomaly hot Spots of Agricultural Production, v4.0, https://agricultural-production-hotspots.ec.europa.eu/files/asap_warning_classification_v_4_0.pdf

[18] ACLED datasets are transformed into time series of events and fatalities. Monthly fatalities are calculated as the rolling sum of fatalities in the previous 30 days. Armed conflict | HDX Signals. (2024, June 12). https://un-ocha-centre-for-humanitarian.gitbook.io/hdx-signals/datasets/armed-conflict

[19] Cadre Harmonisè is a tool for analysing and identifying food and nutrition insecurity risks in the Sahel and West Africa produced by the Aghrymet Regional Centre of the of the Permanent Interstate Committee for Drought Control in the Sahel (CILSS). The CH uses the same analytical framework and severity scales as the IPC. Cadre Harmonisé. (n.d.). Manuel Version 3.0. Analyse et identification des zones à risque et des populations en insécurité alimentaire et nutritionnelle. Retrieved July 23, 2024, from http://cadreharmonise.org/

[20] IDMC data is transformed by removing duplicate events based on the event's ID. Data is then transformed into time series and monthly displacements are calculated as the rolling sum of displacements during the previous 30 days.

Internal Displacement Monitoring Centre's (IDMC)	one or three years, and are higher than 5,000 for conflict driven events or 50,000 for disaster-driven events.	
Global Market Monitor dataset, World Food Programme (WFP).	Signals are triggered when month-to-month price increases in a given location reach the "high" or "severe" level from a lower classification in the previous month, indicating a worsening situation.[21]	Monthly

Source: JRC elaboration based on OCHA (2024)

The system complements its automated signals with AI-generated background information, using Open AI's GPT-4.0 to produce headlines and summaries for signals emerging from datasets containing publicly available text-based information (OCHA, 2024).

Users can subscribe to the service's e-mail notifications by selecting datasets and regions of interest. HDX Signals e-mails include a headline, location of the signal, information on the signal's trend, links to further information, and contextual information. A signals' map is also available on HDX's website.[22]

2.3 Advantages, limitations and alternative approaches

As shown in Section 2.2, early warning algorithms used in the policy context related to migratory events into the EU often rely on one criterion, and tend to compare figures with either the same period in the previous year or with the previous period. These methods can provide immediate inputs for analysts working in the policy context to monitor trends and anomalous patterns in data, facilitating their work. Additional straightforward algorithms to support policymakers could include identifying the highest or lowest value within a given timeframe (e.g., the past quarter or year). The simplicity of these approaches can be an added value, as they can be easily understood and interpreted by policy officials.

However, these types of algorithms have inherent limitations. Using only growth rates to define anomalies, for example, might undermine the identification of different types of evolving situations: they are sensitive to changes at lower numbers, yet they might fail to detect anomalies when numbers are high but not sufficiently growing compared to the reference period used (e.g., previous year, previous month - Barker and Bijak, 2023). In other words, growth rates can be helpful to detect significant changes in the numbers compared to previous levels, but less reliable in capturing whether numbers represent a high volume or not.

Evaluating figures in terms of standard deviations[23] above the historical average is an approach used in the literature to assess the volume of current trends. However, historical standard deviations can be affected by outliers in the time series. Analysing the case of Syria, for instance, Barker and Bijak

Internal displacements | HDX Signals. (2024, July 22). https://un-ocha-centre-for-humanitarian.gitbook.io/hdx-signals/datasets/internal-displacement

[21] The Global Market Monitor of the WFP monitors market food prices across a range of countries. Price changes are classified as 'normal' (between 0%-3%), 'moderate (3%-10%), 'high' (10%-25%) or severe (≥25%).

[22] See: United Nations Office for the Coordination of Humanitarian Affairs (OCHA), Centre for Humanitarian Data. (n.d.). HDX Signals Map. Retrieved July 26, 2024, from https://data.humdata.org/visualization/signals/

[23] The standard deviation is a measure of how spread out a set of numbers is from its average value. It shows how much individual values tend to differ from the average, with a higher standard deviation meaning more variation.

(2023) show that due to the 2015 peak, an important outlier in the overall historical trend, the number of monthly asylum applications in the following years was falling below one standard deviation from the historical average despite often exceeding 7,500, which is a rather high number of monthly applications. In addition, these types of algorithms may require more time and expertise to be interpreted, which may hinder their usability in policy environments, where timely decision-making is crucial.

In summary, **different methods are better suited to meet certain needs at the policy level or to capture different features in the time series (whether these are significant changes or high volumes)**. Some flexibility is necessary when defining the time frame used to assess current figures. On the one hand, algorithms need to avoid detecting too many changes that seem significant in the short term but are not when put into context: for instance, IBCs and asylum applications in 2021 might report a large percentage change compared to 2020, but only because of generally very low arrivals during 2020 due to the COVID-19 pandemic. On the other hand, algorithms should avoid missing changes that would be relevant but are not signalled because of significant outliers in the past: for instance, considering the historical standard deviation, as often done in the literature, can hide new developments due to significant variations in outlier years (e.g., "migration crisis" of 2015 and 2016).

Additionally, **single-criteria algorithms tend to produce binomial "yes" or "no" warnings, which may fail to capture the nuances and complexities of evolving situations across different contexts**, potentially leading to oversimplification and reduced interpretability of the anomalies. **An algorithm which results in different warning levels could allow for more meaningful insights.**

In addition, **the early warning methodology proposed in this report (see Section 4) opts for an unsupervised approach** for two key reasons:

- It can be applied without the need for data labels, making it more adaptable to different sets of indicators;
- It is not prone to bias due to misclassification of the data instances (i.e. when the label "anomalous" is not anomalous in a statistical sense).

Data labels are used, however, to evaluate the performance of our approach, comparing the output of our models to labels from human analysts to ensure the results meet user expectations. This evaluation will be discussed further in Section 5.

Before presenting in detail the proposed early warning methodologies, the next chapter identifies the key criteria that such a system should meet. Building on the literature review conducted in this section, it integrates insights from policy experts to further refine these criteria.

3 Early warning for migration policy. Key requirements

Building on the review of early warning methodologies presented in Chapter 2, the JRC consulted with European Commission policy experts specialising in migration and asylum. The objective of the consultations was to gather feedback from potential end-users to inform the development of an early warning system for mixed migration flows. Through both the methodology review and the exchange with policy analysts, key criteria were identified that a quantitative methodology should fulfil to effectively support policy stakeholders. Additionally, the discussions revealed specific use cases where policy officers would have benefited from early alerts.

3.1 Key criteria

To effectively support policy stakeholders, an early warning system for mixed migration flows should be designed with the following key criteria in mind:

1. **Simplicity**: the early warning system should be grounded on a methodology that is transparent and easy to interpret. The system should provide clear and concise explanations for why particular values are flagged as anomalies, allowing system users to effectively analyse its outputs and use them for informed decision-making.

2. **Timeliness**: by definition (see Section 1.1), the early warning system should detect anomalies early enough to allow for corrective actions. Users of the system should be made aware of anomalous situations before they escalate into crisis or pose serious challenges. The selected use cases (see Section 3.2) illustrate anomalous situations in which policy analysts would have wanted to have received an early warning notification.

3. **Relevance**: indicators covered by the system should be closely related to migration, providing a logical link between the data analysed and potential inflows to the EU. To achieve this, the focus should be placed on indicators that capture both mobility patterns and drivers of migration. While the early warning methodology itself does not allow to establish causal links between these indicators, the system should provide a coherent narrative that supports users in understanding the relationship between the monitored indicators and emerging situations related to migration into the EU.

4. **Flexibility**: the early warning methodology should be flexible and adapt to:

    (a) Specificities of different data series: the system should be able to cater for different datasets, with different orders of magnitude, fluctuations and seasonality patterns.

    (b) Uncertainty: an early warning system for migration operates in a dynamic and unpredictable environment, where sudden, unexpected changes in data cannot be excluded. To navigate this uncertainty, the system's parameters for anomaly detection and the selected thresholds should be designed to be flexible and adaptable, allowing for adjustments when faced with rapidly changing situations.

    (c) Benchmarks: in line with the inherent uncertainty mentioned above, it is important that the benchmarks used to define "normal behaviour" in the data can be adjusted to reflect changing circumstances.

5. **Anomaly detection goal:** all early warning systems should be designed to minimise false positives (that is, pointing at anomalies that aren't there) and false negatives (failing to detect genuine anomalies). On the one hand, excessive false positives may erode confidence in an Early Warning System. On the other hand, false negatives could result in delayed responses to real challenging situations or crisis. For instance, if an early warning system failed to detect a surge in asylum applications from a specific country (false negative), it would fail to alert policy stakeholders about a potential vulnerability, which could lead to untimely resource allocation and response planning. If the system incorrectly identified a period of stability as anomalous (false positive), this could lead to unnecessary resource mobilisation and preparedness. It is also crucial that users of the system are made aware that all anomalies may be isolated incidents that may not necessarily translate into an emerging situation. The role of contextualising these anomalies and take appropriate responses is therefore delegated to the final users, not to the Early Warning System in itself.

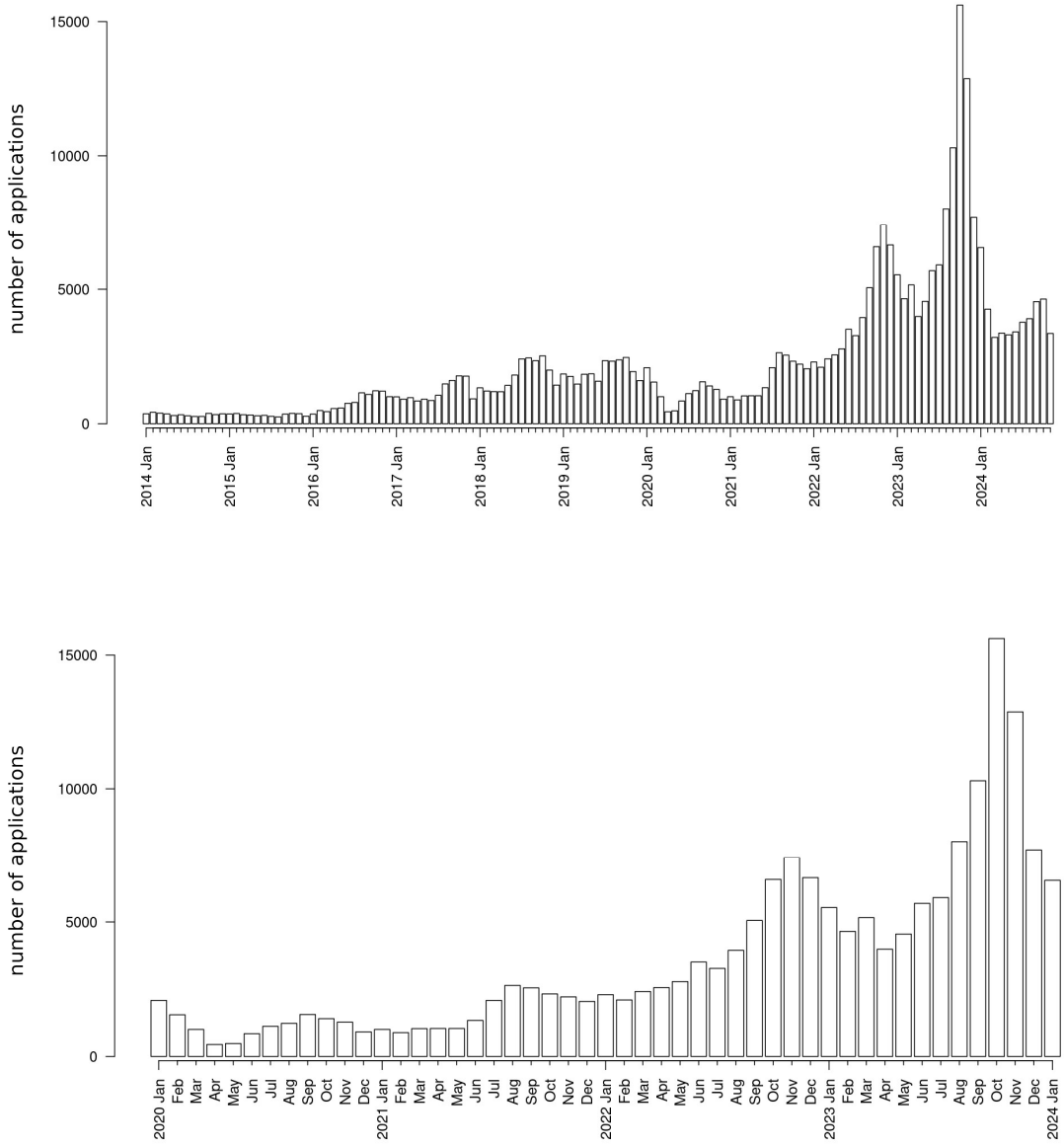
3.2 Use cases

Two use cases were selected, focusing on recent migration trends involving specific nationalities for which the issuance of early warning notifications could have been useful. These cases were selected based on their potential relevance to policy makers and were discussed with European Commission policy analysts during consultations to gather their input and perspectives. The focus on asylum applications was also informed by these exchanges, who highlighted the need and usefulness of early warning on this indicator.

Asylum applications from Turkish nationals

According to Eurostat data, in 2023, Turkish nationals submitted a record 94,000 asylum applications in the EU, making Türkiye the third country of origin that year. This was the second consecutive year that Turkish applications nearly doubled, continuing a steady upward trend since 2020. Given the relevance and steadiness of this increase, and Türkiye's proximity to the EU, it can be safely assumed that policy analysts could have benefitted from a warning on anomalous data patterns in applications from Turkish nationals, especially **prior to the all-time high of 15,600 in October 2023.** This peak, which far exceeded historical levels, **could have been anticipated to some extent by analysing the trends in the previous two years (2021-2022) and considering qualitative information on events and developments linked to migration drivers in Türkiye.**

Figure 3. Asylum applications in the EU by Turkish nationals, 2014-2024



Source: JRC elaboration based on Eurostat migr_asyappctzm

These insights further reinforce the value of using the country as a case study and highlight how analysts could have benefited from early warnings on anomalous asylum application trends. For example, economic instability, particularly rising inflation since mid-2020, has been a recurring theme in news reports. The Turkish lira was severely impacted during the Covid-19 pandemic, underperforming compared to other emerging market currencies.[24] Economic turbulence was followed by social unrest and increasing protests, some of them including police confrontation with protesters.[25] At the beginning of 2023, the country was hit by two strong earthquakes which resulted in over 50 thousand fatalities and hundreds of thousands of people displaced.[26] This non-exhaustive overview serves to illustrate that **by combining early warnings on anomalous asylum application numbers with qualitative assessments of events linked to drivers of migration, analysts could have been better equipped to anticipate and interpret the surge in applications earlier on**.

