

# Using an interpretable Machine Learning approach to study the drivers of International Migration

Harold Silvère Kiossou<sup>1\*</sup>, Yannik Schenk<sup>1</sup>, Frédéric Docquier<sup>1</sup>, Vinasetan Ratheil Houndji<sup>1,2</sup>, Siegfried Nijssen<sup>1</sup>, Pierre Schaus<sup>1</sup>

<sup>1</sup>UCLouvain, Belgium

<sup>2</sup>UAC, Benin

{harold.kiossou, yannik.schenk, frederic.docquier, vinasetan.houndji, siegfried.nijssen, pierre.schaus}@uclouvain.be

## Abstract

Globally increasing migration pressures call for new modelling approaches in order to design effective policies. It is important to have not only efficient models to predict migration flows but also to understand how specific parameters influence these flows. In this paper, we propose an artificial neural network (ANN) to model international migration. Moreover, we use a technique for interpreting machine learning models, namely Partial Dependence Plots (PDP), to show that one can well study the effects of drivers behind international migration. We train and evaluate the model on a dataset containing annual international bilateral migration from 1960 to 2010 from 175 origin countries to 33 mainly OECD destinations, along with the main determinants as identified in the migration literature. The experiments carried out confirm that: 1) the ANN model is more efficient w.r.t. a traditional model, and 2) using PDP we are able to gain additional insights on the specific effects of the migration drivers. This approach provides much more information than only using the feature importance information used in previous works.

## 1 Introduction

Today more people are on the move than ever before. The current number of international migrants, defined as individuals residing outside of their country of birth, is estimated to be almost 272 million people globally, representing 3.3% of the world’s population [DESA, 2019]. Drivers of migration, voluntary or forced, are numerous and often interrelated. Factors linked to economic prosperity, inequality, demography, conflict and persecution, environmental change and natural disasters create a complex interplay of migration incentives and opportunities. On the other hand, migration flows have large implications for receiving countries, impacting economic and demographic structure, culture, environment, spread of infectious diseases, etc. Against this context, deepening the understanding of *which* factors determine international migration

and *how* they do so is crucial to build effective governance strategies and policies.

Modeling human mobility is usually done using extended gravity models [Zipf, 1946; Letouzé *et al.*, 2009; Anderson, 2011] or radiation models [Simini *et al.*, 2012; Masucci *et al.*, 2013]. Both are analytical methods which rely on linear or log-linear relationships between the independents (migration flows) and a set of covariates (features). While these methods often attempt to identify causal effects, they lack predictive power and struggle to match important non-linearities in the migration nexus.

In this paper, we combine a long series of *annual* flows of international migrants with a full set of the key migration drivers identified in previous research. Our contribution is twofold. First, we use this unique data set to show that our model of artificial neural networks (ANN) performs well in estimating yearly migration flows and outperforms the canonical gravity model. Second, we demonstrate how partial dependence plots as an ML interpretability technique can be used to explore complex non-linearities [Friedman, 2001; Molnar, 2018], allowing for insights that go beyond traditional radiation or gravity models. While we do not claim causality, our results highlight the potential of ML-based techniques as an exploratory tool to enrich the usual econometric analysis in a variety of fields.

The remainder of this paper is organized as follows. We first describe some related work as well as the data used in Section 2. Then, in Section 3, we explain our model and the learning process. Finally, we show the results of the experiments carried out in Section 4.

## 2 Background

The aim of predicting human mobility is to estimate the migrant flow  $T_{i,j}$  from an origin country  $i$  to a destination country  $j$ . Let  $\hat{T}_{i,j}$  denote this estimation.

**Related works.** Estimating migration flows using ML tools is a recent approach. To our knowledge, the only work published on this is Robinson and Dilkina [2017].

