MODELING AND SIMULATION OF FORCED POPULATION DISPLACEMENT FLOWS USING COLORED PETRI NETS

Jenaro Nosedal-Sanchez[(a)], Ali Asgary[(b)], Adriano O. Solis[(c)], Francesco Longo[(d)]

[(a),(b),(c)] School of Administrative Studies, York University, Toronto, Ontario, Canada
[(d)]DIMEG, University of Calabria, Rende (CS), Italy

[(a)]jnosedal@yorku.ca, [(b)]asgary@yorku.ca, [(c)]asolis@yorku.ca, [(d)]francesco.longo@unical.it

ABSTRACT
In this paper, we present a discrete event model, developed and implemented using colored Petri nets. The proposed model uses empirical data distributions to simulate the forced population displacement flows observed during the Mosul military operations in 2016-2017. The simulation model allows estimation of the magnitude and destination of the flows within identified routes (between districts in Iraq) and the simulation results appear to reproduce the aggregate displacements actually experienced.

Keywords: colored Petri nets, stochastic modeling, discrete event simulation, forced population displacement, Mosul

1. INTRODUCTION

An internally displaced person (IDP) is someone who has moved within the bounds of his or her own country, as a result of threats due to natural or man-made events such as earthquakes, floods, pandemics or war.
In June 2014, the Islamic State of Iraq and the Levant (ISIL), also known as the Islamic State of Iraq and Syria (ISIS), invaded and took control over the city of Mosul (Figure 1) in Northern Iraq. Since then, the city had become a major hub for terrorist activities and subsequently became ISIS' de facto capital in Iraq (Reddy 2017). Mosul's population, among the most diverse in Iraq, includes substantial populations of Sunni Arabs, Assyrian Christians, Kurds, Yazidis, Turkmen, and other minorities.
After the invasion of Mosul by ISIL/ISIS in 2014, millions of Iraqi people have been displaced. In the case of Mosul and adjacent areas, a large fraction of the displaced population have decided to flee from their homes to other areas inside the country (including other areas within Mosul).
Iraqi armed forces started the military operation to retake the city of Mosul on October 16, 2016 (BBC.com 2016; CNN.com 2016). At that time, there were about 1.5 million people living in the city.
The Displacement Tracking Matrix (DTM) emergency tracking (ET) system for Iraq is used by the International Organization for Migration (IOM, formerly the UNHCR), to track "sudden displacement or return movements triggered by [the Mosul crisis]" (http://iraqdtm.iom.int/EmergencyTracking.aspx).

Figure 1: City of Mosul, Northern Iraq *(encircled in red)*

2. DATA SET AND ANALYSIS

Based on the information available from the ET system, it is possible to track the evolution of the displacements in terms of number of people displaced on a specific date and the accumulated number of IDPs.
Figure 2 provides a view of the displacements observed over the time interval from October 27, 2016 to May 04, 2017 (observations are reported every 3-4 days on average, twice a week). A quick glance at the slope of the *SUM* function seems to provide a clear and predictable pattern of accumulated IDPs. However, by checking exclusively the accumulated IDPs, the magnitudes of the variations are diminished. Measuring the variations ($IDP_t - IDP_{t+1}$) provides a view of the fluctuations between successive observations (number of IDPs from date/observation t to $t+1$), represented as N.
These fluctuations in population displacements over time are triggered by specific circumstances, such as changes in the current situation in terms of security and

other social, political and demographic indicators (Davenport, Moore, and Poe 2003; Shellman and Stewart 2007; Asgary, Solis, Longo, Nosedal, Curinga, and Alessio 2016). In certain situations, the identification of appropriate explanatory variables and development of causal models may be possible (e.g., Collmann, Blake, Bridgeland, Kinne, Yossinger, Dillon, and Zou 2016; Engel and Ibáñez 2007). However, as in the current case, there is an inherent complexity in identifying relevant explanatory variables and having access to their values over sufficiently long time periods.