Asylum applications from Egyptian nationals:

Although smaller in scale compared to Turkish applicants, applications by Egyptian nationals also showed unprecedented numbers in the EU in both 2022 (15,600) and 2023 (26,600). Historically, Egyptian applications to the EU averaged around 4,800 per year from 2014 to 2019, but declined to 3,360 in 2020 due to the Covid-19 pandemic. However, this trend reversed in 2021, with applications reaching a record high of 6,700, paving the way for the substantial increases seen in the following 2 years. First-time applications from Egyptian nationals reached **several monthly all-time highs during this period,** the most significant peaks being March (1,315) and November (1,730) 2022, and January (2,165), March (2,405) and October (2,585) 2023. As was the case with Türkiye, policy experts **could have anticipated this trend to some extent had they been alerted of anomalous data patterns as early was 2021**, potentially allowing for earlier intervention and a more proactive response.

Similarly to the Turkish case, early alerts on anomalies in applications from Egyptian applicants could have been cross-checked against developments related to drivers of migration in this country. For instance, Egypt's economic instability, including a significant depreciation of the currency and rising inflation since 2021, have been widely reported. The country's economic issues were exacerbated by the Russian invasion of Ukraine, which led to disruptions in wheat imports and further exacerbated food price increases and water scarcity (EUAA, 2022). Additionally, the human rights situation in Egypt has been a subject of concern.[27]**This is another case in which early alerts on anomalous data**

[24] See: DW, "Why the Turkish currency is in free fall", 17th August 2020. https://www.dw.com/en/erdogans-credit-binge-fuels-lira-depreciation/a-54524078

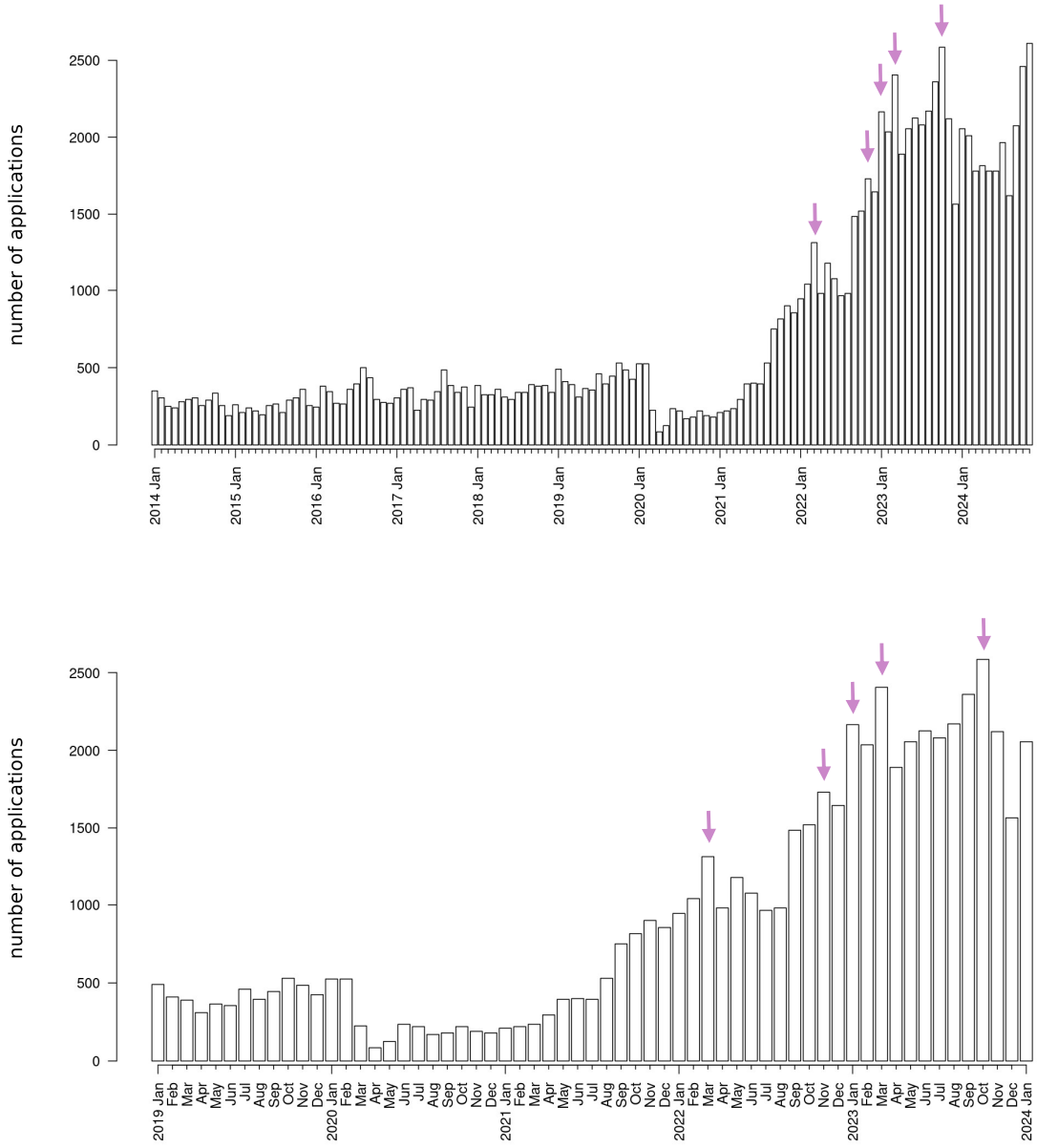
[25] See: Focus - Inflation in Turkey: Soaring prices, housing crisis and student protests. (2021, October 18). France 24. https://www.france24.com/en/tv-shows/focus/20211018-inflation-in-turkey-soaring-prices-housing-crisis-and-student-protests; Turkish police tear-gas women protesting over violence. (2021, November 25). AP News. https://apnews.com/article/europe-middle-east-violence-turkey-recep-tayyip-erdogan-37dc9465b1f867b00465f8006a74d733; Protests in Turkey's cities after lira plunges to new historic low. (2021, November 25). Euronews. https://www.euronews.com/next/2021/11/24/turkish-lira-plunges-to-a-historic-low-and-protests-erupt-after-erdogan-defends-rates-cuts;

[26] See: Human Rights Watch, World Report 2024: Türkiye. Retrieved 28 March 2025 from: https://www.hrw.org/world-report/2024/country-chapters/turkey

[27] See: 46th session of the United Nations Human Rights Council: General Debate Item 4: Human Rights Situations that require the Council's attention, Joint Statement on Egypt. 12 March 2021. https://reliefweb.int/report/egypt/46th-session-united-nations-human-rights-council-general-debate-item-4-human-rights

points, combined with insights on relevant developments in the country of origin, could have supported policy analysts in the monitoring of growing situations.

Figure 4. Asylum applications by Egyptian nationals in the EU, 2014-2024



Source: JRC elaboration based on Eurostat migr_asyappctzm

Having outlined the essential requirements for an early warning system and case studies that can support the testing of the proposed methodology, Chapter 4 places the focus on two algorithms that can underpin the development of a methodology for an EWS for migration.

4 Algorithm proposals

This section begins with a proposal for an algorithm that considers **multiple criteria at the same time** with the aim of providing a more complete picture of anomalies compared to single-criterion algorithms (e.g., percentage changes). The proposed algorithm also adopts **different thresholds** for each criterion, so to provide further flexibility in detecting new trends as well as mitigate the uncertainty that comes with defining strict thresholds that need to be applied across different contexts. Finally, the three proposed criteria combine **different time frames** to assess current figures against previous values.

In addition, **Section 4.2 proposes a complementary approach based on cumulative sum (CUSUM) control charts**, a statistical process control method capable of identifying changes in the data generating process and issuing timely warnings of emerging anomalies.

4.1 A multi-criteria algorithm

The proposed **multi-criteria algorithm** aims, on the one hand, at capturing anomalies that reflect both high volumes and significant changes and, on the other hand, at better contextualising current figures with respect to recent developments and historical trends.[28]

The multi-criteria algorithm uses the following three criteria:

- Monthly year-on-year percentage change.
- Percentiles of the same month in the previous three years.
- Rolling standard deviations above the mean of the previous twelve months.

The proposed algorithm uses a set of thresholds for each criterion to assign an overall "intensity" score to each observation, allowing for a nuanced interpretation of evolving situations across different contexts and mitigating the uncertainty of pre-defining strict thresholds.

4.1.1 Criterion 1: Percentage change

The first criterion, the percentage change[29] compared to the previous year, is among the most used in methodologies applied in the policy context. It captures the **significance of a change** in a given month compared to the same month of the previous year. Among the three criteria used in the proposed methodology, it is the one with a **medium-term** reference period (i.e., last year), comparing figures in the same month to account for possible seasonality in the data.

[28] As shown by Barker and Bijak (2023) for the Syrian case, using single criterion on entire time-series can affect detection of anomalies due to past outliers.

[29] $(X_{n,m} - X_{n,m-12}) / X_{n,m-12}$; where X is the data series, n the nationality and m the month.

4.1.2 Criterion 2: Percentile distribution

The second criterion, the percentile distribution[30] of the same month in the last three years, instead, aims at capturing the **size of current figures** compared to the historical trend, thus adopting a **long-term** reference period. A 'high' volume is relative to each context (e.g., nationality, route, etc.) and, thus, absolute values might not be the best option to define thresholds. For instance, when analysing the case of Syria and Ukraine, Barker and Bijak (2023) applied two different thresholds of monthly asylum applications (300 for Syria and 50 for Ukraine) in their model depending on the historical trend of each nationality. In other words, **there cannot be a one-solution-fits-all using absolute values**. The use of the percentile distribution (e.g., above the 80% of previous values) thus tries to overcome this issue: assessing current figures against historical consolidated values should help evaluate whether a figure is 'high' enough given each specific context. Of course, on a case-by-case evaluation, systems can be tailored to exclude values considered to be low by using additional filters on absolute figures. Similarly, while the tested approach compares the same month in the previous three years to account for seasonality of data, alternative specifications considering more months could be tested in future applications to evaluate possible differences in anomalies detected.

4.1.3 Criterion 3: Rolling standard deviation

Finally, the third criterion, the rolling standard deviation[31] above the mean over the previous year, aims at **evaluating current figures against recent developments**. Among the three criteria, in fact, it is the one with the **shortest reference period**,[32] thus adding information not considered in the other two criteria. Significant deviations above the mean of the past year can help understand whether figures are starting to be high compared to the recent past, thus capturing emerging shifts in time series. In addition to account for recent developments, using the previous twelve months as reference period helps eliminating the bias introduced by past outliers with exceptional variation (e.g., 2015 for the case of Syria).

4.1.4 Thresholds and scoring

Besides relying on three criteria, the proposed algorithm also uses a set of thresholds for each criterion, resulting in an overall score reflecting the "intensity" of anomalies detected (Table 4). This approach tries to overcome binomial "yes" or "no" types of warnings and to add nuances that might help the interpretation of evolving situations across different contexts. Moreover, using a set of thresholds can help mitigate the uncertainty that comes with defining strict thresholds that need to be applied across different contexts and, possibly, types of data.

Table 4 Criterion and thresholds of the multi-criteria algorithm

criterion \ score	**1**	**2**	**3**
% change $_{t-12}$	25%	50%	75%

[30] $X_{n,m}$ above N^{th} percentile of $X_{n,m-12}$, $X_{n,m-24}$ and $X_{n,m-36}$; where *X* is the data series, *n* the nationality and *m* the month.

[31] $X_{n,m}$ above N standard deviations above the mean of $X_{n,m-1}$ to $X_{n,m-12}$; where *X* is the data series, *n* the nationality and *m* the month

[32] The rolling standard deviation criterion relies on data from the previous 12 months, meaning information up to the previous month are taken in consideration, therefore looking at the short-term developments.

percentile $_{t-12,\ t-24,\ t-36}$	70th	80th	90th
rolling sd $_{t-1\ to\ t-12}$	1	1.5	2

Source: JRC elaboration. Note: for each threshold the condition is strictly above, not equal or above

For each criterion, three increasing thresholds are defined and assigned a score from one to three (Table 4). For instance, a month recording a percentage change with the previous year above 75% (3rd threshold, +3) could be signalled as an anomaly with a score of 7 when either of the other two criteria overcomes its third threshold (+3) and the other one its first one (+1). For instance, figures can be above the 90th percentile (thus in the top 10%) of the same month in the past three years (3rd threshold, +3) and above one standard deviation from the last 12 months' mean (thus top 16%)33 (1st threshold, +1).

Hence, **the algorithm assigns an overall "intensity" score from 0 to 9 to each observation summing up each criterion's score, with the observation being flagged as anomaly if the overall score is at least three** (**Figure 5**).

Figure 5 Intensity score and anomaly detection – multi-criteria algorithm

Score	0	1	2	3	4	5	6	7	8	9
Anomaly	No			Yes						

Source: JRC elaboration

The decision to set a minimum value of three for the algorithm to signal an anomaly is driven by the need to balance, on the one side, the potential of the system to detect very different situations at early stages with, on the other side, avoiding too many anomalies.[34]

4.1.5 A flexible algorithm

In order to provide further flexibility in defining anomalies, a final distinction is made between a "loose" and a "strict" algorithm. While for the "loose" only the condition of an overall score of at least three is necessary, for a "strict" signal each of the three criterion must overcome one of its thresholds. In other words, an overall score of three due to one of the three criterion overcoming its third threshold would be flagged as anomaly under the "loose" algorithm but not under the "strict" one. On the contrary, if each criterion has overcome at least its first threshold, then the overall score of three would trigger a signal in both the "loose" and the "strict" multi-criteria algorithms.

[33] In a normal distribution, 68% of observations fall within + or - one standard deviation from the mean (84% within one and a half, and 95% within two standard deviations). Looking only at the right side of the distribution where values are higher than the mean, observations falling more than one standard deviation above the mean belong to the highest 16% (i.e., (100-68)/2). While these specific percentages are valid for data with a normal distribution, the standard deviation remains a measure of the deviance from the mean, thus indicating how spread out are the data from the mean (and thus how 'exceptional' a value is) regardless from the specific distribution of each data series.

[34] As for defining thresholds, some degree of uncertainty applies also when selecting a cut-off for the intensity scale. The design choice of one third of the intensity scale in this approach could be further validated in future research to ensure that the number of anomalies detected reflect users' needs.

The combination of "loose" and "strict" versions can be helpful in monitoring different types of situations evolving as well as changes in a situation that is already detected as being of relevance. For instance, the "loose" version can be more helpful in detecting early on signs of changes while the "strict" can be used to understand whether a situation is further intensifying since it was first detected or if it is losing momentum over time. A performance comparison of the loose and strict models against single-criteria algorithms can be consulted in Annex 5.