The authors implement two ML-based techniques to forecast internal and international migration: the “*extreme*” *gradient boosting regression (XGBoost)* model, and a deep learning based *artificial neural network (ANN)* model to estimate  $\hat{T}_{i,j} = f(\text{features})$  from a set of features. They estimate

---

\*Contact Author

Input features $_{i,j,t}$	Description
$gdp_{i,t}$	Gross Domestic Product per Capita of origin country $i$ during the year $t$
$gdp_{j,t}$	Gross Domestic Product per Capita of destination country $j$ during the year $t$
$pop_{i,t}$	Population size of origin country $i$ during year $t$
$pop_{j,t}$	Population size of destination country $j$ during year $t$
$density_{i,t}$	Population per $km^2$ at origin country $i$ during year $t$
$density_{j,t}$	Population per $km^2$ at destination country $j$ during year $t$
$dep\_ratio_{i,t}$	Dependency ratio at origin country $i$ during year $t$
$dep\_ratio_{j,t}$	Dependency ratio at destination country $j$ during year $t$
$inttot_{i,t}$	Magnitude score of episode of international warfare involving origin country $i$ during year $t$
$civtot_{i,t}$	Magnitude score of episode of civil warfare involving origin country $i$ during year $t$
$disaster_{i,t}$	Natural disaster event in origin country $i$ during year $t$
$disaster_{j,t}$	Natural disaster event in origin country $j$ during year $t$
$Stock_{i,j,d}$	Migrant stock from $i$ living in $j$ at the beginning of decade $d$
$distw_{i,j}$	Population weighted distance between origin country $i$ and destination country $j$
$comlang\_of_{i,j,t}$	Common official language between origin country $i$ and destination country $j$ during year $t$
$comrelig_{i,j,t}$	Share of people with the same religion between origin country $i$ and destination country $j$ during year $t$
$droughts_{i,t}$	Drought event at origin country $i$ during year $t$
$droughts_{j,t}$	Drought event at destination country $j$ during year $t$
$T_{i,j,t}$	Migration flow from country $i$ to country $j$ during year $t$

Table 1: The input features used for the model. Each feature spans from 1960 to 2010 for a pair of origin-destination country.

internal migration across US counties and, closer related to our analysis, changes in bilateral international migrant stocks over decades. In addition, based on the trained XGboost model, they identify the most important features affecting migration, ordering them by their percentage of importance. By doing so, one can identify the drivers that mostly influence the model’s predictions. Unfortunately, this method does not allow to understand *how* these features influence migration flows. Further, the low frequency data on international migration used makes it difficult to derive insights that are policy relevant in the short-run. Estimates may fail to account for the timing of events causing *direct* migratory responses, such as natural disasters or conflicts, and miss a potentially important share of temporary migrants.

**Our approach.** As in Robinson and Dilkina [2017], we want to predict  $\hat{T}_{i,j}$  from a set of features. However, here, we propose a ML model that is able to estimate international human mobility on an annual basis.

To capture the complexity behind international migration dynamics, we add a set of different features of sending and receiving countries (conflict, disaster, etc.) and importantly account for the existence of initial migrant networks.

We analyze the covariates (features) in detail with interpretable ML techniques that go beyond the typical “black box” estimates. Our approach is based on partial dependence plots, allowing us to gain additional insights regarding international migration patterns, especially in the presence of strongly non-linear effects.

**Data and features set.** Table 1 describes the features used to train and test our model. The dataset used is a combination of data from several sources. We combine data from the Determinants of International Migration C2C database (DEMIG) of historic migration flows [Vezzoli *et al.*, 2014],

with the more recent International Migration Data database (IMD) [Stat, 2009]. To the best of our knowledge, this leads to the most comprehensive data set on annual bilateral migration flows taken to analyse in this context, covering flows from 175 origins to 33 mainly OECD destinations over the years 1960 to 2010<sup>1</sup>. Both data sets are based on population registers and residence permits, and are strongly coherent with a Pearson correlation coefficient of over 0.99 for the observations reported in both sources. We draw on the IMD data as primary source and append missing observations with DEMIG data.