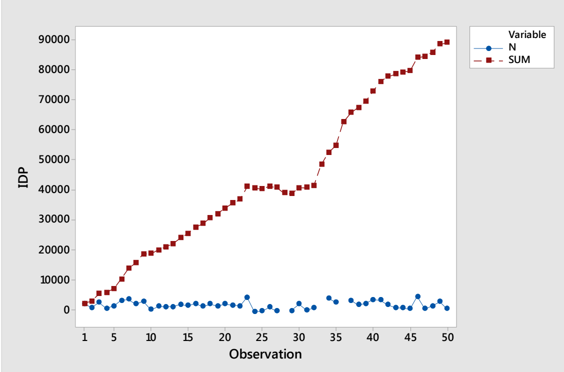

Figure 2: Number of Persons Displaced (per Observation and Cumulative)

In this paper, rather than focusing on the causal approach, the model is developed using a stochastic approach, in order to reproduce the evolution and uncertainty of displacements over time.
Figure 3 shows the distribution of values of *N*, while Figure 4 shows the cumulative distribution function (CDF). Besides the variability observed for *N*, in this case it is also critical to determine the direction of the displacement described by the origin and destination of the individual that decides to flee (in terms of routes taken).

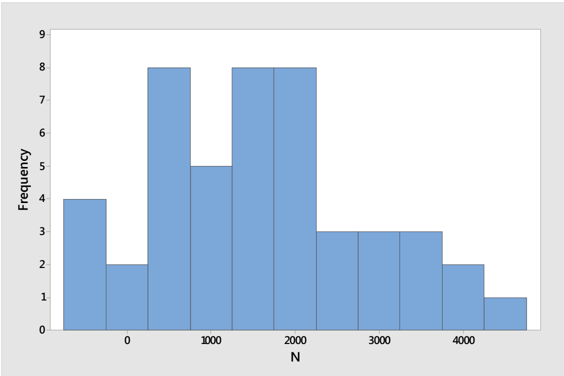

Figure 3: Distribution of Values of *N*

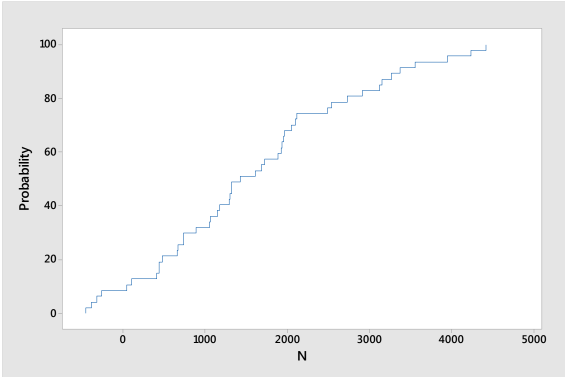

Figure 4: CDF for *N*

From the available data, a set of 52 routes of displacement – both out of and into Mosul – are observed, including the cases of displacement within Mosul itself (denoted as R37). However, the vast majority (a total of 86%) of the displacements are associated with only three of the 52 routes, while 14% fall into the 49 other routes (see Figure 5).

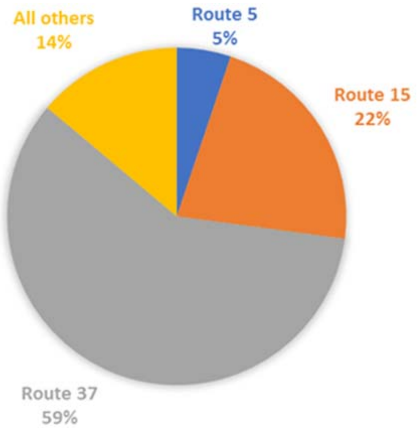

Figure 5: Distribution of IDP Routes

Table 1 summarizes the three principal routes taken by families, while the 49 others are lumped into 'alternative routes' (R_Alt). However, we point out that a more detailed analysis is required of the variation of IDP flows over time.

Table 1: Main Routes Taken by IDPs

Routes	Origin	Destination
R37	Mosul	Mosul
R15	Mosul	Al-Hamdaniya
R5	Hatra	Mosul
R_Alt	Various	Various

Instead of trying to build a regression model for the aggregate flows, in this case, based on observed data the empirical distributions for values of interest are obtained (number of people displaced, percentage of IDPs taking each route).
The empirical distributions of the percentage of IDPs observed for each of the listed routes are presented in

ISBN 978-88-97999-90-4; Bruzzone, Cayirci and Sottilare Eds.

Figure 6. In general, these empirical distributions do not fit with theoretical distributions. These empirical distributions will be implemented as inputs in the simulation model (explained in the next section).