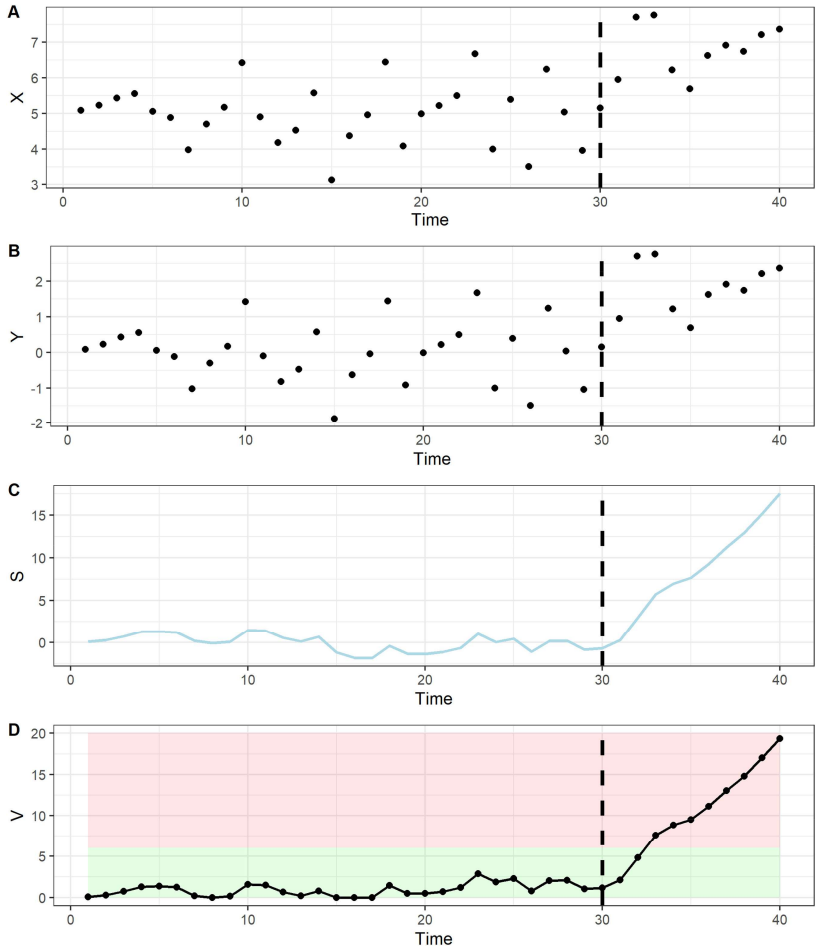
4.2 Cumulative sum (CUSUM) control chart

The **cumulative sum** (CUSUM) **control chart** is a Statistical process control (SPC) originally set out by Page (1954). This tool plots the cumulative sums of the deviations (*scores*) of the sample values from a target value, and an early warning can be triggered when the scores pass a certain limit (*threshold*). The CUSUM is effective for change-point detection of regime shifts at an early stage of the process being monitored, because it combines information from several samples (cumulative sums) rather than looking at each sample individually as some more basic SPCs do (e.g. Shewhart control charts).[35] It is used, for instance, to monitor whether a process is drifting away from its mean, and it is being applied in many different domains and applications, such as surveillance, quality control, and clinical-care processes, to name only a few.

The scores of a CUSUM model are calculated as cumulative deviations from a target value. For instance, if we are interested in understanding if there was a positive shift in the mean of a normally distributed process, we can set our target value to the mean (or median, μ_0) and the deviations are calculated as the difference between each *i*th sample and the mean (or median, $y_i = x_i - \mu_0$). A cumulative sum of such scores (s_i) would fluctuate around a central line (*i.e.* a random walk with mean or median zero) if the process is in control. If, instead, the process shifts upward, then an upward drift will develop in the cumulative sums, and a signal is emitted when the score exceeds the decision interval (H). In this case the process is said to be out of control. It is also common to reset the scores to zero (v_i) if they become negative because they are not relevant for determining when action should be taken. Figure 6 shows an example of the workflow to obtain a one-sided upper CUSUM chart from a time series X with a target value $\mu_0 = 5$, and an upper control limit (threshold) $h^+ = 6$. Points within the red area in the last plot are the signal that the process is out of control.

[35] Family of models often used for on-line process surveillance to monitor if the sample values fall within some control limits and do not exhibit any systematic pattern.

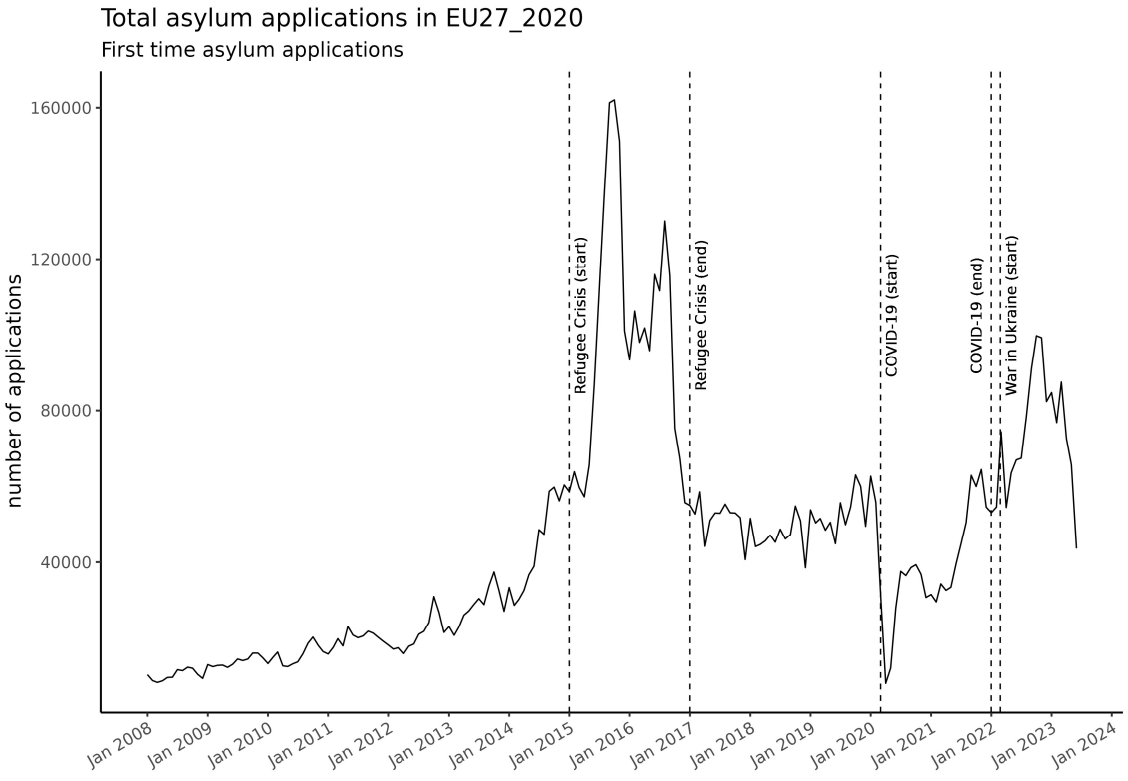
Figure 6. One-sided upper CUSUM control chart applied to data with a change point at time 30 (vertical dashed line). (A) input time series X; (B) score function Y; (C) cumulative sum S; (D) cumulative sum V, restarted at negative values, and upper control limit h = 6.



Source: Figure 1 in Napierała et al. (2022).

The crucial aspects for designing an effective CUSUM control chart are the choice of the target value and the setup of the control limits, because they will affect how sensitive (or not) the model will be to detect important change in the pattern of the data. In the previous example, we used the mean as the target value. Note that any value can be used as target as long as it is meaningful for the shift we are interested to monitor. However, when dealing with migration flows data and related indicators like the ones this report is focusing on, the idea of a constant and clear "target value" becomes vague as these processes are often not normally distributed and their values do not fluctuate around a constant mean or with a constant variance (*i.e.* they are "non-stationary"). To understand this concept better, let's consider the number of first-time asylum applications lodged in the European Union since January 2008 (Figure 7).

Figure 7. Total number of first-time asylum application in the European Union (monthly data). The vertical dashed lines represent key critical past or ongoing events.



Source: Eurostat.

In the figure above we see that the number of monthly asylum applications lodged in the EU has a complex and far-from-linear pattern: it started with a constant and steady increase in the period 2008 to 2014; saw a sudden surge between 2015 and 2016, a period knows as the “refugee crisis"; a rather “stable” period between 2018 and 2019, followed by the COVID-19 pandemic, that caused a marked drop in the series in early 2020, followed by an increase of the values in the subsequent months which most likely reflect the activity of the relevant asylum authorities that process the applications; finally, in February 2022, there is an increase due to the start of the war in Ukraine. Given the great variability of this time series, it would be very hard to set a static and clear threshold for the CUSUM model. There are various strategies to address these issues. First, even if the target value is unknown, it is still possible to estimate it “on the fly” using the samples in the data. Second, it is common to adjust it using a reference value K as multiplier (also known as the allowance, or the slack value). For instance, we can design a CUSUM control chart which monitors shifts that are of a distance of some multiple of the standard deviation (or the MAD) σ of the process from the mean (or median). As said, these values can be calculated on the fly using the historical data. Third, it is often advisable to restrict the information used for calibration–the data used to estimate the CUSUM parameters on the fly–to periods prior to the points we need to monitor. On the one side, longer calibrations allow for better understanding past developments and detect possible features such as seasonality. On the other hand, however, observations that are too far from the present might not reflect the current situation and can add important bias to the system. For this reason, it is always important to have a good understanding of the data being monitored and the context in which they are used to find an ideal calibration period.

4.2.1 CUSUM in the context of migration

Bijak et al. (2017) and later Napierała et al. (2022) assessed the potential of the CUSUM model in aiding the management of mixed migration flows, and a preliminary application of this technique was employed by the EUAA to monitor the number of asylum applications lodged by different nationalities in EU+ countries (Melachrinos et al., 2023). In this work, we are using the CUSUM control chart in a similar way, but with some modifications concerning the way we select the calibration period and the transformation of the data before they enter the model.

The way the CUSUM model is used in the context of this study is a real-time change-point detection system. In other words, **we are interested in monitoring a continuous stream of data and detect possible signals of change in pseudo real-time**. Therefore, the focus is not on the historical data, but in **understanding if the most recent observation is in control or not**. For this reason, we implement a moving window CUSUM model, where for each indicator to be monitored we set a fixed number of calibration samples which precede the most recent observation in the time series, and we move this window one step ahead each time a new observation is made available to the system. Each time this is done, we estimate new parameters on the fly from the calibration set. The parameters used in this report are reported in Annex 5. Finally, we reset the CUSUM value to zero every time it exceeds $h^{\pm}$. By doing so, we focus on providing a timely alert of a potential regime shift.

4.2.2 Dealing with seasonality, trends, autocorrelation and stationarity in time series

One important aspect common to many migration time series is that of seasonality. A seasonal pattern occurs when a time series is affected by seasonal factors such as the time of the year or the day of the week. If we consider the number of asylum applications as an example, there is usually a drop at the end of each year when the processing capacity of asylum authorities tends to be limited, followed by a subsequent increase (Carammia et al., 2022). Another example is given by the asylum flow from the Western Balkan to EU+ countries, where the highest levels generally occur in October of each year, plus, usually, lows in applications during June or July (EUAA, 2015). Therefore, when seasonality is an important and known feature of a time series, it is extremely important that an Early Warning model account for it, otherwise we risk triggering alerts which are false positives due to the seasonal effect.

One strategy to deal with seasonality is to remove the seasonal component from the time series before passing it to the Early Warning model. The "Seasonal and Trend decomposition using Loess" (STL) is a versatile and robust method for decomposing time series developed by Cleveland et al. (1990) and later extended to handle multiple seasonal patterns by Bandara et al. (2025). If the data has a clear seasonal pattern (e.g. with peaks and droughts occurring each year between Winter and Summer), an Early Warning model would probably trigger many alerts if we were not to account for these expected seasonal patterns. By adjusting for the seasonality effect, we can instead focus on the other types of variations in the data.

As the name suggests, STL decomposition also provides a trend component using the LOESS method, a locally-weighted regression for estimating nonlinear relationships. LOESS fits a smoothed function across the data points representing the low frequency variation in the data together with nonstationary, long-term changes (Cleveland at al., 1990). In anomaly detection, we are not really interested in these low frequency and long-term variations, rather in the sharp and short-term ones. Therefore, we can remove this component from the time series to improve the performance of the CUSUM model (and of other anomaly detection algorithms in general). By doing this, we also account

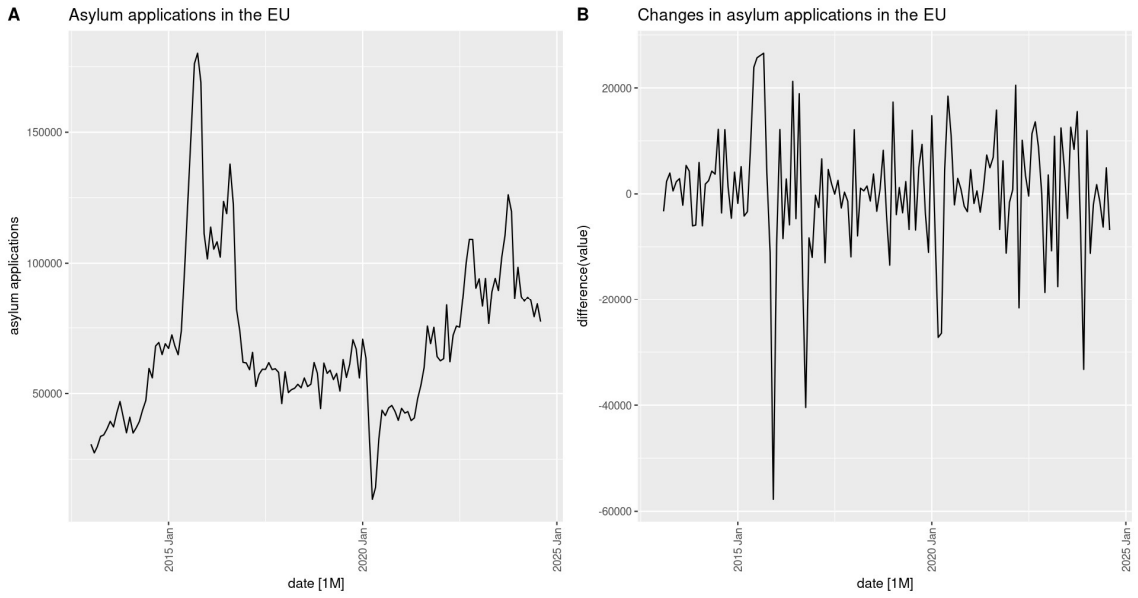
for the issue of autocorrelation to some degree, which is the linear relationship between lagged values of a time series, that may be the result of dynamics that are inherent to the process.

Positive autocorrelation can produce severe negative bias in traditional estimators of the process standard deviation, and this bias produces control limits that are much tighter than desired. This is because in presence of autocorrelation, closer observations tend to form clusters which mask the dispersion of the distribution around the mean, making it look smaller than it really is. This in turn would imply the estimation of a smaller standard deviation, on which the control limits are calculated. For these reasons, the presence of autocorrelation can result in an average false alarm rate much higher than expected or desired (Lu et al., 2001). In fact, the performance of CUSUM charts deteriorates when the process observations are autocorrelated. The popular way of monitoring changes in mean of an auto correlated series is to apply the CUSUM on the residuals of a chosen time-series model that best explains the dynamics of the process observations (Runger et al., 1995; Atienza et al., 2002). The LOESS method provided within the STL decomposition can be used in this regard, as it fits a time-series model and the residuals from the prediction of this model can be used for process monitoring.

