Forecasting Migration Patterns and Land Border Encounters

Raquib Bin Yousuf*, Shengzhe Xu*, Patrick Butler*, Brian Mayer*,
Nathan Self*, David Mares† and Naren Ramakrishnan*
*Department of Computer Science, Virginia Tech, Arlington, VA
†Department of Political Science, University of California San Diego, San Diego, CA
Email: raquib@vt.edu, naren@cs.vt.edu

Abstract—**This paper leverages open source "big data" intelligence to develop predictive models that can provide timely, relevant and accurate indications, warning, and tracking of migration flows / movements of large groups (> 100 persons) through South and Central America to the southwest border of the United States. We describe experiments with a live forecasting setup, development and refinement of predictive models, and how machine learning models can yield insight into the factors underlying mass migration.**

Index Terms—**Anticipatory intelligence, forecasting, machine learning, mass migration.**

I. Introduction

Much of Latin America is at risk of natural disasters as well as political and social turmoil [1]. Governments are not equipped to handle these emergencies and are often incapable of being effective in safe conditions [2]. Political upheaval increases the potential for social unrest causing many citizens to consider migrating to other safer countries, often north and eventually to the southwest border of the United States of America (USA) [3]. As a result forecasting such surges in migration and land border encounters is a timely research problem [4].

This paper proposes a migration forecasting system from open source migration data, utilizing the Customs and Border Protection's (CBP's) southwest land border encounters data [5]. In this paper, we formalize and solve the forecasting algorithm in terms of the unique limitations of the migration scenario and develop predictive models to forecast land border encounters. Migration data from CBP are available only as a monthly aggregate with an approximately two months lag. Thus we must exploit surrogate indicators from textual and digital media resources, i.e., news (both English and Latin American sources in countries of origin). These surrogate indicators enable the creation of algorithms to provide anticipatory intelligence in a different granularity than the ground truth time series data, i.e., the daily encounter values.

Our models are intended to predict sudden changes in the anticipated encounters at the border, thus allowing the rapid and efficient deployment of available resources. We test our models for forecasting up to three months into the future from the available data timestamps but can predict further into the future. We also model spatial dependencies across different components and captured area-specific patterns in a multi-task multi-level learning framework. Additionally, we adapted a multi-faceted evaluation methodology to perform both retrospective and prospective assessments of forecasts using measures such as lead time, accuracy, and self-consistency.

II. Background & Literature Review

A. Migration Triggers

Migrations responds to push and pull factors [6]; the push factors force people to seek safety or a better life elsewhere, while the pull factors make certain places attractive for relocation. In certain places and at certain times either push or pull factors may be relatively more important but they always work in tandem. Push and pull factors operate on both legal and undocumented migration. Most migration is undertaken by individuals or small groups of people and readily absorbed by receiving countries in need of labor because of rapid economic growth or decreasing demographic trends. Mass migrations, however, can number into the millions and occur over a brief period of time, thereby overwhelming the capacity of countries to which they flow either in transit or as a final destination.

Mass migrations are not new phenomena, but they may be increasing because of three factors. Climate change is now producing a plethora of natural disasters (drought, flooding, wildfires, rising sea levels, etc.) more often and with greater intensity, thereby making some geographic spaces either uninhabitable or significantly difficult for people to have an acceptable standard of living [7]. The increasing shift to authoritarian governments (including geographic spaces effectively governed by organized crime) around the world is another push factor with increasing impact, often operating in combination with natural disasters. Some of these governments take advantage of the weakening of global norms against gross violations of human rights, ideological justifications for violence, and ethnic cleansing to create situations of extreme insecurity for large segments of the national population [2]. There is also a third factor of increasing importance and it operates on the pull side: demographic decline in wealthy nations. Their economies as well as their aging populations need a significant influx of labor and offer higher wages as well as social support than elsewhere, thereby becoming attractive destinations for desperate migrants.

Latin America is among the regions of the world that is most vulnerable to natural disasters increasingly stimulated by climate change [8]–[10]. Caribbean islands and Central America have been devastated by recent hurricanes, drought afflicts South America, and rising sea levels impact all of its coastal communities. Most governments have neither the infrastructure, technical capacities nor funds to adequately respond to these vulnerabilities and the crises that develop when disaster hits. In addition, the return of authoritarian government has produced massive out-migration from Venezuelans seeking to escape economic collapse and political retribution. Geographic proximity, a growing economy and demographic decline make the U.S. an attractive destination for migrants.