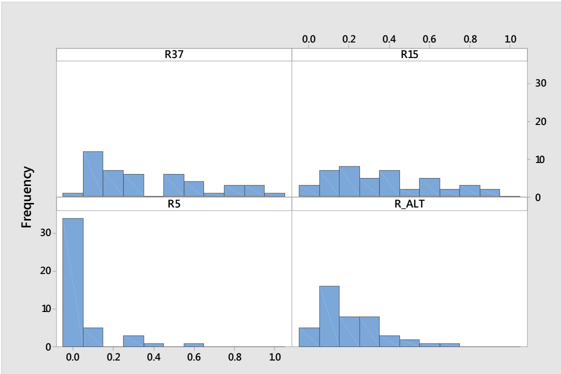

Figure 6: Empirical Distribution of the Percentage of IDPs taking Main Routes

Additionally, temporal fluctuations are illustrated by means of time series for each route (Figure 7). These values exhibit the variation over time without showing any clear trend or tendency.
Finally, to show the probability to observe a specific percentage of displacement for each case the CDF for each of the main routes is presented in Figure 8.

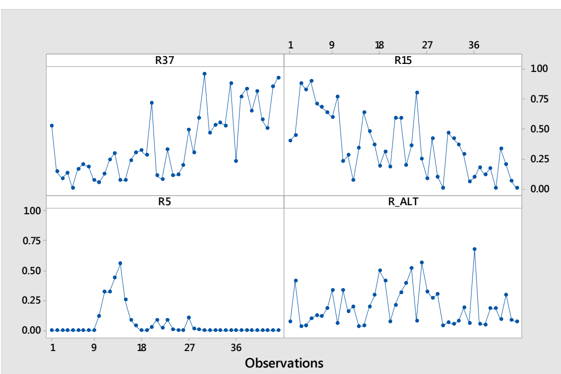

Figure 7: Time Series for Main Routes

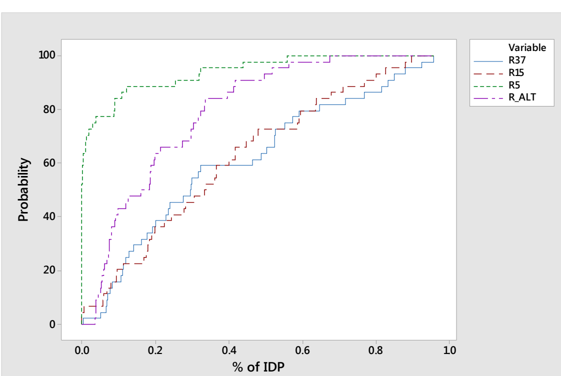

Figure 8: CDFs for Main Routes

3. MODEL DESCRIPTION AND SIMULATION RESULTS

3.1. Model Description

Petri nets (PNs) are an extension of state graphs which provide a compact and effective representation of real systems (Jiménez and Pérez 2004).
The graphical representation by directed bipartite schemas and the mathematical expression by state matrix or state equation support their wide application for modeling, simulating and analyzing the evolution of discrete event systems (Nosedal, Baruwa, and Piera 2013).
This paper presents the application of colored Petri nets (CPNs), which are an enriched type of PN. CPNs allow a higher level of modeling by using colors that enable the representation of entity attributes (Pennisi, Cavalieri, Motta, and Pappalardo 2016).
A full description of the CPN modeling approach and its applications can be found in Jensen, Kristensen, and Wells (2007) and the description of a graphical language for its representation and simulation in Jensen and Kristensen (2015).
In this section, our model is introduced by first listing and providing descriptions of its main elements: colors (Table 2), transitions (Table 3), and places (Table 4). The model (see Figure 9) has been implemented using CPN Tools (Jensen and Kristensen 2015).

Table 2: Colors

Color	Description
K37	Value for % of flow in route 37
K15	Value for % of flow in route 15
K5	Value for % of flow in route 5
Ka	Value for % of flow in the remaining routes
D	Current available % of flow to be assigned
N	Value of the overall flow magnitude simulated
c	Replications sequential number

Table 3: Transitions

Transition	Description
S_R37	Simulation of the flow in route 37
S_R15	Simulation of the flow in route 15
S_R5	Simulation of the flow in route 5
S_RALT	Simulation of the flow in the remaining routes (accumulate)
INTEGRA-TION	Standardization for the values to assure consistent % distribution of the flows and generation of simulation records.