The features used as independent variables in the model can be split in two main categories: (a) country specific indicators such as population and population density, GDP per capita, dependency ratio<sup>2</sup>, internal and international conflicts, natural disasters (earthquakes, storms, floods, volcanic eruptions) and droughts<sup>3</sup> for both origin and destination countries at each time step; and (b) bilateral indicators containing distances between origin and destination at time, e.g. geographical distance, common language and common religion. Finally, migrant networks have been shown to be important drivers of subsequent migration flows [Beine *et al.*, 2009]. We account for the presence of initial migrant networks by adding the stock of migrants born in  $i$  and living in  $j$  at the beginning of each decade to our feature set. The corresponding data stems from the Global Bilateral Migration Database

<sup>1</sup>We limit the analysis to the period 1960-2010 to avoid the turmoil related to the recent refugee crisis that took off with the beginning of the Syrian civil war in 2011.

<sup>2</sup>Dependency ratios are defined as the ratio between young (< 15) and old (> 65) to the working age population.

<sup>3</sup>We use gridded temperature and rainfall data to calculate SPEI-based relevant yearly drought indicators by country [Beguería *et al.*, 2014]

Models	CPC		MAE		RMSE		$r^2$	
	train	test	train	test	train	test	train	test
PPML	0.537	0.591	1157	819	8211	4230	0.237	0.481
ANN	<b>0.776</b>	<b>0.617</b>	<b>545</b>	<b>724</b>	<b>5434</b>	<b>4088</b>	<b>0.670</b>	<b>0.520</b>

Table 2: Comparison of the 2 models for the specified metrics. The values shown are by pair (train - test).

(GBMD) [Özden, 2011].

### 3 Our ML model

We use an ANN to estimate the annual migration flows. Our ANN is composed of 3 densely connected layers with rectified linear unit (ReLU) activation layers, which allows us to catch complex, non-linear interactions between the features and flows. The output layer uses the sigmoid activation function for the estimation. We use the same model for all predictions  $\hat{T}_{i,j}$ . The ANN receives then at each time step the set of features as described in the Table 1 and returns the forecast migration flow  $\hat{T}_{i,j,t+1}$ .

We evaluate the predictive power of the model with several commonly used metrics: the Mean Absolute Error (*MAE*), the Root Mean Square Error (*RMSE*), the Coefficient of determination ( $r^2$ ) and the Common Part of Commuters (*CPC*) also used in [Robinson and Dilkina, 2017].

**Learning process.** To train our model we use three sets. A training set, a validation set, and a test set [Goodfellow *et al.*, 2016]. The training and validation sets consist of the features and flows from 1960 to 2000 and the test set consists of the features and flows for the period from 2001 to 2010. We perform regularization in the model using dropout layers to prevent overfitting and ensure a better generalization [Srivastava *et al.*, 2014]. From the experiments optimal values for the hyperparameters are: number of hidden layer - 3, number of epoch - 200, dropout - 0.10, batch size - 32, loss function - *CPC*, found using the *RMSProp* optimizer [Tieleman and Hinton, 2012].

**Partial Dependence Plots.** The partial dependence plots (PDP) [Hastie *et al.*, 2009; Molnar, 2018; Zhao and Hastie, 2019] attempt to better understand the nature of dependence of the approximation  $f(\text{features})$  on their joint values. The graphical renderings of the partial dependence provide then the marginal effect of one or two features on the target outcome of a machine learning model. Consider the  $X_S$  of  $l < p$  of the input predictor variables  $\text{features} = (X_1, X_2, \dots, X_p)$ , indexed by  $S \subset \{1, 2, \dots, p\}$ . Let  $C$  be the complement set, with  $S \cup C = \{1, 2, \dots, p\}$ . Partial dependence functions for the subset  $S$  can be estimated by the following equation:

$$\hat{f}_S(x_S) = \frac{1}{n} \sum_{i=1}^n \hat{f}(x_S, x_C^{(i)}), \quad (1)$$

where  $x_S$  are the values of  $X_S$  in the training set,  $\{x_C^{(1)}, x_C^{(2)}, \dots, x_C^{(n)}\}$  are the values of  $X_C$  occurring in the training data and  $n$  is the training data size.