B. Forecasting using open source data

Time series forecasting has been extensively researched due to its widespread applicability in a variety of fields, such as finance [11], weather [12], traffic [13], and others. In the past decade, deep learning has been thoroughly researched and has led to the advancement of new technical modeling approaches. These approaches have moved away from purely traditional statistical methods, such as AutoRegressive Integrated Moving Average (ARIMA) [14] and generalized additive models (GAM) [15]. Such traditional methods often face difficulties in capturing complex patterns and non-linearities that exist in real-world data. Recent progress in deep learning, exemplified by Long Short-Term Memory (LSTM) networks [16], [17] and Transformers [18], has greatly enhanced the accuracy of forecasting by adeptly capturing long-term dependencies and non-linear patterns in time series data. There also exist composite approaches such as Prophet [19] and timeGPT [20] which aim to provide a richer vocabulary for modeling and forecasting. However, merely optimizing the time series prediction algorithm is invariably insufficient in practical applications due to hierarchical and bursty patterns of behavior. The classic paper of Kleinberg [21] detects bursts in a stream of events. The EMBERS series of methods [22]–[27] implements event-level modeling and forecasting by utilizing numerous information sources across granularities as real time input factors.

III. Methodology

A. Data Collection

1) Data Sources: The ground truth data was collected from the CBP Data Portal [28]. Monthly encounters (including Title 8 and Title 42) from CBP and Office of Field Operations (OFO), e.g., ports of entry, dating back to October 2020 were collected and updated each month through March of 2023. Additional information is provided with the encounter data such as demographic information (categories indicating age, admissibility, and familial relationship), location of encounter, citizenship, title of authority (the authority under which the non-citizen was processed), and encounter type (the category of encounter based on Title of Authority and component, i.e., CBP or OFO). This information was collected to enhance predictions and provide more specific predictions, e.g., by location or citizenship.

We investigated several data sources to improve our understanding of various aspects of migration and specifically identified innovative surrogate data sources. Specifically, geo-located social media data support early event detection and event prediction as well as serve as indicators for ongoing conflict on the ground. Online media content, such as news media and blog posts, indicate events that already took place and may act as early warning indicators for predicting future events. It should be noted that “Big (Crisis) Data” does have its problems. The data is noisy, sometimes inaccurate, and can contain bias. It can also be difficult to combine with traditional data sources e.g., different frequencies, unstructured data sources, or different units of measurement.

Our data ingest for training and running models was derived from two primary sources, NewsAPI, and webz.io, using keyword data queries. These sources include news articles, i.e., long form text. We collected English and Spanish data using both English and Spanish keywords. Some queries required a Central/South American country name to be listed in the article/tweet to limit and narrow the results on the topics and events of interest. Approximately 5.6 GB of data was originally downloaded from the 2 sources dating back to October 1, 2020: i) 2.8M+ articles from NewsAPI, and ii) 3.4M+ articles from webz.io.

2) Data Queries: The keyword list used to query the data was developed using guidance from social and political science subject matter experts. The keyword list focused on “triggers” to migration. As stated earlier, increases and decreases in migration are driven by push (domestic violence, economic collapse, natural resource issues) and pull (jobs, benefits, community) factors. These push and pull factors could either encourage (policy, domestic violence, environmental stressors) or discourage (policy, transit issues, blockages) migration. The triggers also assume that migration flows are well organized because they are high-risk movements. Triggers also match the forced displacement system that the United Nations High Commissioner for Refugees recommends as the theoretical basis to evaluate possible contributions of “Big (Crisis) Data” to predictive models of forced displacement [29].

Additionally, combinations of words, words in certain contexts, or words from certain sources may have meaning but a holistic count of those words might not, e.g., a count of the words “Cubans,” “Venezuelans,” “Nicaraguans,” or “Haitians” versus those words in a Costa Rican newspaper. We addressed these issues and generated a list consisting of 481 keywords.

B. Data Enrichment

Data is ingested via multiple APIs and undergoes several enrichment steps. Entities (people, locations, organizations, dates) are extracted and labeled. Keyword and country name occurrences are tracked for each article. Additionally, detailed distributed word vector modeling is applied.