A time series may present either a seasonal and long-trend component, just one, or neither of these two. It's up to the modellers and the analysts to find the best combination of parameters that would eventually lead to the desired anomaly detection results. In Section 5, we provide the results of applying the CUSUM model on a selection of use cases as identified in Section 3.2.

Finally, if the time series is not stationary (even after removing the seasonal and trend component, which should help in this regard), one common strategy to make it stationary is to compute the differences between consecutive observations (a technique known as *differencing*), which helps stabilise the mean of a time series (Hyndman and Athanasopoulos, 2021), essential condition for the CUSUM model. An example of the use of this application is given by **Figure 8**.

Figure 8. (A) Asylum applications in the EU; (B) Monthly change in asylum applications in the EU.



In **Figure 8** A, the asylum applications in the EU series represents a non-stationary series, with various trends and changing levels throughout the period. In **Figure 8** B, the monthly change in asylum

applications is calculated as the difference between the asylum applications at month *t* minus the asylum applications in the previous month (*t-1*). The result in this case is a stationary series, where the values are centred around a constant mean, and the variance is rather stable.

All the features described in this section are not only related to the CUSUM model, but they have implications for early warning methods more generally. We have implemented specific parameters to account for them and provide meaningful and unbiased results (see also Annex 5 for a list of these parameters).

4.2.3 Three levels of warnings based on severity scores

As mentioned in Section 4.1.4, intensity scores or severity levels can enhance the effectiveness of the EWS. A three-tiered scale classifying anomalies as "moderate", "serious" or "critical" provides a straightforward reference to convey the severity of the warning. To operationalise this, we assigned different scores to each CUSUM value: if the CUSUM value is beyond 1σ, we mark this as an early warning with score 1 corresponding to the "moderate" level in the EWS; if it is beyond 2σ the score is 2 corresponding to the "serious" level in the EWS; if it is beyond 3σ (which is also the control limit $h^{\pm}$), the score is 3 corresponding to the "critical" level in the EWS.

The next chapter places the focus on the validation of the effectiveness of the algorithms presented in this section. In order to do so, the focus will be placed on asylum applications by two key countries of origin: Türkiye and Egypt.

5 Validation

The development of the proposed methodology was informed by consultations with policy experts in DG HOME, including through an exercise that required them to label anomalies in selected quantitative and univariate time series (as described in 3.2).

By comparing the results of an anomaly detection model with the labels provided by human experts, we can measure the performance of the algorithms proposed in this report. However, it must be noted that this only works when the human has correctly and completely labelled the anomalies, and it also requires that what a human considers an anomaly is also anomalous in a statistical sense (Hyndman, 2024). For this reason, the exercise was done after the algorithms were presented to the respondents. This allowed us to gather a *ground truth* dataset[36] to test the performance of our models.

Since the idea of anomaly can vary amongst respondents, it can happen that the same observation is marked as anomaly by some respondents but not by others. For this reason, and for the sake of testing the performance of the EW models, **we consider an instance to be "significant" when it has gathered more consensus among the respondents, indicating broader agreement among experts**. We can therefore define a consensus score, which is calculated as the sum of the number of respondents who marked an observation as an early warning (*i.e.* the number of votes). The total number of respondents for this exercise was five. Finally, we assume that instances with no votes are implicitly *not anomalous*.

Using the ground truth and the outcome of the models, we can build a ***confusion matrix***[37] and calculate the *confusion metrics* of the models, which allows us to **evaluate certain performances**. In particular, by looking at the rate of *true positives* and *false positives*—respectively, how often the model identified a real anomaly as such or not—we can measure the *Precision* of an algorithm using the formula *Precision=TP/(TP+FP)*, where *TP* and *FP* are respectively true and false positives (the latter also known as Type I error). The higher the Precision value, the better the algorithm is in limiting the occurrence of false positives, the maximum being 1 and the minimum 0. This is of great interest since false positives can jeopardise the credibility of an EWS by issuing false alarm notifications to the users.

Similarly, we look into the ***Recall* metrics**, **which measure how well the model identifies all actual positive cases**. Its formula is *Recall=TP/(TP+FN)*, where *FN* stands for f*alse negative*, namely the case in which the model has flagged an anomaly as *not anomalous* (also known as Type II error). A higher *Recall* value means a more performant algorithm, a lower value means the opposite, the maximum being 1 and the minimum 0. Contrary to Precision, this metrics focuses on missing positive cases and can be combined with the consensus score, since ideally we want our models not to miss any true anomaly, and in particular those with a high consensus.

In the following sections, we present the consensus plots and the results of the model validation for each use case identified in 3.2 using the parameter configuration as reported in Annex 4 and the results reported in Annex 6. In general, we can see that there are only a few instances where the

[36] Verified and trusted data used for validation purposes.

[37] A confusion matrix is a table that is used to define the performance of a classification algorithm. A confusion matrix visualizes and summarizes the performance of a classification algorithm. (Singh et al., 2021)

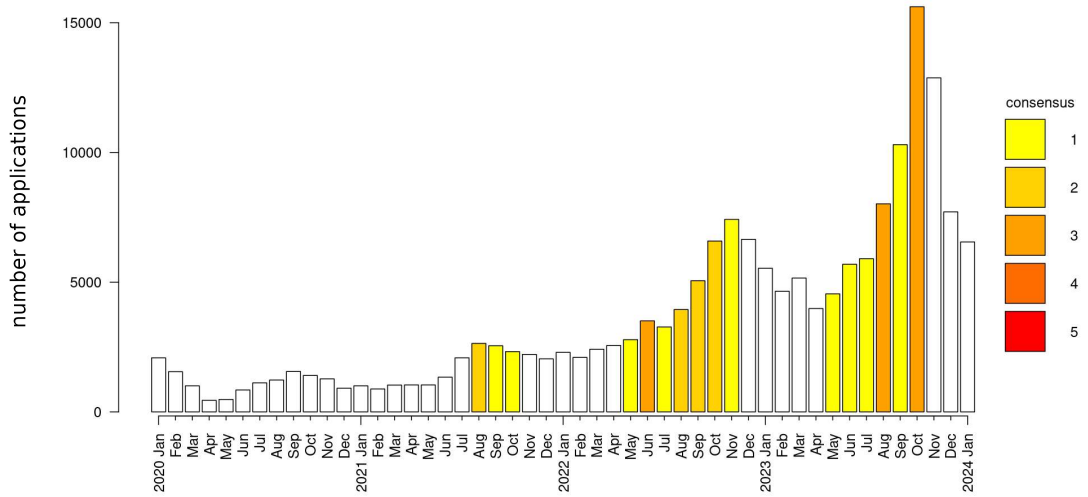
respondents' votes match. This could indicate that each expert has a personal idea of what is an early warning that differs from the others. Even so, there are still observations with high scores (3 or more votes).

Finally, since each EW model assigns a certain score to each anomaly which is different from the consensus score, the error metrics are derived without taking these scores into account but only considering the dichotomy between anomaly and normality. However, in describing the results of the validation, we reintroduce the concept of scores when looking at the anomalies that the model failed to detect, if any (false negative). This is because we want to highlight cases where a possible missed anomaly is also a serious one, which ideally should not happen with a well-performing EWS.

5.1 Asylum applications from Turkish nationals, 2020-2023

Figure 9 shows the number of first-time asylum applications lodged by Turkish nationals since 2020 in the EU, along with the information on the consensus of what is an anomaly for DG HOME experts who participated in the interactive session of the workshop (consensus).

Figure 9. Validation: asylum applications from Turkish nationals 2020-2023



Source: JRC elaboration

The figure shows that the highest score was 3, meaning that 3 respondents out of 5 labelled the same observations as anomalous. These are June 2022, and August and October 2023. No instance received neither 4 nor 5 votes, while the rest of the selected anomalies have low values (between 1 and 2).

Table 5 is the confusion matrix for the CUSUM and multi-criteria models.

Table 5. Confusion matrix for the CUSUM and multi-criteria models on the Asylum applications in the EU by Turkish nationals. P: positive; N: negative; PP: predicted positive; PN: predicted negative; TP: true positive; TN: true negative; FP: false positive; FN: false negative.

		Predicted

		CUSUM		Multi-criteria	
	N+P=49	**Anomaly (PP)** 12	**Not anomaly (PN)** 37	**Anomaly (PP)** 15	**Not anomaly (PN)** 34
Actual	**Anomaly (P)** 16	11 (TP)	5 (FN)	12 (TP)	4 (FN)
	Not anomaly (N) 33	1 (FP)	32 (TN)	3 (FP)	30 (TN)

Source: JRC elaboration

Table 6 shows the performance of the models, displaying the Precision and Recall error metrics.

Table 6. Performance of the models

	CUSUM	**Multi-criteria**
Precision	0.9	0.8
Recall	0.7	0.8

Source: JRC elaboration

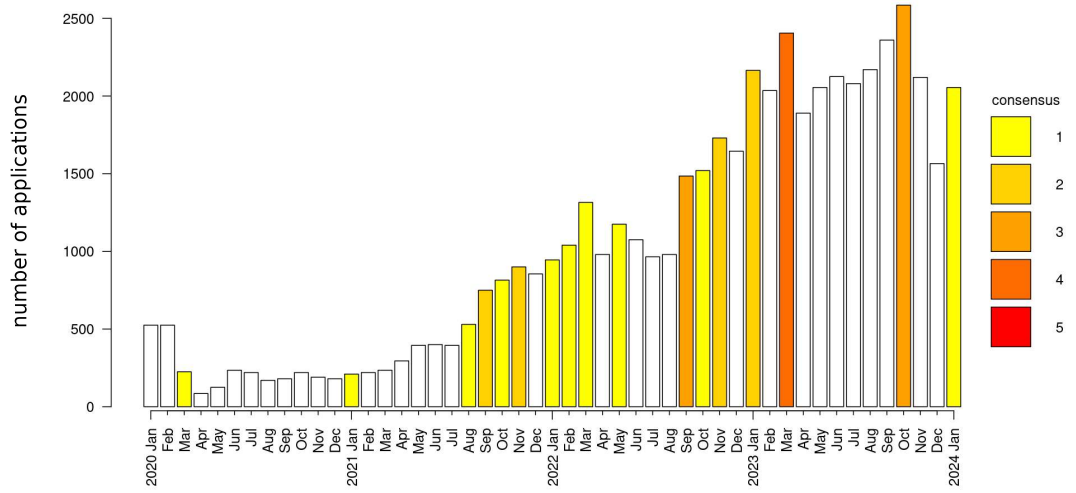
Both models have very high Precision values, with the CUSUM scoring 0.9 and the multi-criteria 0.8 respectively (both close to the maximum score which is 1). This means that **they both trigger very few false alarms.** This is crucial, since an EW model that throws too many false alarms is unreliable and users may end up not trusting its results.

The Recall metrics, namely the probability of catching a real positive case, is also high in both models (0.7 and 0.8 for the CUSUM and the multi-criteria respectively). This means that **the models detected 70-80% of the actual anomalies as identified by HOME experts**. Of the few anomalies not detected (false negatives), nearly all had a consensus score of 1 (the lowest score). The only exception was one anomaly in the CUSUM method, which had a score of 2. **Not only did the models successfully identify most of the actual anomalies, but they also correctly detected the most significant ones, which had high consensus scores**. The few anomalies that were missed were relatively minor, with low consensus scores. Overall, for this specific use case we can see that the **CUSUM tends to give better results in terms of limiting the amount of false alarms, whereas the multi-criteria is better for detecting the real anomalies, although the difference between the two is very subtle.**

5.2 Asylum applications from Egyptian nationals, 2020-2023

Figure 10 shows the number of first-time asylum applications lodged by Egyptian nationals since 2020 in the EU, along with the information on the consensus of what is an anomaly for HOME experts who participated in the interactive session of the workshop (consensus).

Figure 10. Validation: asylum applications from Egyptian nationals 2020-2023



Source: JRC elaboration

The figure shows that the observation with the largest consensus was March 2023, with 4 votes out of 5, followed by September 2022 and October 2023, both with 3 votes each. It also shows that there was no anomaly with unanimous vote (level 5). The rest of the anomalies have low values (between 1 and 2).

Table 7 is the confusion matrix for the CUSUM and multi-criteria models.

Table 7. Confusion matrix for the CUSUM and multi-criteria models on the Asylum applications in the EU by Egyptian nationals. P: positive; N: negative; PP: predicted positive; PN: predicted negative; TP: true positive; TN: true negative; FP: false positive; FN: false negative.

		Predicted			
		CUSUM		**Multi-criteria**	
	N+P=49	**Anomaly (PP)** 12	**Not anomaly (PN)** 37	**Anomaly (PP)** 36	**Not anomaly (PN)** 13
Actual	**Anomaly (P)** 17	11 (TP)	6 (FN)	14 (TP)	3 (FN)
	Not anomaly (N) 32	1 (FP)	31 (TN)	8 (FP)	24 (TN)

Source: JRC elaboration

Table 8 shows the performance of the models, displaying the Precision and Recall error metrics.

Table 8. Performance of the models

	CUSUM	**Multi-criteria**
Precision	0.9	0.6
Recall	0.6	0.8

Source: JRC elaboration

In this use case, the Precision metrics differ to some degree between the two models. On the one hand, the CUSUM model scored 0.9, close to the maximum score of 1. This means that it generated very few false alarms, with only one instance of incorrectly identifying an event as an anomaly, as was the case with the Turkish nationals. As explained previously, minimising false alarms is a key element of an EWS, as excessive false alarms can impact the credibility of the model. In contrast, the Precision for the multi-criteria is rather low instead (0.6), suggesting a rather high rate of false alarms. This might be due to the fact that using multiple criteria can increase the chances of over-detecting anomalies, even though we limited this possibility using the "strict" version of the algorithm (see also Annex 5 for a complete list of its parameters).

The scores for the Recall metrics show the opposite scenario, with the multi-criteria scoring 0.8, while the CUSUM reached 0.6, meaning that 60% of the actual anomalies were detected by the latter. Although the CUSUM model missed 40% of the anomalies (a total of 6 cases), it is notable that all of these undetected anomalies had very low consensus scores, with 5 having the minimum score of 1 and 1 having a score of 2. Therefore, all the important anomalies (those with a high consensus score) were still correctly detected by the model.

Overall, **for this specific use case we can conclude that using the CUSUM model limits the rate of false alarms at the expense of missing real anomalies (although these were of relatively minor importance). On the other hand, the multi-criteria model is more successful in detecting a higher proportion of real anomalies, but this comes at the expense of generating more false alarms. Ultimately, the choice between these two models depends on the users' priorities.**

6 Conclusions

This report proposed an early warning methodology stemming from an empirical analysis of anomaly detection methods for mixed migration flows found in both the policy context and the academic literature.

The proposed approach features two algorithms:

- The **multi-criteria algorithm**, which allows to capture different patterns in the data and to contextualise data points by taking into consideration different time-spans. This approach allows for a nuanced interpretation of anomalies, accounting for both short-term and long-term trends, and can be tailored to specific contexts.
- The **CUSUM method**, a statistical process control technique capable of identifying early changes in data which had been previously tested in migration studies.