## 4 Experimental Results

The first step of our experiments aims to guarantee the effectiveness of our ML model. To do so, we compare it to a traditional gravity model estimated by a Poisson Pseudo Maximum-Likelihood approach (PPML) [Beine *et al.*, 2016]. As mentioned above, the test set contains all the migration flows from 2001 to 2010 representing about 31% of our full sample of data. Table 2 shows the results of the comparison. One can see that the ANN model performs better than the gravity model on all metrics for both the training and the test sets. Using the same information, our ANN model is thus more reliable on the task of predicting international migrant flows.

Next, we analyze how the different features influence the flows. We wish to focus our attention on the most important features. While ideally we would use the ANN to determine which features are important, doing so is challenging, due to the black box nature of ANNs. We decided to use a practical approach in this work in which Random Forests (RFs) are used to filter the initial set of features before analysing these features using more complex approaches. While the performance of this RF is not as good as that of the ANN, this allows to derive some measure of feature importance more easily [Breiman, 2001]. Figure 1a shows the most important features extracted using a RF algorithm. As expected, we find that initial migrant stocks are the most important predictors of subsequent migration flows. Other important features are  $distw_{i,j,t}$ ,  $gdpc_{j,t}$ ,  $density_{j,t}$ ,  $pop_{i,t}$ ,  $pop_{j,t}$ , etc.

To understand the actual impact of single features in the ANN, we subsequently use PDPs; these allow for a better understanding of *how* features and outcome are related in the ANN itself. The figures from 1b to 1f present the PDP of selected features<sup>4</sup>. On each figure we represent estimated migration flows  $\hat{T}_{i,j,t}$  as a function of the feature we examined. The x-axis represents the natural logarithm of the features  $Stock_{i,j,t}$ ,  $distw_{i,j,t}$ ,  $gdpc_{i,t}$  and the absolute value for  $droughts_{i,t}$  and  $inttot_{i,t}$ . The y-axis is the flow value estimated by the partial dependence.

The PDP in Figure 1b illustrates how the initial migrant stocks relate to subsequent flows. International migration flows grow exponentially with the evolution of the migration stock in the destination country. This highlights the potential of migrant networks to foster migration, either through family reunification programs or through easier integration and enhanced flow of information [Beine *et al.*, 2009].

The second most important feature is  $distw_{i,j}$ , the population weighted distance between the two countries  $i$  and  $j$ . The associated PDP in Figure 1c shows that the farther away

<sup>4</sup>The full set of results is available upon request from the authors.