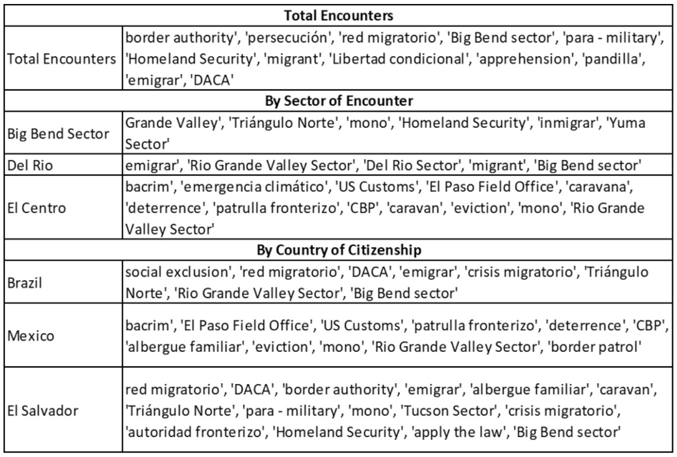
Geocoding ensures an accurate understanding of data origin and identifies the most impacted locations. The article’s location is predicted using document metadata and content, with the enrichment process determining where it was published, the reporting source, and topical content. Conflicts between these locations are resolved using models inspired by the

EMBERS system [22]–[27]. We also extract South/Central American country names from the articles. After enrichment, the data size grew to 650GB, a 100-fold increase from the original ingested data.

C. Analysis and Prediction Models

1) Keyword Analysis: From the aforementioned 481 keywords, we used regression models for feature selection, identifying the top 60 most positively and 60 most negatively correlated keywords with total encounters and specific demographic segments (sectors and citizenship).

Our prediction models handle keyword counts differently: some use individual keyword counts as features, while others use aggregated counts. Individualized counts help models detect specific changes (e.g., "Title 42"), whereas aggregated counts pool variance to capture broader trends. When predicting by citizenship, keywords must be mentioned with the country of interest, ensuring relevance. A sampling of the top 100 correlated keywords for specific demographics are shown in Figure 1.



Total Encounters	
Total Encounters	border authority', 'persecución', 'red migratorio', 'Big Bend sector', 'para - military', 'Homeland Security', 'migrant', 'Libertad condicional', 'apprehension', 'pandilla', 'emigrar', 'DACA'
By Sector of Encounter	
Big Bend Sector	Grande Valley', 'Triángulo Norte', 'mono', 'Homeland Security', 'inmigrar', 'Yuma Sector'
Del Rio	emigrar', 'Rio Grande Valley Sector', 'Del Rio Sector', 'migrant', 'Big Bend sector'
El Centro	bacrim', 'emergencia climático', 'US Customs', 'El Paso Field Office', 'caravana', 'deterrence', 'patrulla fronterizo', 'CBP', 'caravan', 'eviction', 'mono', 'Rio Grande Valley Sector'
By Country of Citizenship	
Brazil	social exclusion', 'red migratorio', 'DACA', 'emigrar', 'crisis migratorio', 'Triángulo Norte', 'Rio Grande Valley Sector', 'Big Bend sector'
Mexico	bacrim', 'El Paso Field Office', 'US Customs', 'patrulla fronterizo', 'deterrence', 'CBP', 'albergue familiar', 'eviction', 'mono', 'Rio Grande Valley Sector', 'border patrol'
El Salvador	red migratorio', 'DACA', 'border authority', 'emigrar', 'albergue familiar', 'caravan', 'Triángulo Norte', 'para - military', 'mono', 'Tucson Sector', 'crisis migratorio', 'autoridad fronterizo', 'Homeland Security', 'apply the law', 'Big Bend sector'

Fig. 1: Sampling of keyword sets that appear in every update of the Top 100 correlated keywords with a specific demographic of encounters.

2) Burst Analysis: This analysis uses the Bursty and Hierarchical Structure algorithm described in [21] to determine a baseline level of keyword counts and uses this to determine when bursts of certain keywords or aggregations of keyword groups (e.g., Top 60 correlated keywords) occur.

When there is a sustained burst in an aggregation of keywords on a daily level we then predict a burst in encounters (in the proceeding month) for the applicable demographic. Additionally, when there is a sustained burst in a specific keyword on a daily level, the keyword is added to a daily list of important keywords. These provide context for an analyst to better understand the trends that are occurring in the ingested data and the predicted encounters. These keywords can be sorted or filtered into categories, such as geo-political locations, people, nationalities, etc.