Both methods are tailored to assign an overall "intensity" score or warning level to each observation, based on a set of thresholds. This allows to better capture the severity of changes detected for evolving situations across different contexts.

The performance of these algorithms was evaluated through two case studies, selected according to their potential relevance to policy experts and discussed with European Commission experts: asylum applications from Turkish and Egyptian nationals. The validation was done by comparing the results of the models with the labels provided by human experts, who were asked to identify anomalies in the data.

The experts' labels were used to create a ground truth dataset, which was then used to evaluate the performance of the models. The performance was measured using two metrics: Precision and Recall. Precision measures the proportion of true positives (correctly identified anomalies) among all predicted anomalies, while Recall measures the proportion of true positives among all actual anomalies.

The results show that:

- For asylum applications from Turkish nationals, both models performed well, with high Precision and Recall values. The CUSUM model had a slightly higher Precision value, indicating that it generated fewer false alarms.
- For asylum applications from Egyptian nationals, the CUSUM model had a high Precision value, but a lower Recall value, indicating that it missed some real anomalies. The multi-criteria model had a lower Precision value, but a higher Recall value, indicating that it detected more anomalies, but also generated more “false alarms”.

Overall, the results suggest that, for the specific use cases presented in this report and possibly other indicators with similar distributions, **the choice between the two models depends on the users' priorities**. If minimising false alarms is more important, the CUSUM model may be preferred. If detecting a higher proportion of real anomalies is more important, the multi-criteria model may be preferred.

To further develop and refine these methodologies, it is essential to **test them on a broader range of indicators**, including those related to mobility (outside or inside the EU) or those concerning drivers of migration or triggering factors.

These methods can provide the basis for the development of an Early Warning System that follows the key principles outlined in Section 3:

- **Simplicity**: the multi-criteria algorithm and the CUSUM method are transparent and their results can be interpreted in a straight-forward manner by policy analysts, allowing them to analyse their outputs to inform decision-making.
- **Timeliness**: the algorithms' ability to detect anomalies early on, allowing for close to real-time monitoring, combined with the use of intensity scores and warning levels, can allow for timely interventions.
- **Relevance:** the methodology proposed can be applied to indicators that are essential for monitoring migration flows, ensuring relevance to the policy context, even if this report has placed the focus on asylum applications. This also relates to the next key criterion:
- **Flexibility:** the proposed methodology offers flexibility in several aspects, including the option to choose an algorithm according to specific priorities, to assign intensity scores and warning levels which allow to prioritise different types of anomalies, or to adjust thresholds.
- **Accuracy**: both methods proved effective in detecting significant anomalies, and can satisfy different users' needs: while the CUSUM model limits the rate of false alarms at the expense of missing (few and relatively unimportant) real anomalies, the multi-criteria model detects a higher proportion of real anomalies but generates more false alarms.

In addition to the above, users of an EWS based on these algorithms should be encouraged to integrate information on quantitative anomalies with qualitative sources. The latter can significantly enhance the effectiveness of early warning algorithms by providing valuable contextual background information.

Moreover, these inputs can identify emerging trends and developments that warrant attention, even if quantitative indicators do not trigger alerts. Insights and judgements of experienced analysts are crucial for anticipatory decision-making in migration policy, as they allow to leverage both specialised knowledge and different perspectives on potentially interlinked developments. Techniques that allow to detect emerging issues and first signals of change, such as horizon scanning, could be a suitable solution for tapping into this expertise.

References

Studies

- Atienza, O. O., Tang, L. C., & Ang, B. W. (2002). A CUSUM Scheme for Autocorrelated Observations. Journal of Quality Technology, 34(2), 187–199. https://doi.org/10.1080/00224065.2002.11980145
- Bandara, K., Hyndman, R. J., & Bergmeir, C. (2025). MSTL: A seasonal-trend decomposition algorithm for time series with multiple seasonal patterns. International Journal of Operational Research.
- Barker, E., & Bijak, J. (2023). Could We Have Seen it Coming? Towards an Early Warning System for Asylum Applications in the EU (SSRN Scholarly Paper 4605897). https://doi.org/10.2139/ssrn.4605897
- Bijak, J., Hilton, J. & Forster, J. (2017). Quantitative assessment of asylum-related migration: a survey of methodology, Publications Office. https://data.europa.eu/doi/10.2847/642161
- Carammia, M., Iacus, S. M., & Wilkin, T. (2022). Forecasting asylum-related migration flows with machine learning and data at scale. Scientific Reports, 12(1), 1457. https://doi.org/10.1038/s41598-022-05241-8
- Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. ACM Comput. Surv., 41(3), 15:1-15:58. https://doi.org/10.1145/1541880.1541882
- Cleveland, R. B., Cleveland, W. S., McRae, J. E. & Terpenning, I. J. (1990). STL: A seasonal-trend decomposition procedure based on loess. Journal of Official Statistics, 6(1), 3–33. http://bit.ly/stl1990
- Constantinos Melachrinos, Marcello Carammia, & Teddy Wilkin. (2023). An innovative framework for analysing asylum-related migration. In Harnessing data innovation for migration policy: A handbook for practitioners. https://euaa.europa.eu/publications/innovative-framework-analysing-asylum-related-migration
- Crippa, A., d'Agostino, G., Dunne, J. P., & Pieroni, L. (2024). Conflict as a cause of migration. Oxford Economic Papers, 1–23. https://doi.org/10.1093/oep/gpae037
- European Union Agency for Asylum, Bijak, J., Hilton, J., & Forster, J. J. (2017). Quantitative assessment of asylum-related migration: A survey of methodology. Publications Office of the European Union. https://data.europa.eu/doi/10.2847/642161
- Gordenker, L. (1984). Early Warning of Forced Migration. In Defense of the Alien, 7, 68–77.
- Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J. & Stahel, W.A. (1986). Robust Statistics: The Approach Based on Influence Functions. Wiley Series in Probability and Statistics. 10.1002/9781118186435
- Hoffmann, M., Rosenbaum, M. & Yoshida, N. Estimation of the lead-lag parameter from non-synchronous data. Bernoulli 19, 426–461 (2013).

- Human Rights Watch (2024), World Report 2024: Türkiye. https://www.hrw.org/world-report/2024/country-chapters/turkey
- Hyndman, R.J. & Athanasopoulos, G. (2021) Forecasting: principles and practice, 3rd edition, OTexts: Melbourne, Australia. OTexts.com/fpp3. Accessed on 20th November 2024.
- Hyndman, R.J. (2023) fpp3: Data for "Forecasting: Principles and Practice" (3rd Edition). R package version 0.5. https://CRAN.R-project.org/package=fpp3
- Hyndman, R.J. (2024). That's Weird: Anomaly Detection Using R. OTexts, Melbourne, Australia. In preparation, https://OTexts.com/weird/
- Lu, C. W. & Reynolds, M. R. (2001). CUSUM Charts for Monitoring an Autocorrelated Process. Journal of Quality Technology, 33(3), 316–334. https://doi.org/10.1080/00224065.2001.11980082
- Macis, L., Tagliapietra, M., Meo, R., & Pisano, P. (2024). Breaking the trend: Anomaly detection models for early warning of socio-political unrest. Technological Forecasting and Social Change, 206, 123495. https://doi.org/10.1016/j.techfore.2024.123495
- Melachrinos, C., Carammia, M. & Wilkin, T. (2023). An Innovative Framework for Analysing Asylum-Related Migration. International Organization for Migration (IOM), *Harnessing Data Innovation for Migration Policy. A Handbook for Practitioners*, Geneva: International Organization for Migration, pp. 54-65.Marcucci, S., & Verhulst, S. (n.d.). Towards a Taxonomy of Anticipatory Methods: Integrating Traditional and Innovative methods for Migration Policy. Big Data for Migration Alliance. Retrieved July 30, 2024, from https://www.data4migration.org//articles/blog-series-part-3-towards-a-taxonomy-of-anticipatory-methods-integrating-traditional-and-innovative-methods-for-migration-policy/
- Murphy, J. J. Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications (New York Institute of Finance, 1999)
- Napierała, J., Hilton, J., Forster, J. J., Carammia, M., & Bijak, J. (2022). Toward an Early Warning System for Monitoring Asylum-Related Migration *Flows in Europe.* International Migration Review, 56(1), 33–62. https://doi.org/10.1177/01979183211035736
- Page, E. S. (1954). *Continuous Inspection Schemes*. Biometrika, 41(1/2), 100–115. https://doi.org/10.2307/2333009
- Runger, G. C., Wlllemain, T. R., & Prabhu, S. (1995). Average run lengths for CUSUM control charts applied to residuals. Communications in Statistics - Theory and Methods, 24(1), 273–282. https://doi.org/10.1080/03610929508831487
- Schölkopf, B., Williamson, R. C., Smola, A., Shawe-Taylor, J., & Platt, J. (1999). Support Vector Method for Novelty Detection. Advances in Neural Information Processing Systems, 12. https://papers.nips.cc/paper_files/paper/1999/hash/8725fb777f25776ffa9076e44fcfd776-Abstract.html
- Science Direct. (n.d.). Early Warning System—An overview. Retrieved June 11, 2024, from https://www.sciencedirect.com/topics/social-sciences/early-warning-system

- Schvitz, G., Weiler, S. M., Van, D. M.-S., & Bardelli, G. (2023). *Science4Peace: A quantitative toolkit for conflict early warning and crisis management.* JRC Publications Repository. https://publications.jrc.ec.europa.eu/repository/handle/JRC135125
- Sfyridis, A., Cheng, T., & Vespe, M. (2017). Detecting vessels carrying migrants using machine learning. https://publications.jrc.ec.europa.eu/repository/handle/JRC106832
- Sharifi, A., Pathak, M., Joshi, C., & He, B.-J. (2021). A systematic review of the health co-benefits of urban climate change adaptation. Sustainable Cities and Society, 74, 103190. https://doi.org/10.1016/j.scs.2021.103190
- Singh, P., Singh, N., Singh, K. K., & Singh, A. (2021). Diagnosing of disease using machine learning. In Machine learning and the internet of medical things in healthcare (pp. 89-111). Academic Press. https://doi.org/10.1016/B978-0-12-821229-5.00003-3

Policy and Technical Documents

- Agrhymet. (n.d.). Cadre Harmonisé. Retrieved July 23, 2024, from http://cadreharmonise.org/
- Aghrymet. (n.d.). Cadre Harmonisé. Manuel Version 3.0. Analyse et identification des zones à risque et des populations en insécurité alimentaire et nutritionnelle. Retrieved July 23, 2024, from https://www.cadreharmonise.org/web/content/6811?unique=961409fc03685a9362c9161ec92eb2f6c0a0bbee&download=true
- Council of the European Union. (2016). The EU Integrated Political Crisis Response – IPCR– Arrangements. In Brief. https://www.consilium.europa.eu/media/29699/web_ipcr.pdf
- European Border and Coast Guard Agency (Frontex). (n.d.). Risk analysis for 2023/2024. Retrieved December 14, 2024, from https://www.frontex.europa.eu/assets/Publications/General/ARA_2023.pdf
- European Commission. (n.d.-a). Horizon Scanning. Retrieved July 30, 2024, from https://knowledge4policy.ec.europa.eu/foresight/topic/horizon-scanning_en
- European Commission. (n.d.-b). Strategic foresight. Retrieved September 25, 2024, from https://commission.europa.eu/strategy-and-policy/strategic-foresight_en
- European Commission, Directorate-General for International Partnerships (INTPA). (n.d.-a). INTPA Strategic Foresight. Retrieved September 25, 2024, from https://capacity4dev.europa.eu/resources/INTPA-Foresight_en
- European Commission, Directorate-General for International Partnerships (INTPA). (n.d.-b). Youth4Foresight Toolkit | Capacity4dev. Retrieved September 25, 2024, from https://capacity4dev.europa.eu/resources/INTPA-Foresight/info/youth4foresight-toolkit_en
- European External Action Service (EEAS). (2022). Early Warning System. Fact sheet. https://www.eeas.europa.eu/sites/default/files/documents/Factsheet%20-%20EWS.pdf

- European Migration Network (EMN). (n.d.). Mixed Migration Flow. Retrieved December 16, 2024, from https://home-affairs.ec.europa.eu/networks/european-migration-network-emn/emn-asylum-and-migration-glossary/glossary/mixed-migration-flow_en
- European Parliament, Committee on Foreign Affairs. (2022). REPORT on the 2021 Commission Report on Turkey | A9-0149/2022. https://www.europarl.europa.eu/doceo/document/A-9-2022-0149_EN.html
- European Strategy and Policy Analysis System (ESPAS). (n.d.). Horizon Scanning. Retrieved September 25, 2024, from https://espas.eu/horizon.html
- European Union Agency for Asylum (EUAA) (2015). Asylum Applicants from the Western Balkans: comparative analysis of trends, push-pull factors and responses – Update. https://euaa.europa.eu/sites/default/files/public/Asylum-Applicants-from-the-Western-Balkans_Update_r.pdf.
- European Union Agency for Asylum (EUAA). (2023). An innovative framework for analysing asylum-related migration, May 2023.https://euaa.europa.eu/sites/default/files/publications/2023-05/An_innovative_framework_analysing_asylum-related_migration.pdf
- European Union Agency for Asylum (EUAA). (2024a). Asylum Report 2024. Retrieved December 14, 2024, from https://euaa.europa.eu/publications/asylum-report-2024
- European Union Agency for Asylum (EUAA). (2024b). Data analysis of decisions on asylum applications and recognition rates at second or higher instances in 2023. https://euaa.europa.eu/publications/data-analysis-decisions-asylum-applications-and-recognition-rates-second-or-higher-instances-2023
- European Union Agency for Asylum (EUAA). (2024c). Sudan: Country focus : security situation in selected areas and selected profiles affected by the conflict : April 2024. Publications Office. https://data.europa.eu/doi/10.2847/759057
- European Union Agency for Asylum (EUAA). (2022). Migration Drivers Report: Egypt as a Country of Origin. https://euaa.europa.eu/publications/migration-drivers-report-egypt-country-origin
- European Union Agency for Asylum (EUAA). (n.d.-b). Data Analysis and Research. Retrieved December 11, 2024, from https://euaa.europa.eu/asylum-knowledge/data-analysis-and-research
- European Union Agency for Asylum (EUAA). (n.d.-c). EASO Asylum Report 2020. Retrieved December 14, 2024, from https://euaa.europa.eu/asylum-report-2020/43-push-factors-third-countries
- European Union Agency for Asylum (EUAA). (n.d.-d). Latest Asylum Trends. Retrieved December 14, 2024, from https://euaa.europa.eu/latest-asylum-trends-asylum
- Food and Agriculture Organization of the United Nations (FAO). (n.d.). *Integrated Food Security Phase Classification—IPC*. Retrieved June 17, 2024, from https://www.fao.org/policy-support/mechanisms/mechanisms-details/en/c/428641/