Table 4: Places

Place	Description
ED_R37	Contains the values that describes the empirical frequency distribution for route 37
ED_R15	Contains the values that describes the empirical frequency distribution for route 15
ED_R5	Contains the values that describes the empirical frequency distribution for route 5
ED_RALT	Contains the values that describes the empirical frequency distribution for the 49 remaining routes (aggregate)
S_R37	Contains the simulated value for the flow using route 37
S_R15	Contains the simulated value for the flow using route 15
S_R5	Contains the simulated value for the flow using route 5
S_RALT	Contains the simulated value for the flow using the 49 remaining routes (aggregate)
PERCENT-AGE	Contains the memory for the current available % of flow to be assigned during the simulations
ED1	Contains the values that describes the empirical frequency distribution for the aggregate value of the flow (overall magnitude)
counter	Assigns a sequential number for each replication
S_RESULTS	Keeps the records of the simulated values (sequence, overall magnitude, % for route37, % for route 15, % for route 5, and % in the 49 remaining routes).

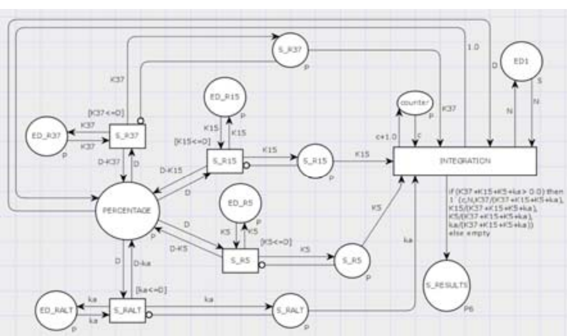

Figure 9: CPN Model for Simulation of Displacement Flows

3.2. Simulation Results

For aggregate IDP flows and flows along each of the four main routes as listed in Table 1, 30 replications of the model were run, with 50 time periods in each replication. For instance, in Figure 10, comparisons are provided concerning the time series for the aggregate magnitude the actual vs. simulated flows in the 11th replication. N_11 corresponds to the values for the entire simulated time frame of 50 periods, SUM_11 corresponds to the cumulative value of simulated flows, while N and SUM correspond to the real values (obtained from ET data set).

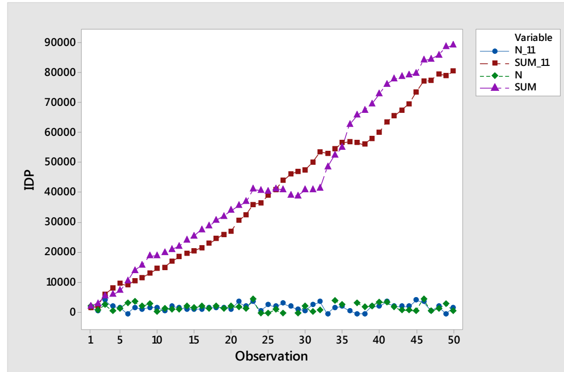

Figure 10: Actual vs. Simulated Flows (Replication 11)

Similarly, in Figure 11, comparison of the time series are shown for values obtained during the 26th replication.

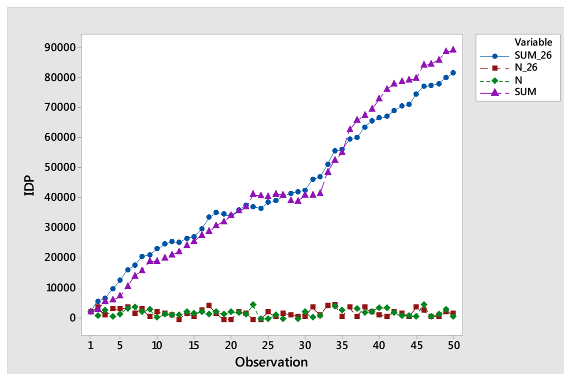

Figure 11: Actual vs. Simulated Flows (Replication 26)

The distributions of the values of *N* for each of the 30 replications (N_1 to N_30) are shown in Figure 12. The actual distribution is included in the top-left corner (indicated by N).

The distribution comparisons for the simulated values for percentage of flow within route R37 (IDP flows within Mosul) are presented in Figure 13.

Despite the fact that routes taken by IDPs follow a very different distribution (as illustrated in Figure 6), the actual and simulated distributions are very similar. Variations among simulations are expected due to the inherent uncertainty of the real data, and it is the intent of this model to reproduce this behaviour.