3) Prediction Models: We begin by formalizing a problem definition, which guides the development of several predictive models and algorithms described in the following subsections and compared in the results section. As mentioned, the ground truth data is available on a two-month delay yet all models are making predictions several months into the future.

Problem Definition: A migration encounter sequence, denoted as $\mathcal{E} = \{e_1, e_2, \ldots, e_m\}$, is provided at the monthly level, where each e_m represents the ground truth label as a numerical value. It is important to note that the ground truth label has a two-month delay. This means that when making model predictions, the available historical sequence $\mathcal{E}$ only includes $\{e_1, e_2, \ldots, e_{m-2}\}$. A news keyword sequence, $\mathcal{K} = \{k_1, k_2, \ldots, k_d\}$, is generated from a mix of news sources and provided daily. The goal is to forecast the migration encounter sequence for the next month at a daily level, denoted as $\mathcal{E}' = \{e'_d, e'_{d+1}, \ldots, e'_{d+30}\}$, given the historical sequences $\mathcal{E}$ and latest $\mathcal{K}$. When forecasting e'_d, the input feature $\mathcal{K}$ will include $\{k_1, k_2, \ldots, k_{d-1}\}$.

M1: Bursty Regression: This model uses correlations between the aggregated counts of the top 60 keywords and encounters and applies that to the daily keyword usage to predict daily encounters. It then aggregates daily counts to monthly counts for comparison to the ground truth. Unfortunately, this model was not as accurate as other models because news cycles decay more rapidly than actual encounters. Therefore the model predicts large drops in encounters before the drops actually are realized resulting in a higher error despite a similar trend/pattern in the keyword counts. This model was used as a starting point for further tuning.

Burst Detection and Enumeration: To detect the bursty nature of the target encounter sequence $\mathcal{E}$ of $m-2$ labels, we adopted a bursty and hierarchical-structure-in-streams design and automaton-based bursty detection [21]. A simple example of this bursty automaton is a two-state model, corresponding to "low" and "high" activity levels. In practice, the infinite-state version $\mathcal{A}^*_{s,\gamma}$ can be defined as:

$$
\begin{aligned}
f(x) &= \alpha_i e^{-\alpha_i x} \\
\alpha_i &= \hat{g}^{-1} s^i \\
\tau(i,j) &= \mathrm{relu}[(j-i)\gamma \ln n]
\end{aligned}
\tag{1}
$$

Eq. 1 describes the infinite automaton within a bursty detection model. The hyperparameters γ and s represent the cost of moving to a higher state and a scale parameter, respectively.

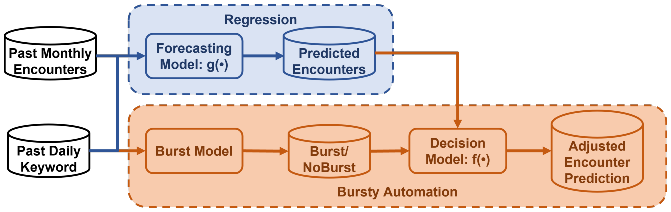


Fig. 2: M2 pipeline.

M2: Bursty Regression with Multi-Armed Bandits: The goal of this model is to predict e_{d+1} (which is daily) and $burst_{d+1}$, i.e., burst prediction for the next day, with two input streams:

1) e_{m-2}: Encounters for month m with a 2-month lag.
2) k_d: Daily keyword count for day d.

The main challenge is that the two inputs e_{m-2} and k_d have asynchronous updates and different granularity. The model was therefore designed to find the direct relationship between the daily keyword and the daily encounters that can be mimicked. Given each state (burst or no burst), we have 3 possible actions:

1) Decrease the predicted encounter with decay.
2) Increase the predicted encounter as normal.
3) Increase the predicted encounter under a burst.