- Integrated Food Security Phase Classification (IPC). (n.d.-a). IPC Overview and Classification System. Retrieved June 17, 2024, from https://www.ipcinfo.org/ipcinfo-website/ipc-overview-and-classification-system/en/
- Integrated Food Security Phase Classification (IPC). (n.d.-b). Population Tracking Tool. Retrieved June 17, 2024, from https://www.ipcinfo.org/ipc-country-analysis/population-tracking-tool/en/
- Integrated Food Security Phase Classification (IPC). (2022). Understanding the IPC Scales. https://reliefweb.int/report/world/understanding-ipc-scales
- Integrated Food Security Phase Classification (IPC). (2024). Food Security Cluster Handbook. https://handbook.fscluster.org/docs/671-integrated-phase-classification
- International Organisation for Migrations (IOM). (2023). Report of the 33rd session of the Standing Committee on Programmes and Finance. Delivering impact through data for action, insight and foresight. https://governingbodies.iom.int/system/files/en/scpf/33rd/s-33-8-delivering-impact-through-data.pdf
- United Nations Development Programme (UNDP). (2022). Foresight Playbook. Overview of Foresight Tools. https://www.undp.org/sites/g/files/zskgke326/files/2022-07/UNDP-RBAP-Foresight-Playbook-Appendix-2022_0.pdf
- United Nations Environment Programme (UNEP), (2023). Strategic Foresight @ UNEP. https://www.unep.org/ocs/strategicforesightatunep
- United Nations Human Rights Council: General Debate Item 4: Human Rights Situations that require the Council's attention, Joint Statement on Egypt. 12 March 2021. https://reliefweb.int/report/egypt/46th-session-united-nations-human-rights-council-general-debate-item-4-human-rights
- United Nations High Commissioner for Refugees (UNHCR). (2023a). Digital Transformation Strategy 2022-2026. https://www.unhcr.org/digitalstrategy/wp-content/uploads/sites/161/2023/07/Digital-Transformation-Strategy-2022-2026-UNHCR-Web.pdf
- United Nations High Commissioner for Refugees (UNHCR). (2023b). Emergency preparedness and response. https://www.unhcr.org/sites/default/files/legacy-pdf/640889b14.pdf
- United Nations Office for Disaster Risk Reduction (UNDRR). (2007). Early warning system. https://www.undrr.org/terminology/early-warning-system
- United Nations Office for the Coordination of Humanitarian Affairs (OCHA), Centre for Humanitarian Data. (n.d.-a). HDX Signals Map. Retrieved July 26, 2024, from https://data.humdata.org/visualization/signals/
- United Nations Office for the Coordination of Humanitarian Affairs (OCHA), Centre for Humanitarian Data. (n.d.-b). Intro to HDX Signals .pdf. Google Docs. Retrieved July 26, 2024, from https://drive.google.com/file/d/1tcnhGDqlF_1tq4Q4SFkcXGLXpMRhM4bv/view?usp=embed_facebook

- United Nations Office for the Coordination of Humanitarian Affairs (OCHA), Centre for Humanitarian Data. (2024a). Armed conflict | HDX Signals. https://un-ocha-centre-for-humanitarian.gitbook.io/hdx-signals/datasets/armed-conflict
- United Nations Office for the Coordination of Humanitarian Affairs (OCHA), Centre for Humanitarian Data. (2024b). Datasets | HDX Signals. https://un-ocha-centre-for-humanitarian.gitbook.io/hdx-signals/datasets
- United Nations Office for the Coordination of Humanitarian Affairs (OCHA), Centre for Humanitarian Data. (2024c). Agricultural hotspots | HDX Signals. https://un-ocha-centre-for-humanitarian.gitbook.io/hdx-signals/datasets/agricultural-hotspots
- United Nations Office for the Coordination of Humanitarian Affairs (OCHA), Centre for Humanitarian Data. (2024d). Prompts | HDX Signals. https://un-ocha-centre-for-humanitarian.gitbook.io/hdx-signals/prompts
- United Nations Office for the Coordination of Humanitarian Affairs (OCHA), Centre for Humanitarian Data. (2024e). Internal displacements | HDX Signals. https://un-ocha-centre-for-humanitarian.gitbook.io/hdx-signals/datasets/internal-displacement

Newspaper Articles

- Focus - Inflation in Turkey: Soaring prices, housing crisis and student protests. (2021, October 18). France 24. https://www.france24.com/en/tv-shows/focus/20211018-inflation-in-turkey-soaring-prices-housing-crisis-and-student-protests
- Protests in Turkey's cities after lira plunges to new historic low. (2021, November 25). Euronews. https://www.euronews.com/next/2021/11/24/turkish-lira-plunges-to-a-historic-low-and-protests-erupt-after-erdogan-defends-rates-cuts
- Turkish police tear-gas women protesting over violence. (2021, November 25). AP News. https://apnews.com/article/europe-middle-east-violence-turkey-recep-tayyip-erdogan-37dc9465b1f867b00465f8006a74d733
- "Why the Turkish currency is in free fall" (2020, August 17). DW. https://www.dw.com/en/erdogans-credit-binge-fuels-lira-depreciation/a-54524078

List of abbreviations and definitions

Abbreviations	Definitions
ACLED	Armed Conflict Location & Event Data Project
AFI	Acute food insecurity
AMN	Acute malnutrition
ASAP	Anomaly Hotspots of Agricultural Production
CET	Conflict event trends
CFI	Chronic food insecurity
CH	Cadre Harmonisé
CUSUM	Cumulated sum control chart
DCRM	Dynamic Conflict Risk Model
DG HOME	Directorate-General for Migration and Home Affairs
EC	European Commission
EEAS	European External Action Service
EMLO	European Migration Liaison Officer
EMN	European Migration Network
EPS	EUAA Early Warning and Preparedness System
EU	European Union
Eu-LISA	European Union Agency for the Operational Management of Large-Scale IT Systems
EUAA	European Union Agency for Asylum
Europol	European Union Agency for Law Enforcement Cooperation
Eurostat	Statistical Office of the European Union

Abbreviations	Definitions
EWS	Early Warning System
FN	False negative
FP	False positive
FRA	European Union Agency for Fundamental Rights
Frontex	European Border and Coast Guard Agency
GCA	Global Crisis Atlas
GCRI	Global Conflict Risk Index
GDELT	Global Database of Events, Language, and Tone
HDX	Humanitarian Data Exchange
IBC	Irregular border-crossing
IDMC	Internal Displacement Monitoring Centre
IOM	International Organization for Migration
IPC	Integrated Food Security Phase Classification
JHA	Justice and Home Affairs
LOESS	Locally estimated scatterplot smoothing
MAD	Median absolute deviation
NSC	Non-state conflict
N	Negative
OCHA	United Nations Office for the Coordination of Humanitarian Affairs
OSV	One-sided violence
P	Positive

Abbreviations	Definitions
PN	Predictive negative
PP	Predictive positive
SBC	State-based conflict
SPC	Statistical control process
STL	Seasonal and trend decomposition using Loess
TN	True negative
TP	True positive
UCDP	Uppsala Conflict Data Program
UNDRR	United Nations Office for Disaster Risk Reduction
UNCHR	United Nations High Commissioner for Refugees
WFP	World Food Programme

List of figures

List of tables

Annexes

Annex 1. IPC Scales[38]

Table 9. IPC scales

Scale	Severity phase				
Acute Food Insecurity (AFI)	Minimal/None	Stressed	Crisis	Emergency	Catastrophe
	Essential needs met without atypical and unsustainable methods to access food and income.	Minimally adequate food consumption – inability to afford essential non-food expenditures	Significant food gaps, leading to malnutrition, or minimum food needs met by depleting assets or using crisis-coping strategies.	Large consumption gaps, leading to very high acute malnutrition and excess mortality; or mitigation of food consumption gaps through emergency strategies and asset liquidation.	Extreme lack of food or other basic needs. Evident levels of starvation, destitution and extremely critical acute malnutrition.
Chronic Food Insecurity (CFI)	No/Minimal Chronic Food Insecurity	Mild Chronic Food Insecurity	Moderate Chronic Food Insecurity	Severe Chronic Food Insecurity	
	Access and consumption of a diet of acceptable quantity and quality, sustainable livelihoods, resilience to shocks and low risk of stunted children.	Access to enough food, but not always of acceptable quality. Livelihoods are borderline sustainable and limited resilience to shocks, but risk of stunting in children is low.	Mild deficits in food quantity for 2-4 months each year, consistent lack of food quality. Livelihoods are marginally sustainable, with limited resilience to shock. Moderate risk of stunting in children.	Seasonal deficits in food quantity for more than 4 months per year, consistent lack of food quality. Very marginal livelihoods which are not resilient. Likeliness of severe stunting in children.	
Acute Malnutrition (AMN)	Acceptable	Alert	Serious	Critical	Extremely critical
	Less than 5% of children acutely malnourished.	5-9.9% of children acutely malnourished.	10.14-14.9% children acutely malnourished.	15-29.9% children acutely malnourished.	30% or more children acutely malnourished.

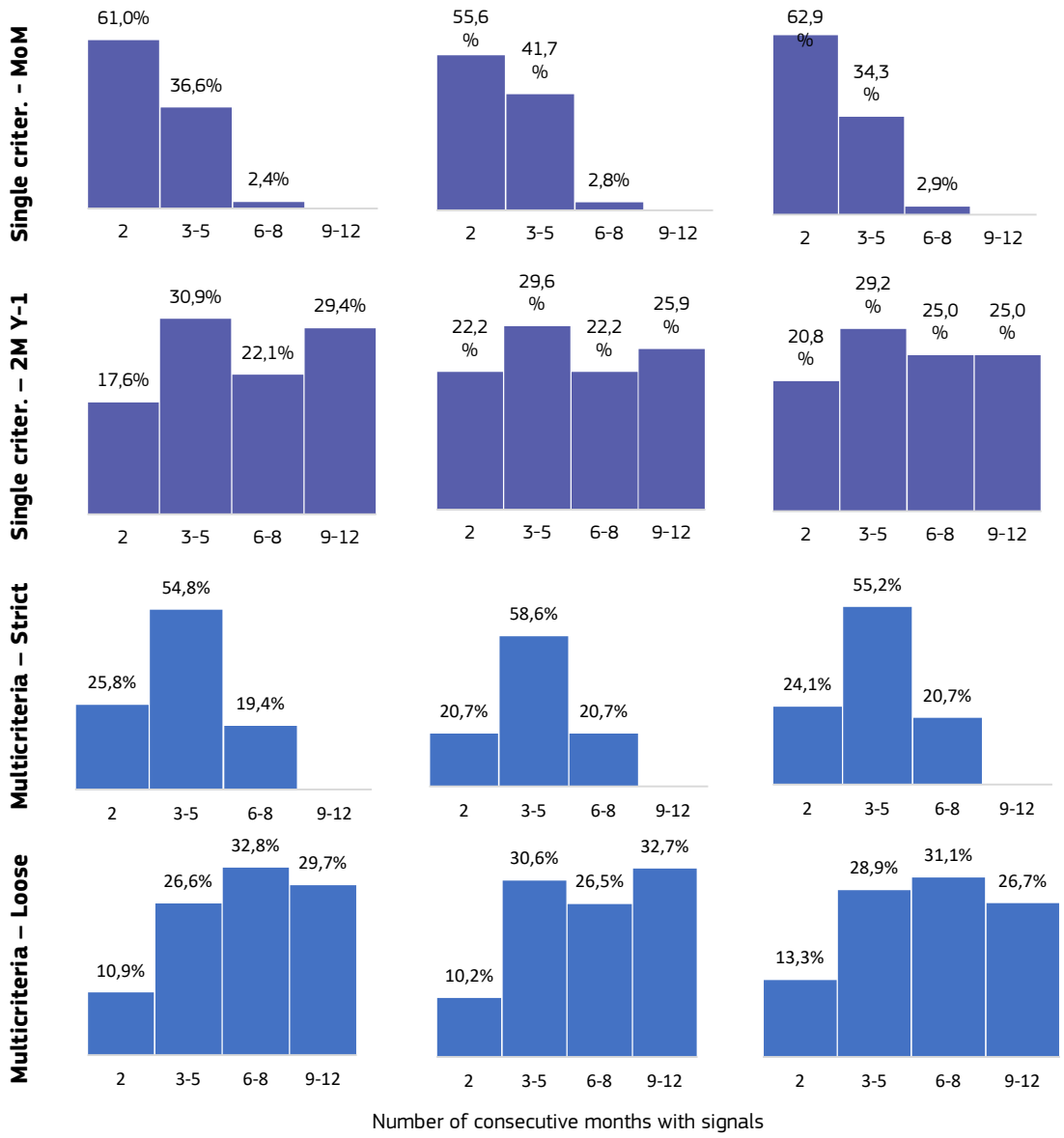
Source: JRC elaboration based on IPC (2022).

Annex 2. Nationality distribution

Figure 11 Distribution of nationalities by number of consecutive months with signals – different algorithms

All months **10 IBCs min** **25 IBCs min**

[38] Integrated Food Security Phase Classification (IPC). (2022, February 23). Understanding the IPC Scales. https://reliefweb.int/report/world/understanding-ipc-scales



Annex 3. Number of months signalled as anomalies

Table 10 Number of months signalled as anomaly for IBCs in 2023 under the four algorithms for detected nationalities

	Multicriteria Strict	Multicriteria Loose	Single crit. MoM	Single crit. 2M Y-1	All four
Syria	2	9	4	5	Yes
Senegal	6	11	7	11	Yes
Tunisia		6	6	2	No
Ivory Coast	5	8	4	9	Yes
Afghanistan		1	4		No
Guinea	6	10	3	10	Yes

Mali	8	12	6	12	Yes
Morocco	1	6	4	6	Yes
Turkey	2	4	4	4	Yes
Pakistan	1	5	5	4	Yes
Sudan	5	9	5	9	Yes
Gambia	7	11	5	12	Yes
Bangladesh	1	7	4	3	Yes
Burkina Faso	7	11	5	12	Yes
Egypt		3	3	4	No
Palestine	3	9	6	7	Yes
Algeria		4	5	2	No
Cameroon	4	8	4	9	Yes
Eritrea	4	10	5	11	Yes
Iraq	1	6	4	4	Yes
Ukraine		6	4	2	No
Somalia	3	4	5	4	Yes
Sierra Leone	5	9	5	10	Yes
Benin	7	12	5	12	Yes
Nigeria	2	4	3	3	Yes
Democratic Republic of the Congo	1	4	5	2	Yes
Iran		2	4	1	No
Ethiopia	5	9	6	8	Yes
Cuba	2	6	4	6	Yes
Mauritania	4	10	6	7	Yes
Yemen	3	11	4	8	Yes
Ghana	5	10	5	8	Yes
Republic of the Congo	2	10	4	9	Yes
Comoros	4	11	6	6	Yes
Liberia	5	8	4	10	Yes
Albania		1	4	1	No
China	6	12	4	11	Yes
Guinea-Bissau	6	10	5	10	Yes
South Sudan	4	10	4	11	Yes
Nepal	3	12	4	10	Yes
India			4	1	No
Libya			4		No
Togo	5	8	6	11	Yes
Sri Lanka	3	8	6	8	Yes
Burundi		5	5	3	No
Chad	4	10	4	9	Yes
Mongolia	1	7	4	5	Yes

