Evaluation and extension of the radiation model for internal migration - Supplemental Material

Lucas Kluge

Potsdam Institute for Climate Impact Research (PIK), Member of the
Leibniz Association, PO Box 60 12 03,D-14412, Potsdam, Germany and
Institute of Physics and Astronomy, University of Potsdam, Karl-Liebknecht-Str. 24/25, 14476, Potsdam, Germany[*]

Jacob Schewe

Potsdam Institute for Climate Impact Research (PIK), Member of the
Leibniz Association, PO Box 60 12 03,D-14412, Potsdam, Germany
(Dated: June 11, 2021)

I. NORMALIZED GRAVITY MODEL

In the following we show the results for a Gravity model constrained to produce the same number of migrants as given by the census data. To calibrate the model we use the same method as for the unconstrained model described in the paper. At each iteration during the calibration, before evaluating the performance for a set of parameters, we normalize the sum of all flows to fit the total number of reported migrants. Due to the normalization, the scaling coefficient A in the original Gravity model is not fitted, but determined directly as:

$$\tilde{A} = \frac{\sum_{i,j} M_{i,j}^{\text{cens}}}{\sum_{i,j} \frac{m_i^a m_j^b}{d_{ij}^g}}, \tag{1}$$

with $M_{i,j}^{\text{cens}}$ being the census flow from origin i to des-

tination j. As in the main paper, m_i is the population of i, m_j the population of j, d_{ij} the distance between i and j, and a, b, g are fitting parameters.

The results of the normalized model can be seen in Fig. S1 for the US and in Fig. S2 for Mexico. The corresponding parameters are displayed in Tab. S1.

Type	Country	$\tilde{A}$	a	b	g
Inter	USA	5.80E-11	1.30	1.10	0.82
Intra	USA	2.30E-05	0.86	0.68	0.69
Inter	MEX	2.97E-08	1.22	0.75	0.69
Intra	MEX	2.18E-04	1.11	0.30	0.99

TABLE S1: Gravity model coefficients for the US and Mexico datasets with Normalization.
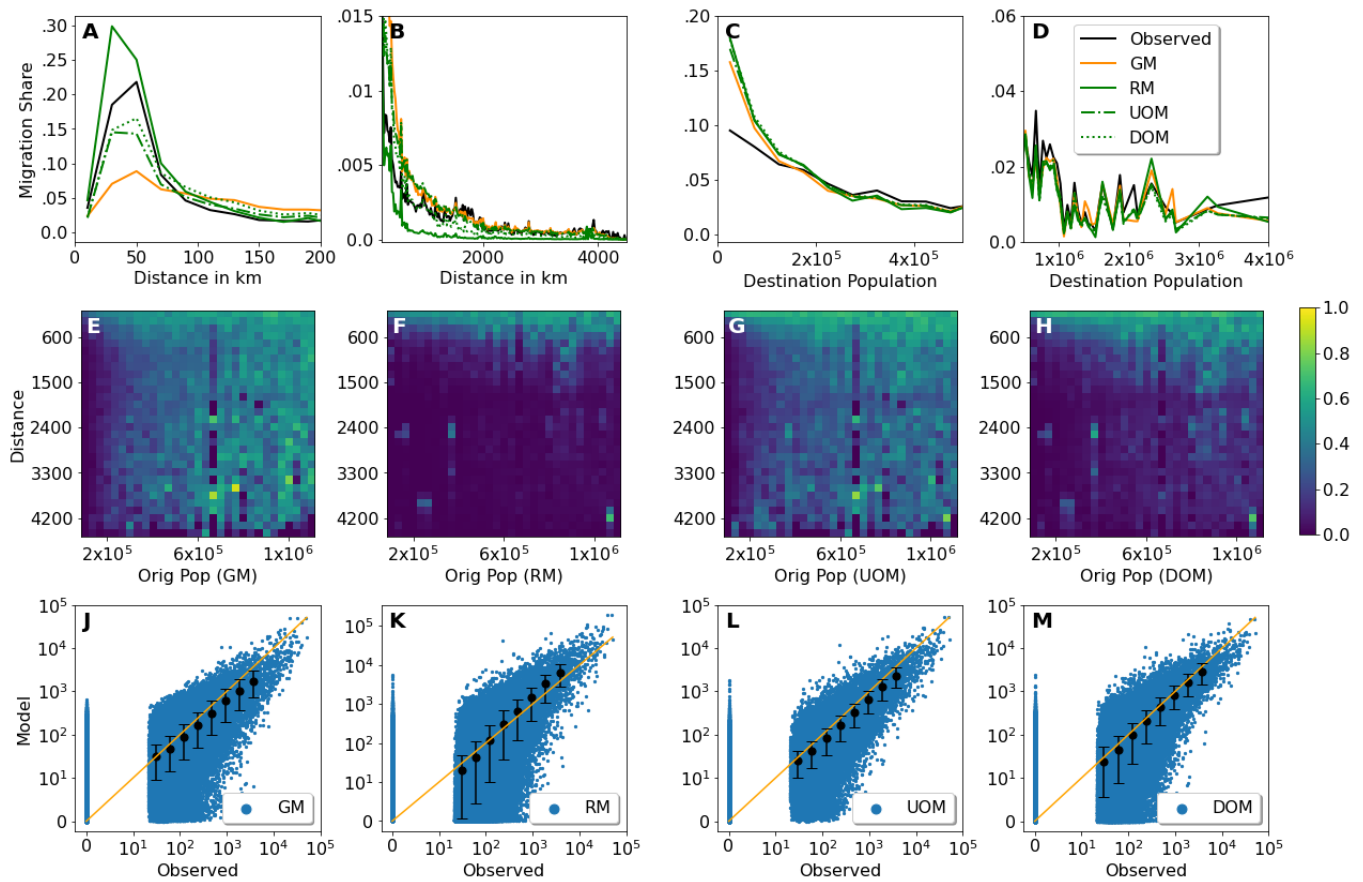
---