The model's input data consists of burst detections based on news keywords. Although daily ground truth encounters are not available, the model retains the memory of the most recent predicted daily encounters, which helps carry forward trends into the next day, denoted as e'_{d+1}. The forecasting model is a regression function $g(\cdot)$, trained using historical data from the labeled period, which includes available monthly encounter trends and daily ground truth data for the keywords. The decision model $f(\cdot)$, shown in Figure 2, includes parameters β, which are learned to capture the effects of decay and the magnitude of bursts. To achieve daily-level predictions, the values in the set $\mathcal{E}' = e'_{m-2}, e'_{m-2+1d}, \ldots, e'_{d-1}, e'_d$ are predicted auto-regressively.

$$\begin{aligned} e'_{d+1} &= f[k_d, g(e_{1,\ldots,m-2}, k_{1,\ldots m-2}), e'_d] \\ &= \begin{cases} e'_d * \beta_1, & \text{if decay} \\ g(k_d) * \beta_2, & \text{if rise, and } \textit{burst}(e'_d) \text{ is True} \\ g(k_d) * \beta_3, & \text{if rise, and } \textit{burst}(e'_d) \text{ is False} \end{cases} \end{aligned} \tag{2}$$

Eq. 2 represents a value function designed to predict future encounters, e'_{d+1}, which is predicted by incorporating inputs with varying levels of granularity. This process involves using ground truth encounter labels e'_m with a two-month delay and combining e'_{m-2}, on a monthly basis, and daily news keywords, k_d. Although the ground truth label, e'_m, is based on coarser, monthly data, the prediction method can still achieve detailed daily forecasts, represented as e'_{d+1}.

To determine the parameters in the equation above, we employ the Multi-Armed Bandits (MAB) framework. Let the action chosen at time t be denoted by A_t, and let the corresponding payoff be R_t. The expected payoff for any given action a is denoted by $q_{(}a)$, as defined in Eq. 3. Our goal is to estimate the value of action a at time t, represented by $Q_t(a)$, and we aim for this estimate to closely approximate $q_*(a)$.

$$q_*(a) = \mathbb{E}[R_t | A_t = a] \tag{3}$$

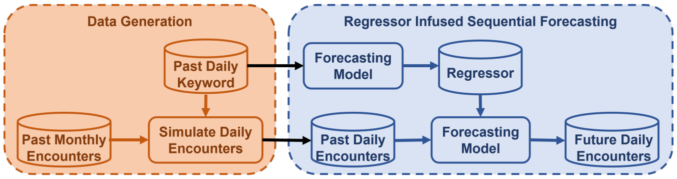


Fig. 3: M3 and M4 pipeline.

M3 and M4: Keyword Infused Forecasting: M3 forecasts future daily encounters for the rest of the month in a sequential manner using regressor-infused time series forecasting. M4 is the same as M3 but it uses the individual keyword counts mentioned in the keyword analysis as multiple separate features. All past observed keywords are modeled to act as regressors for the final forecasting model which outputs the daily encounter values. There are two steps.

1) The first step is data generation. A challenge of time-series forecasting in this scenario is the granularity of the observed data. All of our encounter data are aggregated monthly but we build a daily encounter prediction model. So, to forecast the number of daily encounters we employ a mathematical model to simulate daily encounter data from historical monthly totals. The model assumes that the growth of cumulative daily encounters follows the growth of cumulative daily keyword counts. Curve fitting parameters are initialized from keyword data using this equation.

$$\mathcal{N}(\tau) = \mathcal{N}_0 \cdot (1 + \gamma\xi)^{\tau}$$

We solve for the growth rate γ:

$$\begin{aligned} \left(\left(\frac{\mathcal{N}(\tau)}{\mathcal{N}_0}\right)^{\frac{1}{\tau}} - 1\right) \cdot \xi^{-1} &\equiv \left(e^{\frac{1}{\tau}\ln\left(\frac{\mathcal{N}(\tau)}{\mathcal{N}_0}\right)} - 1\right) \cdot \xi^{-1} \\ &\equiv \left(e^{\frac{1}{\tau}\ln\left(\frac{\sum_{d=1}^{\tau} n(d)}{n(1)}\right)} - 1\right) \cdot \xi^{-1} \end{aligned} \tag{4}$$

where $\mathcal{N}(\tau)$represents the value of the cumulative encounter at time τ, i.e., $\mathcal{N}(\tau) = \sum_{d=1}^{\tau} n(d)$. $\mathcal{N}_0$ is the initial value of the cumulative encounter count for first day of the month, i.e., $\mathcal{N}_0 = \sum_{d=1}^{1} n(d) = n(1)$. γ represents the growth rate of the cumulative sum and $\mathcal{N}$ per time unit τ. ξ is a constant that modifies the growth rate γ with τ as the time variable.