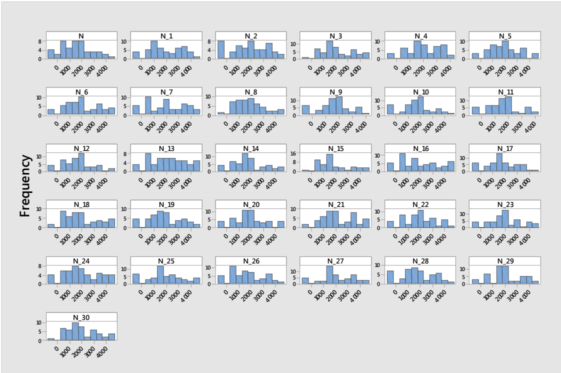

Figure 12: Distributions of Simulated Values of *N* (Replications 1,...,30)

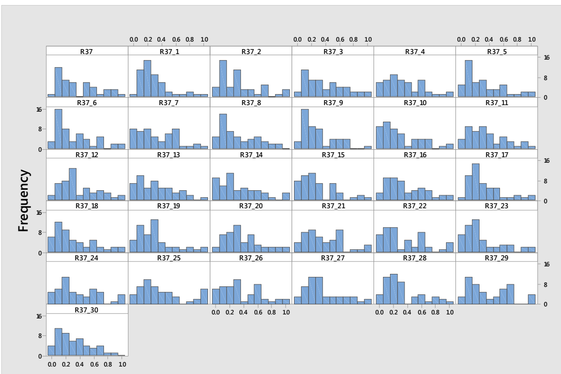

Figure 13: Distributions of Simulated Values of Percentage of IDPs taking Route 37 (Replications 1,...,30)

These simulated values and their stochasticity are illustrated and compared with the actual values by means of cumulative distribution diagrams (Figures 14 to 18).

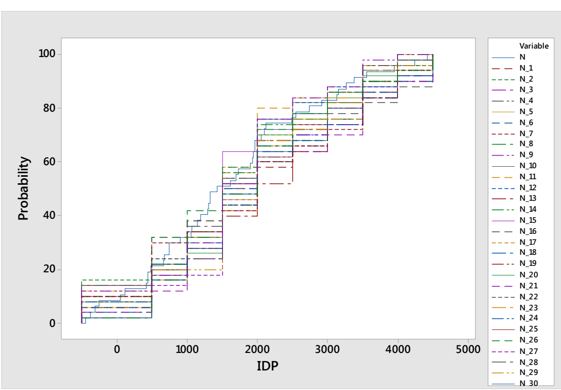

Figure 14: Cumulative Distribution Diagram for *N* – Actual vs. Simulated Values (Replications 1,...,30)

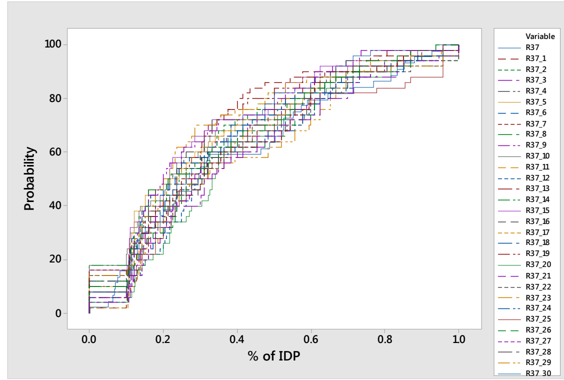

Figure 15: Cumulative Distribution Diagram for Percentage of IDPs taking Route 37 – Actual vs. Simulated Values (Replications 1,...,30)

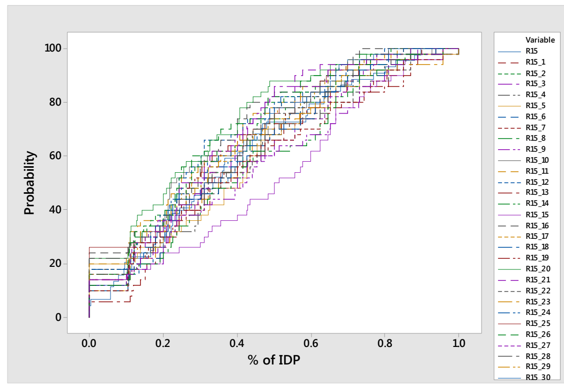

Figure 16: Cumulative Distribution Diagram for Percentage of IDPs taking Route 15 – Actual vs. Simulated Values (Replications 1,...,30)