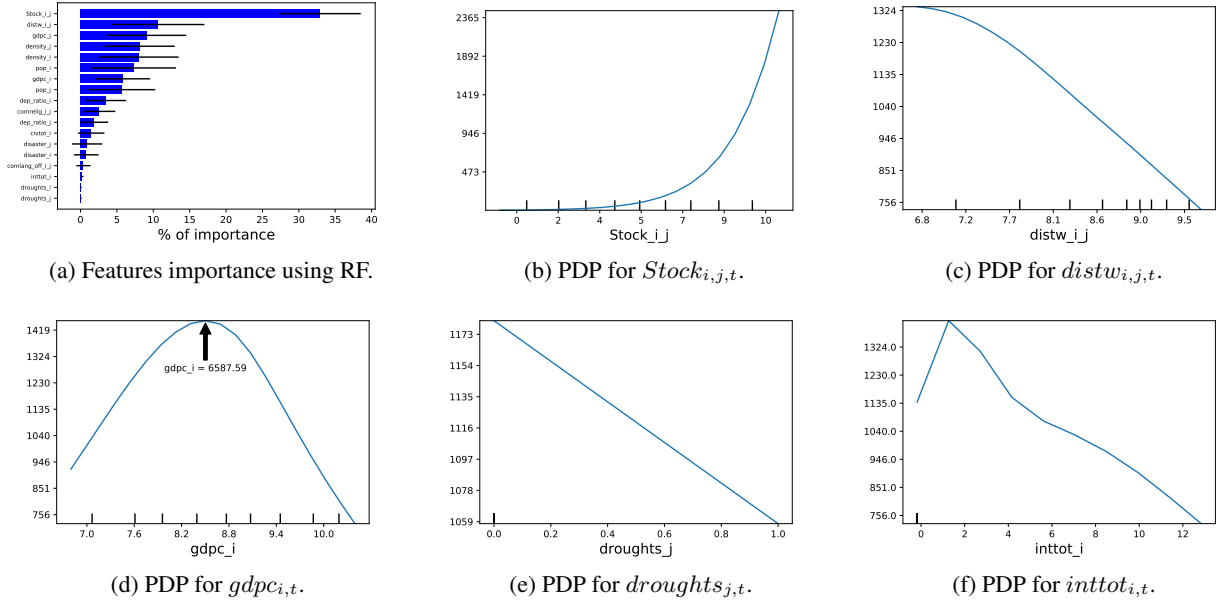


Figure 1: Features importance and Partial dependence plots for input features. On the top left we have the features importance plot from the Random Forest algorithm. The features values on the PDP are expressed using their natural logarithm values.

two countries are from each other, the lower the respective migrant flow. This could be indicative of the impact of increasing migration costs.

Next we look at the impact of droughts in destination countries on the flow of incoming migrants.  $droughts_{j,t}$  reflects the number of drought month during year  $t$  at destination country  $j$ <sup>5</sup>. Figure 1e reveals that destination countries tend to receive less migrants during drought years, assumably driven by worse employment possibilities especially for low skilled labour. This example emphasizes the added value of PDPs: while the overall feature importance is low, suggesting to dismiss droughts as migration drivers, the PDP indicates a significant effect in those countries that are concerned.

Figure 1d reveals a bell shaped relation between GDP per capita at origin and out-going migration flows. This finding is in line with the idea of a mobility transition across different development stages: while the majority of people in origin countries with lower GDP per Capita wants to migrate but faces binding financial constraints, the migration incentive decreases once income disparities between origin and destination countries vanish [Zelinsky, 1971; Dao *et al.*, 2018]. We find this turning point at a value of 6587 \$ in 2011 chained PPP, roughly reflecting the GDP of Kazakhstan in 2001. This illustrates the potential of PDPs to analyse complex, non-linear relationships between covariates and outcome variables.

In the same vein, Figure 1f delineates a non-linear link between conflicts and migration decisions.  $inttot_{i,t}$  reflects the

level of intensity of international conflict during year  $t$  for each origin country  $i$ . While low levels of conflicts can push people towards emigration, less and less people migrate when conflicts intensify and populations may be trapped.

## 5 Conclusion

In this paper, we have proposed an artificial neural network (ANN) model that is able to estimate annual flows better than a classic gravity model. With this trained model, we have applied an interpretability technique (partial dependence plots) which allows us to get deeper insights about how migration is influenced by its drivers. PDPs can reveal interesting, non-linear relationship between covariates and outcome variables. Hence machine learning is not only able to cope with the complex dynamics behind international migration flows, but can also be informative on *how* the flows change w.r.t. the covariates. As further work, we believe that it is possible to increase the interpretability by applying other techniques such as Shapley values or LIME with recurrent neural networks. In another direction, precise data of migration bilateral migration is still scarce. We believe that our proposed model can be useful to predict a broader set of migration flows, for example the important share of south-south migration which we ignored so far, and thus generate comprehensive migration matrices across time and space.

## References

- [Anderson, 2011] James E. Anderson. The gravity model. *Annu. Rev. Econ.*, 3(1):133–160, 2011.
- [Beguería *et al.*, 2014] Santiago Beguería, Sergio M Vicente-Serrano, Fergus Reig, and Borja Latorre. Standardized precipitation evapotranspiration index (spei)

<sup>5</sup>We use the SPEI drought indicator and consider a month as drought month when country aggregated SPEI is at least 1.5 standard deviations lower than its historic mean [Harari and Ferrara, 2018].