2) The second step is sequential forecasting. We use past encounters and present keyword counts to first forecast keywords for the rest of the month, and then, encounters for the rest of the month, using forecasted keywords as a regressor. We adapted the additive forecasting model from [19] to model our forecasting as the following:

$$e(\theta) = \psi(\theta) + \sum_{i=1}^{n} \beta_i k_i(\theta) + \epsilon_\theta \tag{5}$$

where $e(\theta)$ represents the number of encounters at time θ. $\psi(\theta)$ is the combined trend, seasonality, and holiday component, i.e., $\psi(\theta) = g(\theta) + s(\theta) + h(\theta)$. ϵ_θ is the error term. $\sum_{i=1}^{n} \beta_i k_i(\theta)$ encapsulates multiple keywords as a separate regressor to model the forecast, i.e., β_i is the i^{th} coefficient quantifying the influence of i^{th} keyword on encounters; $k_i(\theta)$ is the i^{th} keyword frequency at time θ. The trend component $g(\theta)$ can be modeled as either a piece-wise linear function or a logistic growth curve. $s(\theta)$ is captured using a Fourier series, which allows the model to flexibly fit recurring patterns over different periods. $h(\theta)$ are modeled as indicator variables for known events that can cause spikes or drops in encounters.

Figure 3 shows the pipeline for these two models. Table I shows the difference between the four analytical models in regards to horizon (forecast, condition), granularity and core technique.

TABLE I: Comparison between different models. Each model has the same conditioning horizon of monthly encounters with two months of lag and daily keywords with one day lag.

Models	**Forecasting Granularity**	**Forecast Horizon**
M1: Bursty Regression	Daily	Next day
M2: Bursty Regression with Multi-Armed Bandits	Daily	Next day
M3: Keyword Infused Forecasting	Daily	Rest of month
M4: Multiple Keyword Infused Forecasting	Daily	Rest of month

IV. Results

A. Burst prediction

To evaluate burst prediction we subjectively defined bursts as increases in encounters from one month to the next of 10% or more. This analysis is correct in predicting a burst 91% of the time (precision) and predicts 82% of bursts that actually occurred (recall). Sector and country specific bursts had a precision of around 80% and recall of around 60%. Therefore an analyst can expect to be forewarned that a significant increase will occur and in what specific demographic. Unfortunately, this analysis does not predict how large the bursts in encounters will be. Figure 4 shows the burst prediction for total encounters overlaid on the actual count of total encounters.

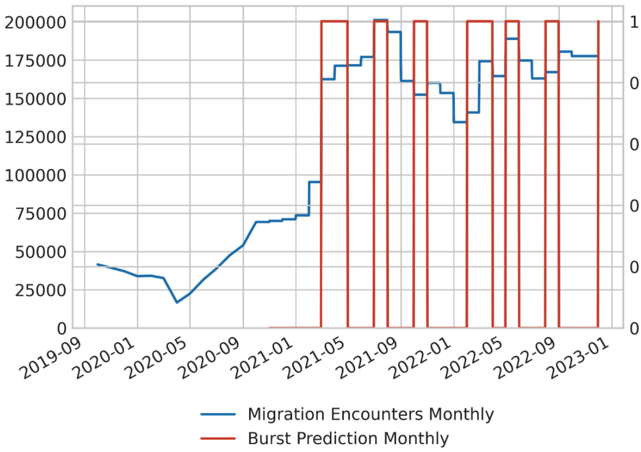


Fig. 4: Burst predictions compared to actual encounters.

B. Encounter Prediction Metrics

Before discussing the results, we describe the metric used to measure the forecasting model. Models are evaluated on three key factors, i) lead time (how far in advance are the forecasts provided in days), ii) mean absolute percentage error (MAPE), and iii) self-consistency (if the forecasting model is stable against changes due to real-time data). MAPE is calculated using the formula:

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{e_t - e'_t}{e_t}\right| \times 100 \tag{6}$$

where n is the number of predictions, e_t is the actual value of encounters at time t, and e'_t is the forecasted value of encounters at time t.

C. Daily prediction

Our four models have different lead times. These are listed in the Forecast Horizon column of Table I. The MAPE for each of the model's ability to predict total encounters over multiple time periods is shown in Table II. The actual and forecasted values for total encounters of each model are plotted in Figure 5.