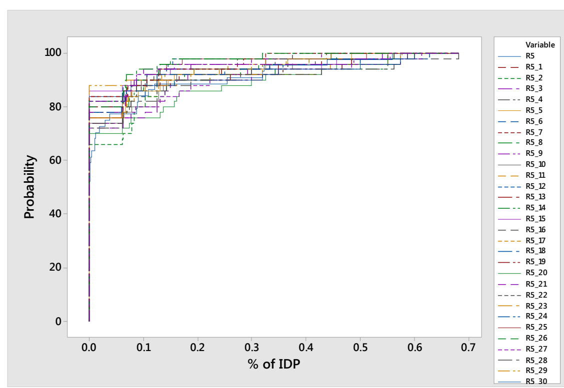

Figure 17: Cumulative Distribution Diagram for Percentage of IDPs taking Route 5 – Actual vs. Simulated Values (Replications 1,...,30)

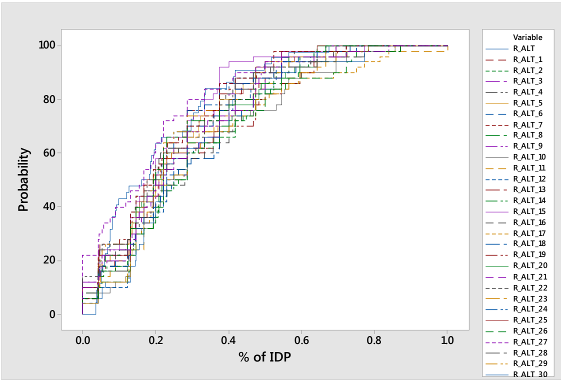

Figure 18: Cumulative Distribution Diagram for Percentage of IDPs taking Alternative Routes – Actual vs. Simulated Values (Replications 1,...,30)

In terms of the reproducibility of the temporal evolution of numbers of IDPs (actual time series were illustrated in Figure 7), Figures 19 and 20 provide examples of time series for the simulated values (in specific replications) for Route 5 and Alternative Routes, (both exhibiting fairly complex actual time series patterns).

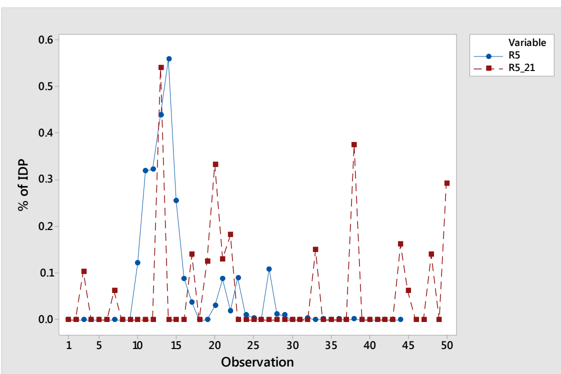

Figure 19: Actual vs. Simulated Values for Route 5 (Replication 21)

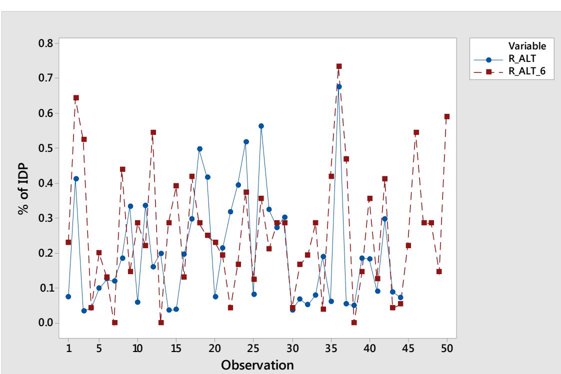

Figure 20: Actual vs. Simulated Values for Alternative Routes (Replication 6)

In both cases, despite the inherent variability and uncertainty of the observed values, our model appears to provide simulated time series that more or less follow the actual patterns as determined from the values reported in the ET system.

4. CONCLUSIONS AND FUTURE WORK

The current model involves stochastic modeling and simulation of forced population displacements based on discrete event modeling with the use of CPN.

The current model represents a single component that may be integrated into a more comprehensive model. The main components and structure for such a comprehensive framework, which integrates other elements and other simulation approaches (e.g., agent-based modeling), is reported and explained by Asgary, Solis, Longo, Nosedal, Curinga, and Alessio (2016).