Niger	3	7	4	9	Yes
Russia	2	6	4	6	Yes
Kosovo*		1	3	1	No
Lebanon	1	3	3	2	Yes
Kuwait	2	8	2	7	Yes
Haiti	1	6	2	6	Yes
Djibouti	1	5	3	5	Yes
Belarus			4		No
Central African Republic	2	7	3	9	Yes
Angola	3	6	2	7	Yes
Rwanda	1	6	3	4	Yes
Serbia	1	2	3	4	Yes
Moldova	1	3	3	1	Yes
Azerbaijan	3	10	3	10	Yes
Uganda		3	2	3	No
Bosnia and Herzegovina	1	5	3	3	Yes
Kenya		5	2		No
Turkmenistan	1	6	3	6	Yes
Gabon		8	2	2	No
Uzbekistan		3	2	3	No
North Macedonia		4	1	3	No
Tajikistan		1	1	2	No
Montenegro		2	1		No
Jordan	2	8	1	6	Yes
Armenia	1	2		2	No
Kazakhstan	1	3		3	No
Bolivia	1	4		3	No
Kyrgyzstan	1	4		2	No
South Africa		1		1	No
United Kingdom	1	2		2	No
Georgia		2		3	No
Zimbabwe	1	5			No
Equatorial Guinea		1			No
Ecuador		2			No
North Korea		1			No
Dominican Republic		4			No
Tanzania		2			No
Mozambique		2			No
Malawi		2			No
Myanmar		3			No
Saudi Arabia		2			No

Colombia		1			No
Papua New Guinea		1			No
South Korea		1			No
Qatar		1			No
Cape Verde		1			No
Singapore		1			No
Japan		1			No
Thailand		1			No
Brazil		1			No
Argentina		1			No
Macao		1			No
Venezuela		1			No
Tonga		1			No
Jamaica		1			No
Mauritius		1			No
Namibia		1			No

Note: For nationalities with extremely low historical IBCs, some months can be flagged as anomalies just because figures are not null for a short period of time

Annex 4. Model parameters for the validation exercise

Table 11 shows the parameters used for the CUSUM model used to calculate the performance in the validation exercise.[39]

Table 11. Validation parameters in the CUSUM model

parameter	**value**
seasonal adjustment	false
residual	false
stationary	true
robust (σ = MAD)	true
window	25
reference value (*K*)	σ/2
control limits ($h^{\pm}$)	σ*3

Table 12 shows the parameters used for the multi-criteria model used to calculate the performance in the validation exercise ("strict" version).

Table 12. Criterion and thresholds of the multi-criteria algorithm

criterion \ score	**1**	**2**	**3**

[39] All time series considered in this report do not have any missing data. It is possible to deal with missing data in the time series, but this is not discussed here as it is out of the scope of this work.

% change $_{t-12}$	25%	50%	75%
percentile $_{t-12,\ t-24,\ t-36}$	70th	80th	90th
rolling sd $_{t-1\ to\ t-12}$	1	1.5	2

Note: for each threshold the condition is strictly above, not equal or above

Annex 5. Performance comparison: multi-criteria (strict and loose) vs. single-criteria

This section presents the results on total IBCs by nationality in 2023 using the two versions of the multi-criteria algorithm (i.e., "loose" and "strict") and two algorithms relying on a single indicator. Following the general approach in the policy context, the two single-criterion algorithms tested are:

- Month on month percentage change[40]
- Two-months cumulated year-on-year percentage change[41]

In both cases, the single-criterion algorithm launches a warning if the change is above 25%. The main difference lies in the reference period used: the month on month looks at very recent developments, while the two-month cumulated change contextualises current figures with respect to past trends (i.e., previous year).

Number of anomalies and nationalities detected

With a similar number of nationalities identified, the multi-criteria loose algorithm detects on average more anomalies when a situation is evolving and worthy of attention than single-criterion algorithms. By posing more restrictions, instead, the multi-criteria strict detects on average fewer anomalies, which can help understand when a new detected trend under the loose version is intensifying.

The first dimension to compare the four algorithms is the number of months signalled as anomalies and the number of nationalities detected (Figure 13). When looking at all observations, the multi-criteria algorithm applying the "loose" condition reports a considerably higher number of anomalies. When months with very few IBCs per nationality are excluded (e.g., less than 10 IBCs), however, differences between algorithms became clearer.

While the number of nationalities detected becomes quite similar across the four algorithms,[42] the number of months signalled as anomalies still differs considerably, despite a significant drop for both the "loose" multi-criteria and the single criterion two-month year-on-year change.

Table 13 Number of signals and nationalities detected across algorithms

	All months		**10 IBCs min**		**25 IBCs min**	
	Signals	*Nationalities*	*Signals*	*Nationalities*	*Signals*	*Nationalities*
Single criterion MoM	277	71	231	56	201	51
Single criterion 2M Y-1	447	76	328	58	282	51
Multi-criteria Strict	174	58	153	45	139	40

[40] $(IBC_{n,m} - IBC_{n,m-1}) / IBC_{n,m-1}$; where n=nationality and m=month of crossing

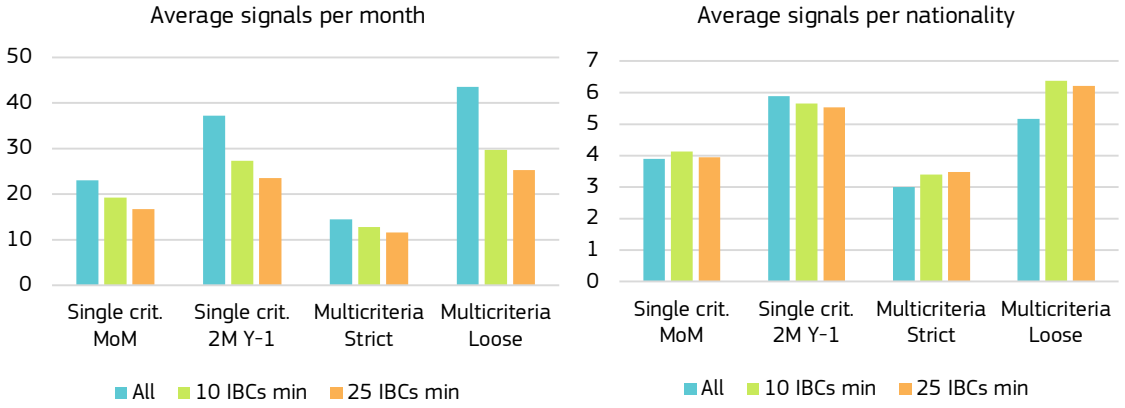
[41] $[(IBC_{n,m} + IBC_{n,m-1}) - (IBC_{n,m-12} + IBC_{n,m-13})] / (IBC_{n,m-12} + IBC_{n,m-13})$; where n=nationality and m=month of crossing

[42] The number of nationalities detected does not decrease significantly when the threshold of IBCs is raised from 10 to 25, thus pointing to stability of results at low level of IBCs.

Multi-criteria Loose	522	101	357	56	304	49

With a **similar number of nationalities** covered, the multi-criteria algorithms tend to **send more signals**. This means that, on average, more months are detected when a situation is evolving and it is worthy of attention. Moreover, **the multi-criteria algorithms do not require significantly more signals per month to detect anomalies across nationalities** (**Figure 12**). For instance, the "loose" multi-criteria reports 6.4 signals per nationality with 29.75 signals per month, compared to 5.6 per nationality and 27.3 per month of the two-month year-on-year change.[43]

Figure 12 Average number of signals per month and per nationality across algorithms



Distribution of anomalies

The multi-criteria algorithm in the "loose" version tends to perform better at detecting and following emerging trends than single-criterion algorithms, especially those using the month-over-month criterion.

When focusing on nationalities with consecutive months detected, both versions of the multi-criteria algorithm show a more normal distribution than single-criterion algorithms. 70% of nationalities with consecutive anomalies lie between 3 and 8 months under the "strict" version, compared, on the one hand, to an even distribution between 2 and 12 months under the two-months year-on-year change algorithm (i.e., not clearly differentiating) and, on the other hand, a concentration in the 2 months group under the month-over-month algorithm (i.e., not following trend).

While averages can help evaluate algorithms' efficiency in identifying anomalies, **the distribution of anomalies detected is relevant to understand what kind of situations are captured by the algorithms**, and how early emerging situations are identified and for how long.

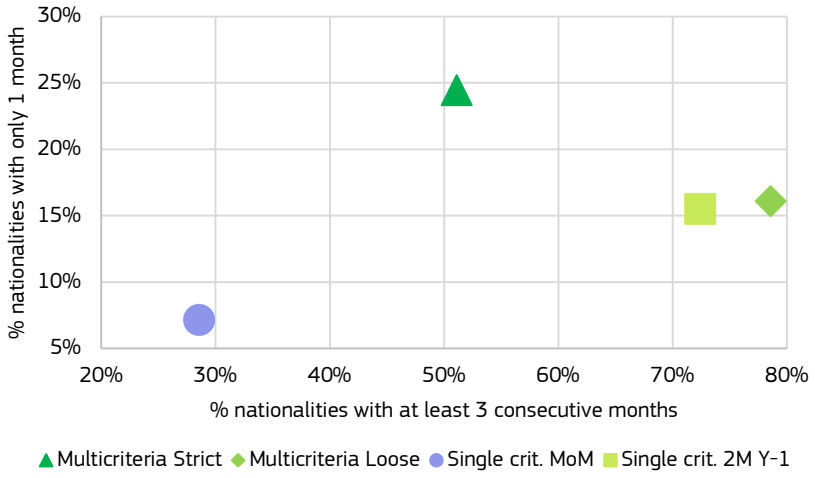
Figure 13 compares the four algorithms checking the percentage of nationalities that are detected only for one month (Y-axis) and the percentage of those detected for at least three consecutive months (X-axis), thus indicating a sustained trend. Ideally, algorithm should lie low on the y axis and high on the x-axis.

[43] Figures refer to the models considering only months with at least 10 IBCs recorded.

Among the four, the month-over-month algorithm reports very low shares of both nationalities detected only for one month (7%), and detected for three (or more) consecutive months (29%) (Figure 13, purple circle). In other words, this algorithm detects mostly for two months, thus meaning that can easily lead to an anomaly, but can hardly detect and follow emerging trends.[44] In fact, among the nationalities with consecutive months with this algorithm, 60% have only two consecutive months (Figure 11 in Annex 2).

The "loose" multi-criteria and the two-months year-on-year change both result in 16% of nationalities detected only for one month, with the former reporting 79% of nationalities detected for three (or more) consecutive months and the latter 72% (Figure 13). Hence, differently from the month-over-month, **both approaches manage to *detect* as well as to *follow* trends for different months**. When looking in details at the distribution of nationalities with consecutive months, **the "loose" multi-criteria seems to perform better in following emerging trends**: nationalities detected only for two consecutive months are 10-13% using this algorithm, compared to 20-22% in the two-month year-on-year change (Figure 11 in Annex).

Figure 13 Distribution of nationalities by detected months: 1 vs at least 3 consecutive months



Note: Months with at least 10 IBCs recorded

Finally, **the "strict" multi-criteria, given the additional imposed conditions, reports the highest share of nationalities detected only in one month** (24%), but with still 50% signalled for three (or more) consecutive months (Figure 14), thus better than the single-criterion algorithm month-over-month (29%). Moreover, the distribution of nationalities with consecutive months shows that 70% are detected between 3 and 8 months in the "strict" multi-criteria, compared to approximately 40% in the month-over-month (Figure 11 in Annex 2).

[44] It is in fact hard to imagine a trend that progressively increases every month by a given percentage, whatever this is. Or, if this happen, then the trend would already been clearly identified as an anomaly without the need of the algorithm.

Hence, from the distribution of nationalities by consecutive months detected, it appears that **the multi-criteria algorithms tend to better differentiate among situations and capturing growing trends than single-criterion algorithms** (Figure 11 in Annex 2). Finally, between the two multi-criteria, the "loose" shows a more right-skewed distribution towards more months reported (i.e., nationalities with 9 to 12 months detected) than the "strict" as it is less restrictive in the conditions necessary to flag an anomaly. As explained, the two versions of the multi-criteria algorithm can be used compared to understand whether a new emerging situation detected with the "loose" version is intensifying further or slowing down by applying the "strict" version.

Number of anomalous months in main nationalities detected

The multi-criteria algorithm can detect anomalies as early or even earlier than the single-criterion algorithms tested, especially the month over month. The "loose" version identifies potential changes worthy of attention while the "strict" version adds information on the severity of the change.

Differences between the four algorithms can be seen also from the number of months signalled as anomalies for a selection of nationalities with significant IBCs in 2023. The single-criterion month-on-month algorithm detects more or less 5 months, considerably below the other single-criterion algorithm and the "loose" multi-criteria (Figure 14). This seems to confirm challenges that come with this approach in following situations over time and/or detect them early on. The "loose" multi-criteria reports the same number of anomalies (or more) than the month-over-month approach in 22 out of 25 nationalities, with larger differences especially in the top eight. The "strict" multi-criteria reports similar numbers, and in some cases even more, than the month-over-month for the main nationalities detected (e.g., Burkina Faso, Guinea, Mali).

The comparison with the single-criterion two-month year-on-year change can help highlight the potential of multi-criteria algorithms in identifying different types of situation that might be either missed or detected only later on using a single indicator. While the number of signalled months in the "loose" and in the two-month year-on-year change is the same in the majority of cases, differences appear for nationalities with lower number of signals. When there are four or less months signalled by the two-month year-on-year change, the "loose" multi-criteria tends to have more signals, even four more in the case of Tunisia, Bangladesh and Ukraine. Hence, it seems that the multi-criteria "loose" detects a new emerging trend earlier on, with differences shrinking as situations develop, with more and more months being signalled under the two-month year-on-year change too.

The comparison between the "strict" multi-criteria and the "loose" further allows to evaluate the different anomalies, with the **"loose" detecting early on changes possibly worthy of attention, and the "strict" adding information on the severity of the change**. For instance, Afghanistan, Algeria and Tunisia, are all detected under the "loose" but not the "strict" algorithm, possibly identifying changes that might be relevant but are not yet quite significant across all dimensions and respect to each nationality's trend (Figure 14). For these three nationalities, for instance, the algorithm using two-month year-on-year change detects either two anomalies (Algeria and Tunisia) or none (Afghanistan). Similarly, the "strict" version can help understand **when a situation is slowing down in intensity**. For instance, in the cases of Guinea, Gambia and Senegal, the "strict" version detects the first anomaly either one month later or in the same month of the "loose" and the two-month year-on-year change, but then identify less months as anomalies.