TABLE II: Evaluation results of models

Model	**Evaluation Time Period**	**MAPE**
M2	Jan 21 - Sep 22	26.3%
M2	Dec 21 - Sep 22	10.5%
M3	Mar 22 - Oct 22	11.3%
M4	Feb 21 - Dec 22	20.2%
M4	Feb 22 - Dec 22	09.3%

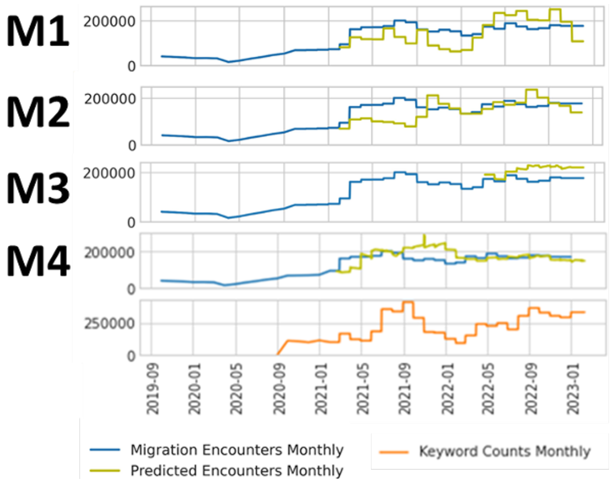


Fig. 5: Actual and forecasted values for total encounters from each model.

D. Self-consistency analysis

This evaluation assesses the change of the monthly prediction as we progress throughout the month. Optimally the predictive models would be fairly stable with changes due to the real-time data (news and corresponding keywords) we are seeing each day. Figure 6 shows the measure for a sample month which indicates that the models are fairly consistent.

E. User interface

We also developed a user interface where we stitch together all the relevant results from all analyses and models to present

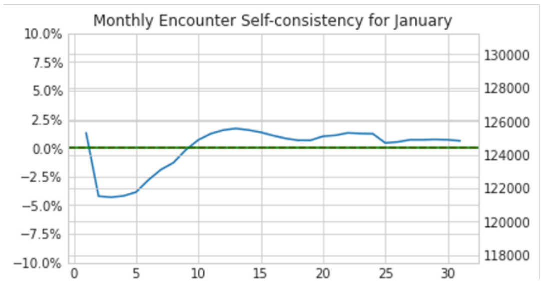


Fig. 6: Self-consistency evaluation for encounter prediction (blue) through the days of the month

to relevant authorities. Users can select and filter the results. We also show the bursting keywords which can be filtered and sorted into categories, e.g., geo-political locations, people, nationalities, etc. An example output of bursting keywords in the user interface is shown in Figure 7.

V. Future Work and Conclusion

Migration on the southwest border of the USA has been fluctuating for many different reasons which raises the necessity of robust and real-time prediction models of migration encounters. In this research, we presented the development and deployment of a predictive system built with the help of open-source data. We also identify key challenges to predicting migration . As future work for this research, we aim to enhance the accuracy of the granular (i.e., sector, country, demographic) models. We also plan to use Dynamic Query Expansion to expand our keyword selection process [22]. Moreover, we aim to develop nested multi-instance learning models (nMIL) [30] which can show the causal and interventional analysis for an important migration event through the temporal sequence of precursors. Migration remains an important challenge for the USA and novel research is important to address this challenge.

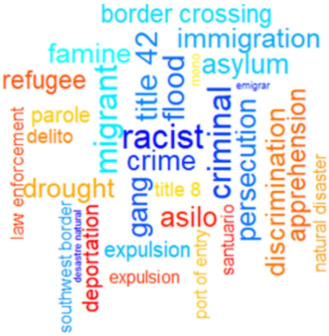


Fig. 7: Keywords contributing to bursts.

References

[1] UNICEF, "9 out of 10 children have been exposed to at least two climate shocks," https://www.unicef.org/lac/en/stories/9-out-of-10-children-have-been-exposed-to-at-least-two-climate-shocks, 2022.

[2] J. E. L. Campos and D. Lien, "Political instability and illegal immigration," *Journal of population economics*, vol. 8, no. 1, pp. 23–33, 1995.

[3] G. Hanson, P. Orrenius, and M. Zavodny, "Us immigration from latin america in historical perspective," *Journal of Economic Perspectives*, vol. 37, no. 1, pp. 199–222, 2023.