[*] kluge@pik-potsdam.de

FIG. S1: Observed (tax return based) and model data for internal migration in the US on a county level [1]. A and B show the distribution of migrant flows over travel distance (in 20km bins; y-axis indicates share of total migrant flow), for distances below (A) and above 200km (B). Analogously, C and D display the distribution of migrant flows over destination population size. E-H display the Sørensen-Dice coefficient for each model as a function of travelled distance (150km bins) and origin population size (33,333 inhabitants bins). J-M compare the census and model value of each individual flow. The orange line is the identity line. Black circles indicate the average, and error bars indicate the 25th and 75th percentile, across for a range of different size classes.

[1] IRS Migration Flows.   https://web.archive.org/web/20120415185331/http://www.irs.gov:80/taxstats/article/0,,id=212695,00.html.   Accessed: 2020-09-21.

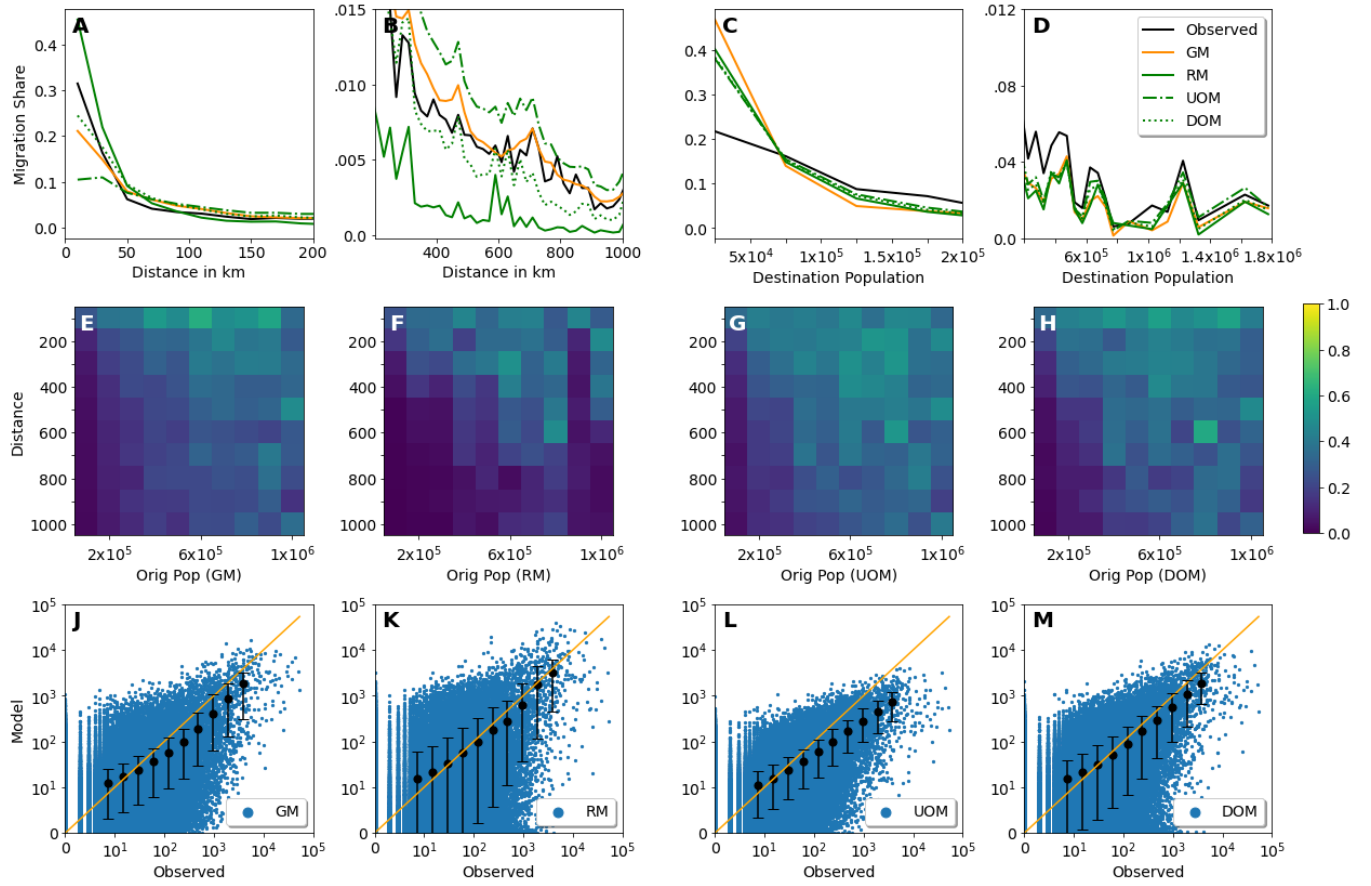[2] Mexico Population Census. http://en.www.inegi.org.mx/temas/estructura/. Accessed: 2020-06-05.

FIG. S2: As Fig. S1 but for Mexico, with observations based on microcensus data [2]. Binning distance in A and B is 20km. Sørensen-Dice coefficients in E–H are shown for 100km and 100,000 inhabitants bins.