Based on the initial results obtained, further work will include extending the CPN model to expand the set of possible routes. A further refinement would include updating the input data set (empirical distribution of IDP flows and the observed percentage distribution among routes). Since the military operation at Mosul has been completed as of July 2017 (CNN 2017), the input data may readily be updated.

ACKNOWLEDGMENTS

This research has been partially funded through the Collaborative Research and Training Experience (CREATE) Program in Advanced Disaster, Emergency and Rapid-response Simulation (ADERSIM) at York University, which is funded by the Natural Science and Engineering Research Council of Canada (NSERC). The research is also part of a Partnership Development Grant (Building Bridges across Social and Computational Sciences: Using Big Data to Inform Humanitarian Policy and Interventions) funded by the Social Science and Humanities Research Council of Canada (SSHRC). The Modeling and Simulation Center–Laboratory of Enterprise Solutions (MSC-LES) of the University of Calabria, Italy, is an international collaborator in the ADERSIM program.

REFERENCES

Asgary A., Solis A.O., Longo F., Nosedal J., Curinga M.C., and Alessio L.E., 2016. An agent-based modeling and simulation tool for estimation of forced population displacement flows in Iraq. Proceedings of the 6th International Defense and Homeland Security Simulation Workshop, pp. 75-81, September 26-28, Larnaca, Cyprus.

BBC.com, 2016. Mosul: Iraq and Kurdish troops make gains in battle (article dated 17 October 2016). http://www.bbc.com/news/world-middle-east-37679325 (accessed 16 April 2017).

Collmann J., Blake J., Bridgeland D., Kinne L., Yossinger N.S., Dillon R., and Zou K., 2016. Measuring the potential for mass displacement in menacing contexts. Journal of Refugee Studies, 29 (3), 273-294.

CNN.com, 2016. Battle for Mosul begins with gunfire and car bombs (article dated 17 October 2016). http://www.cnn.com/2016/10/17/middleeast/mosul

-isis-operation-begins-iraq/ (accessed 16 April 2017).
CNN.com, 2017. Mosul: Iraqi-led forces push into key city (article dated 10 July 2017). http://www.cnn.com/2016/10/17/world/gallery/mosul/index.html (accessed 19 July 2017).
Davenport C., Moore W., and Poe S., 2003. Sometimes you just have to leave: domestic threats and forced migration, 1964-1989. International Interactions, 29 (1), 27-55.
Engel S. and Ibáñez A.M., 2007. Displacement due to violence in Colombia: A household-level analysis. Economic Development and Cultural Change, 55 (2), 335-365.
google.ca, 2017. Map of Iraq and surrounding region. https://www.google.ca/maps/@34.5047751,45.6161276,6z (accessed 19 July 2017).
Jiménez E. and Pérez M., 2004. Simulation and optimization of logistic and production systems using discrete and continuous Petri nets. Simulation, 80 (3), 143-152.
Jensen K., Kristensen L.M., and Wells L., 2007. Coloured Petri nets and CPN Tools for modelling and validation of concurrent systems. International Journal on Software Tools for Technology Transfer, 9 (3-4), 213-254.
Jensen K. and Kristensen L.M., 2015. Colored Petri nets: a graphical language for formal modeling and validation of concurrent systems. Communications of the ACM, 58 (6), 61-70.
Nosedal J., Piera M.A., and Baruwa O., 2013. Concurrent and distributed systems analysis using colored Petri nets. Proceedings of XXXIV Jornadas de Automática, pp. 538-544, September 4-6, Terrassa, Spain.
Pennisi M., Cavalieri S., Motta S., and Pappalardo F., 2016. A methodological approach for using high-level Petri nets to model the immune system response. Proceedings of the 15th International Conference on Bioinformatics (InCoB 2016), pp. 91-105, September 21-23, Queenstown, Singapore.
Reddy P.A., 2017. Battle of Mosul: mass displacement of natives and a blatant violation of international humanitarian laws. LSE Human Rights Blog: http://blogs.lse.ac.uk/humanrights/2017/02/22/battle-of-mosul-mass-displacement-of-natives-and-a-blatant-violation-of-international-humanitarian-laws/ (accessed 19 July 2017).
Shellman S.M. and Stewart B.M., 2007. Predicting risk factors associated with forced migration: an early warning model of Haitian flight. Civil Wars, 9 (2), 174-199.