[4] J. Gramlich, "Migrant encounters at the U.S.-Mexico border hit a record high at the end of 2023 — pewresearch.org," https://www.pewresearch.org/short-reads/2024/02/15/migrant-encounters-at-the-us-mexico-border-hit-a-record-high-at-the-end-of-2023/, 2024.

[5] "CBP Public Data Portal | U.S. Customs and Border Protection." [Online]. Available: https://www.cbp.gov/newsroom/stats/cbp-public-data-portal

[6] R. Parkes, "Migration: new 'push' and 'pull' dynamics," European Union Institute for Security Studies (EUISS), Tech. Rep., 2015. [Online]. Available: http://www.jstor.org/stable/resrep06863

[7] R. Obokata, L. Veronis, and R. McLeman, "Empirical research on international environmental migration: a systematic review," *Population and Environment*, vol. 36, no. 1, pp. 111–135, 2014. [Online]. Available: http://www.jstor.org/stable/24769636

[8] E. Gencer, "An overview of urban vulnerability to natural disasters and climate change in central america & the caribbean region," 2013.

[9] A. Stein, "Natural disasters, climate change and environmental challenges in central america," in *Handbook of Central American Governance*. Routledge, 2013, pp. 59–74.

[10] C. Charvériat, "Natural disasters in latin america and the caribbean: An overview of risk," 2000.

[11] O. B. Sezer, M. U. Gudelek, and A. M. Ozbayoglu, "Financial time series forecasting with deep learning: A systematic literature review: 2005–2019," *Applied soft computing*, vol. 90, p. 106181, 2020.

[12] Z. Karevan and J. A. Suykens, "Transductive lstm for time-series prediction: An application to weather forecasting," *Neural Networks*, vol. 125, pp. 1–9, 2020.

[13] J. Zhang, Y. Zheng, and D. Qi, "Deep spatio-temporal residual networks for citywide crowd flows prediction," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 31, no. 1, 2017.

[14] A. C. Harvey, *ARIMA Models*. London: Palgrave Macmillan UK, 1990, pp. 22–24.

[15] T. J. Hastie, "Generalized additive models," in *Statistical models in S*. Routledge, 2017, pp. 249–307.

[16] S. Hochreiter, "Long short-term memory," *Neural Computation MIT-Press*, 1997.

[17] A. Sherstinsky, "Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network," *Physica D: Nonlinear Phenomena*, vol. 404, p. 132306, 2020.

[18] A. Vaswani, "Attention is all you need," *Advances in Neural Information Processing Systems*, 2017.

[19] S. J. Taylor and B. Letham, "Forecasting at scale," *The American Statistician*, vol. 72, no. 1, pp. 37–45, 2018.

[20] A. Garza and M. Mergenthaler-Canseco, "Timegpt-1," 2023.

[21] J. Kleinberg, "Bursty and hierarchical structure in streams," in *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, 2002, pp. 91–101.

[22] N. Ramakrishnan *et al.*, "'beating the news' with embers: forecasting civil unrest using open source indicators," in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2014, pp. 1799–1808.

[23] A. Doyle *et al.*, "The embers architecture for streaming predictive analytics," in *2014 IEEE international conference on big data (big data)*. IEEE, 2014, pp. 11–13.

[24] ——, "Forecasting significant societal events using the embers streaming predictive analytics system," *Big data*, vol. 2, no. 4, pp. 185–195, 2014.

[25] N. Ramakrishnan *et al.*, "Model-based forecasting of significant societal events," *IEEE Intelligent Systems*, vol. 30, no. 05, pp. 86–90, 2015.

[26] P. Saraf and N. Ramakrishnan, "Embers autogsr: Automated coding of civil unrest events," in *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 599–608.

[27] S. Muthiah *et al.*, "Embers at 4 years: Experiences operating an open source indicators forecasting system," in *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 2016, pp. 205–214.

[28] U.S. Customs and Border Protection, "CBP Public Data Portal — cbp.gov," https://www.cbp.gov/newsroom/stats/cbp-public-data-portal, 2024.

[29] "Big (Crisis) data for predictive models – A literature review." [Online]. Available: https://www.unhcr.org/media/big-crisis-data-predictive-models-literature-review

[30] Y. Ning, S. Muthiah, H. Rangwala, and N. Ramakrishnan, "Modeling precursors for event forecasting via nested multi-instance learning," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 1095–1104. [Online]. Available: https://doi.org/10.1145/2939672.2939802