- revisited: parameter fitting, evapotranspiration models, tools, datasets and drought monitoring. *International Journal of Climatology*, 34(10):3001–3023, 2014.
- [Beine *et al.*, 2009] Michel Beine, Frédéric Docquier, and Caglar Ozden. *Diasporas*. The World Bank, 2009.
- [Beine *et al.*, 2016] Michel Beine, Simone Bertoli, and Jesús Fernández-Huertas Moraga. A practitioners’ guide to gravity models of international migration. *The World Economy*, 39(4):496–512, 2016.
- [Breiman, 2001] Leo Breiman. Random Forests. *Machine Learning*, 45(1):5–32, October 2001.
- [Dao *et al.*, 2018] Thu Hien Dao, Frédéric Docquier, Chris Parsons, and Giovanni Peri. Migration and development: Dissecting the anatomy of the mobility transition. *Journal of Development Economics*, 132:88–101, 2018.
- [DESA, 2019] UN DESA. World population prospects 2019: Highlights. *New York (US): United Nations Department for Economic and Social Affairs*, 2019.
- [Friedman, 2001] Jerome Friedman. Greedy Function Approximation: A Gradient Boosting Machine. *Annals of Statistics*, 29:1189–1232, October 2001.
- [Goodfellow *et al.*, 2016] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. The MIT Press, 2016.
- [Harari and Ferrara, 2018] Mariaflavia Harari and Eliana La Ferrara. Conflict, climate, and cells: a disaggregated analysis. *Review of Economics and Statistics*, 100(4):594–608, 2018.
- [Hastie *et al.*, 2009] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning*. Springer Series in Statistics. Springer New York, New York, NY, 2009.
- [Letouzé *et al.*, 2009] Emmanuel Letouzé, Mark Purser, Francisco Rodriguez, and Matthew Cummins. Revisiting the Migration-Development Nexus: A Gravity Model Approach. *Human Development Report Office (HDRO), United Nations Development Programme (UNDP), Human Development Research Papers (2009 to present)*, January 2009.
- [Masucci *et al.*, 2013] A Masucci, Joan Serras, Anders Johansson, and Michael Batty. Gravity versus radiation models: On the importance of scale and heterogeneity in commuting flows. *Physical review. E, Statistical, nonlinear, and soft matter physics*, 88:022812, August 2013.
- [Molnar, 2018] Christoph Molnar. *Interpretable Machine Learning*. Leanpub, February 2018.
- [Özden, 2011] Maurice Walmsley Terrie L. Çağlar Parsons Özden, Christopher R. Schiff. Where on Earth is Everybody? *World Bank Economic Review*, 25(1):12–56, January 2011.
- [Robinson and Dilkina, 2017] Caleb Robinson and Bistra Dilkina. A Machine Learning Approach to Modeling Human Migration. *arXiv:1711.05462 [physics]*, November 2017.
- [Simini *et al.*, 2012] Filippo Simini, Marta C. González, Amos Maritan, and Albert-László Barabási. A universal model for mobility and migration patterns. *Nature*, 484(7392):96–100, April 2012.
- [Srivastava *et al.*, 2014] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. page 30, 2014.
- [Stat, 2009] OECD Stat. International migration database, 2009.
- [Tieleman and Hinton, 2012] Tijmen Tieleman and Geoffrey Hinton. Rmsprop: Divide the gradient by a running average of its recent magnitude. *COURSERA: Neural networks for machine learning*, 4(2):26–31, 2012.
- [Vezzoli *et al.*, 2014] Simona Vezzoli, Maria Villares-Varela, and Hein De Haas. Uncovering international migration flow data: Insights from the demig databases. 2014.
- [Zelinsky, 1971] Wilbur Zelinsky. The hypothesis of the mobility transition. *Geographical review*, pages 219–249, 1971.
- [Zhao and Hastie, 2019] Qingyuan Zhao and Trevor Hastie. Causal Interpretations of Black-Box Models. *Journal of Business & Economic Statistics*, pages 1–10, July 2019.
- [Zipf, 1946] George Kingsley Zipf. The P1 P2/D Hypothesis: On the Intercity Movement of Persons. *American Sociological Review*, 11(6):677–686, 1946.