Figure 14 Number of signalled months in 2023 (first month in brackets)- selection of nationalities with significant IBCs

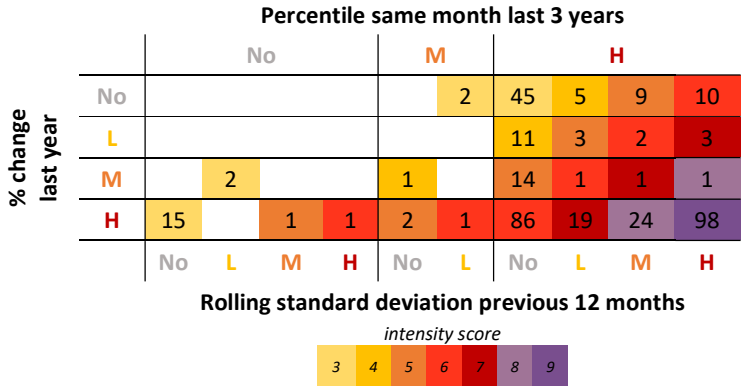
	Multicriteria Strict	Multicriteria Loose	Single crit. MoM	Single crit. 2M Y-1
Syria	2 (Sep)	9 (Jan)	4 (Mar)	5 (Jan)
Senegal	6 (Feb)	11 (Feb)	7 (Jan)	11 (Feb)
Guinea	6 (Feb)	10 (Jan)	3 (Feb)	10 (Jan)
Ivory Coast	5 (Feb)	8 (Jan)	4 (Feb)	9 (Jan)
Mali	8 (Feb)	12 (Jan)	6 (Feb)	12 (Jan)
Gambia	7 (Feb)	11 (Jan)	5 (Feb)	12 (Jan)
Burkina Faso	7 (Jan)	11 (Jan)	5 (Feb)	12 (Jan)
Sudan	5 (Apr)	9 (Feb)	5 (Jun)	9 (Feb)
Tunisia		6 (Feb)	6 (Feb)	2 (Mar)
Morocco	1 (Aug)	6 (Mar)	4 (Apr)	6 (Apr)
Turkey	2 (Aug)	4 (Aug)	4 (Mar)	4 (Sep)
Afghanistan		1 (Sep)	4 (May)	
Bangladesh	1 (Dec)	7 (Feb)	4 (Feb)	3 (Mar)
Pakistan	1 (Mar)	5 (Jan)	5 (Feb)	4 (Jan)
Cameroon	4 (Feb)	8 (Jan)	4 (Feb)	9 (Jan)
Egypt		3 (Mar)	3 (Mar)	4 (Jan)
Palestine	3 (Jul)	9 (Jan)	6 (Mar)	7 (Jan)
Eritrea	4 (Apr)	10 (Jan)	5 (Jan)	11 (Feb)
Iraq	1 (Jun)	6 (Jan)	4 (Mar)	4 (Jan)
Sierra Leone	5 (Jan)	9 (Jan)	5 (Mar)	10 (Jan)
Somalia	3 (Jul)	4 (Apr)	5 (Mar)	4 (May)
Benin	7 (Jan)	12 (Jan)	5 (Jan)	12 (Jan)
Algeria		4 (Mar)	5 (Apr)	2 (Mar)
Ukraine		6 (Jan)	4 (Jan)	2 (Jan)
Nigeria	2 (Jul)	4 (Feb)	3 (Mar)	3 (Jul)

Applying intensity scores

The multi-criteria algorithm provides further flexibility in evaluating anomalies by using an overall intensity score, which allows for a more nuanced interpretation and tries to overcome binomial "yes" or "no" types of anomalies.

The overall intensity score obtained by the combination of the three criteria, and their thresholds (see Section 4.1.4), used in multi-criteria algorithms provides further flexibility in evaluating the anomalies detected. The matrix below reports the distribution of months with at least 10 IBCs that were signalled as an anomaly under the "loose" multi-criteria algorithm. While the majority of anomalies is clustered in the lower-right-hand corner of the matrix reporting the higher overall intensity scores, several observations lie over intermediate scores given by different combinations across the three criteria.

Figure 15 Anomalies by criteria and overall intensity score – "loose" algorithm on months with at least 10 IBCs



Source: JRC elaboration. Note: The three thresholds for each criterion are reported as Low (L, score 1), Medium (M, score 2) and High (H, score 3). When anomalies are not overcoming the first threshold for a given criterion, they are listed under the "No" category.

Finally, the distribution of anomalies under the "loose" multi-criteria algorithm highlights the importance of **looking at different trends at the same time** in order to avoid missing potential important changes. About 20% of anomalies are reported in the first row, which means the year-on-year percentage change was below the first threshold (i.e., 25%). Hence, these months would have not been signalled under the "strict" multi-criteria algorithm, which imposes that at least one threshold is overcome for each criterion. Similarly, they would not be identified under a single-criterion algorithm using year-on-year change with the same threshold. However, since these observations overcame thresholds in either the percentile and standard deviation criteria (or both), they **might indicate very first signs of change** and thus possibly requiring further attention in the following months, and qualitative information to better contextualize situation-specific developments.

Key takeaways

- Multi-criteria algorithms offer flexibility in defining anomalies compared to single-criterion algorithms.
- There is no one-size-fits-all approach to defining thresholds across nationalities or data series.
- Flexibility should be prioritised to minimise false signals while considering different types of data.

This overview has shown the flexibility of multi-criteria algorithms in defining possible anomalies compared to single-criterion algorithms, while at the same time being able to better capture the evolution of situations. Moreover, the comparison between the "loose" and the "strict" multi-criteria further highlights the importance of considering different dimensions at once to lower the risk of exclusion due to excessively high thresholds. As explained, different observations marked as anomalies under the "loose" are not under the "strict" as they do not overcome the first threshold for the year-on-year criterion, and would have not been detected in a single criterion algorithm. Flexibility can thus be introduced, for instance, in terms of lowering thresholds, or considering two out of three criteria to be necessary to launch a "strict" signal instead of all three.

There is not a clear-cut approach to define thresholds across nationalities and, even less, across different types of data series (e.g., migration flows, drivers and triggering factors). Hence, flexibility should be prioritized, for instance through the use of "loose" and "strict" models, intensity scores, and/or changing thresholds, without compromising the need to minimise false signals.

Annex 6. Results of Early Warning methods for the selected use cases

In this section, the anomalies and relative scores detected by the multi-criteria and CUSUM models for the two use cases are presented. Figures 16 and 17 show the results for the first-time asylum applications in the EU lodged by Egyptian nationals (for the CUSUM and multi-criteria models, respectively). Figures 18 and 19 show the results for the first-time asylum applications in the EU lodged by Turkish nationals (for the CUSUM and multi-criteria models, respectively).

Figure 16. Warnings detected by the CUSUM model for first-time asylum applications in the EU lodged by Egyptian nationals. The fill colours represent the three severity level of the warning: "moderate" (yellow), "serious" (orange), and "critical" (red).

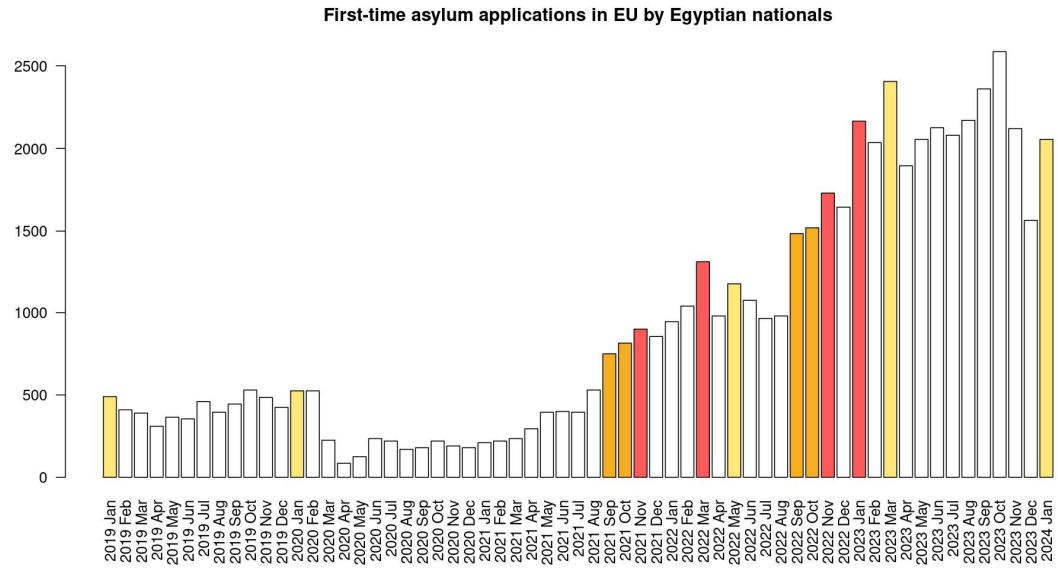


Figure 17. Warnings detected by the multi-criteria model for first-time asylum applications in the EU lodged by Egyptian nationals. The fill colours represent the intensity score of the warning.

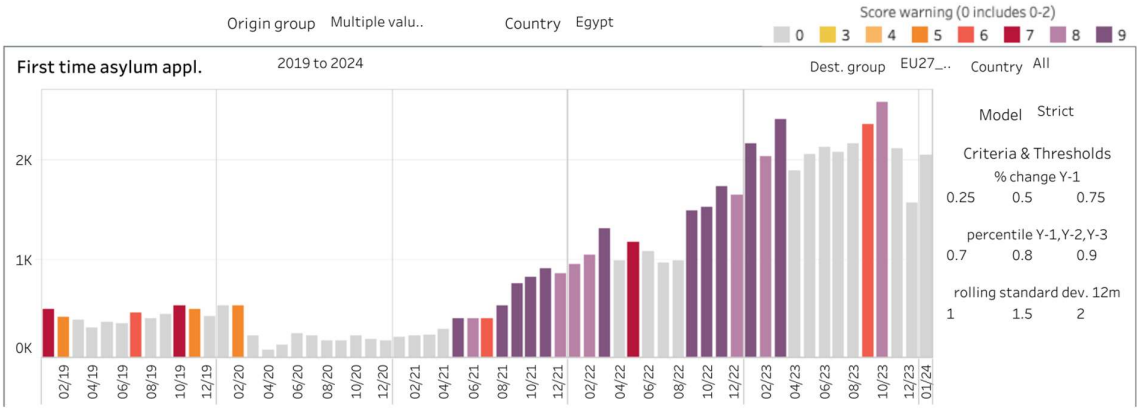


Figure 18. Warnings detected by the CUSUM model for first-time asylum applications in the EU lodged by Turkish nationals. The fill colours represent the three severity level of the warning: "moderate" (yellow), "serious" (orange), and "critical" (red).

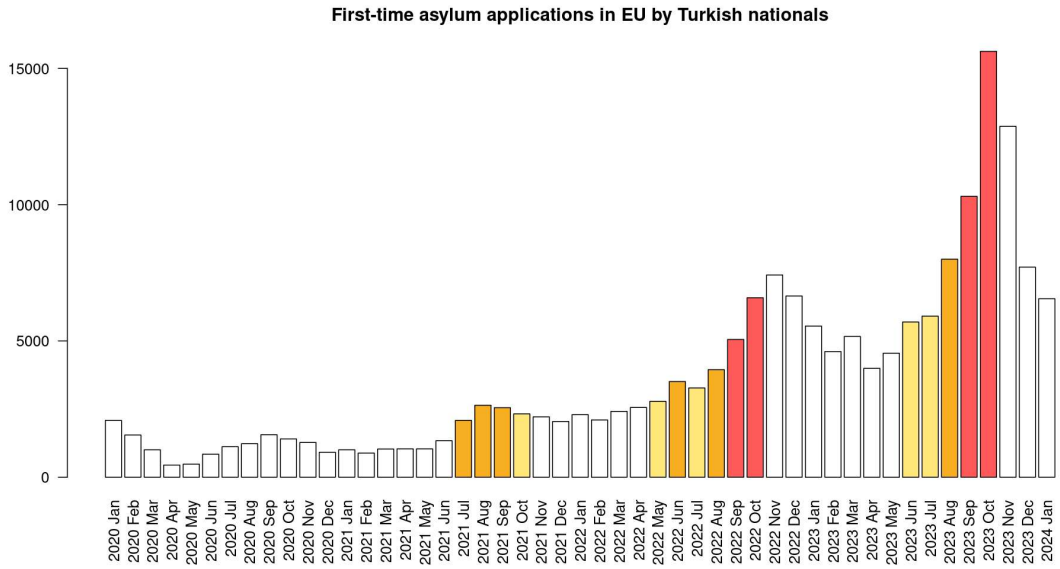
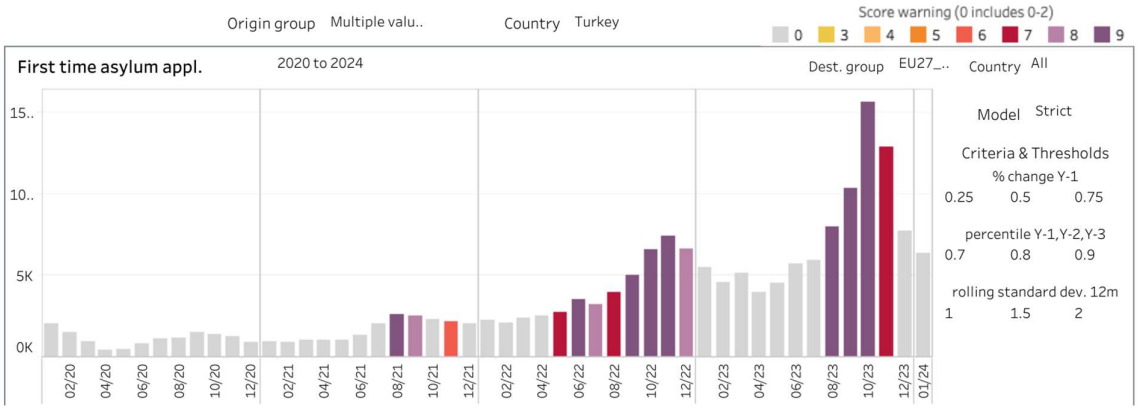


Figure 19. Warnings detected by the multi-criteria model for first-time asylum applications in the EU lodged by Turkish nationals. The fill colours represent the intensity score of the warning.



Getting in touch with the EU

In person

All over the European Union there are hundreds of Europe Direct centres. You can find the address of the centre nearest you online (european-union.europa.eu/contact-eu/meet-us_en).

On the phone or in writing

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

— by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),

— at the following standard number: +32 22999696,

— via the following form: european-union.europa.eu/contact-eu/write-us_en.

Finding information about the EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website (european-union.europa.eu).

EU publications

You can view or order EU publications at op.europa.eu/en/publications. Multiple copies of free publications can be obtained by contacting Europe Direct or your local documentation centre (european-union.europa.eu/contact-eu/meet-us_en).

EU law and related documents

For access to legal information from the EU, including all EU law since 1951 in all the official language versions, go to EUR-Lex (eur-lex.europa.eu).

EU open data

The portal data.europa.eu provides access to open datasets from the EU institutions, bodies and agencies. These can be downloaded and reused for free, for both commercial and non-commercial purposes. The portal also provides access to a wealth of datasets from European countries.

Science for policy

The Joint Research Centre (JRC) provides independent, evidence-based knowledge and science, supporting EU policies to positively impact society

EU Science Hub
Joint-research-centre.ec.europa.eu

Publications Office
of the European Union