Bayesian multiregional population forecasting: England

Arkadiusz Wiśniowski*[*], James Raymer**[†]
*University of Manchester
**Australian National University

14th December, 2015

Abstract

In this paper, we extend the well-known multiregional population projection model developed by Andrei Rogers and colleagues to be fully probabilistic. Multiregional models provide a general and flexible platform for modelling and analysing population change over time. They allow the combination of all the main components of population change by age with various transitions that population groups may experience throughout their life course. What distinguishes these models from ordinary projections is that they include transition matrices of interregional migration by age. This information is an important component of subnational population change yet models for forecasting the patterns for use in population projections are largely non-existent. National statistical offices tend to rely on simple deterministic assumptions regarding net migration or gross flows of in-migration and out-migration. These models do not take into account the linkages between origins and destinations and often have to be adjusted to ensure zero net migration and the same totals for in-migration and out-migration. In this paper, we focus on the full matrix of flows to avoid this problem. To deal with the large number of possible flows, we develop a Bayesian hierarchical model to forecast age-specific interregional migration in England, and then include this information with probabilistic forecasts of regional births, deaths, immigration and emigration. The results demonstrate the differences that arise from different models specifications and the promise of the general approach.

Key words: probabilistic population forecasts, Bayesian inference, multiregional projections

1 Introduction

This research substantially extends earlier efforts for multiregional estimation and forecasting population, namely Gullickson and Moen (2001), Sweeney and Konty (2002), Raymer et al. (2006), Wilson and Bell (2007), and (Bryant et al., 2013), by the inclusion of probabilistic information within the multiregional population projection model. Multiregional models include transition matrices of interregional migration by age, which is an important component of subnational population change yet probabilistic models for forecasting the patterns for use in population projections are largely non-existent. National statistical offices tend to rely on simple deterministic assumptions regarding net migration or gross flows of in-migration and out-migration. These models do not take into account the linkages between origins and destinations and often have to be adjusted to ensure zero net migration and the same totals for in-migration and out-migration.

This paper will be extended to include forecasts of births, deaths, internal and international migration for England. These forecasts will subsequently be combined within a multiregional population projection model in a similar fashion as in Wiśniowski et al. (2015). The current version of the model has been tested using the data from Italy as in Raymer et al. (2006). They represent interregional migration in the years 1970–1971 to 2000–2001 in five-year intervals, disaggregated by 20 five-year age groups for the Northwest, Northeast, Centre, South and Islands regions.

In our approach, we concentrate on the multiplicative component models such as those developed in, e.g., Raymer et al. (2006) and Smith et al. (2010), and bilinear forecasting models such as Lee and Carter method (Lee and Carter, 1992). Disaggregation of the flows into multiplicative components is useful for identifying the important underlying structures in the multi-dimensional data on migration.

[*]Contact email at A.Wisniowski@manchester.ac.uk
[†]Contact email at James.Raymer@anu.edu.au

2 Methods of forecasting interregional migration

The forecasting approach for interregional migration developed in this paper utilises Bayesian inference techniques, which treats all unknown parameters as random and having probability distributions. This allows us to incorporate uncertainty from the data and model parameters in the forecasts. Additionally, the Bayesian approach enables us to include in the model any informative and subjective information that can be elicited from the experts.

2.1 Statistical model

We assume that the time t, origin o, destination d and age a specific migration flows y_{odat} are Poisson distributed:

$$y_{odat} \sim \text{Poisson}\left(\mu_{odat}\right). \tag{1}$$

The logarithm of the mean μ_{odat} is then assumed to follow normal distribution with a mean being a bilinear model:

$$\log \mu_{odat} \sim \text{Normal}\left(\mathcal{M}, \tau_{odat}\right). \tag{2}$$

$\mathcal{M}$ represents a mean that includes a time trend for forecasting. In this paper, we propose three models that include the bilinear time effect:

$$\mathcal{M}_1 = \alpha_{oa} + \alpha_{da} + \alpha_{od} + \beta_{od}\kappa_t, \tag{3}$$

$$\mathcal{M}_2 = \alpha_{oa} + \alpha_{da} + \alpha_{od} + \beta_o\kappa_{1t} + \beta_d\kappa_{2t}, \tag{4}$$

$$\mathcal{M}_3 = \alpha_{oa} + \alpha_{da} + \alpha_{od} + \beta_{od}\kappa_{1t} + \beta_a\kappa_{2t}, \tag{5}$$

where we assume that

$$\alpha_{oa} \sim \text{Normal}\left(\alpha_a, \tau_{OA}\right), \tag{6}$$

$$\alpha_{da} \sim \text{Normal}\left(\alpha_a, \tau_{DA}\right), \tag{7}$$

$$\alpha_a \sim \text{Normal}\left(0, \tau_A\right). \tag{8}$$

In the first model, $\mathcal{M}_1$, the age profiles of migration are captured by the age specific term α_a, which is constant over time and across regions. The deviations from this age profile which are specific to each origin are then captured by the coefficient α_{oa}. Analogously, the destination-specific deviations of the age profile are reflected by α_{da}. Further, we introduce the two-way interaction term α_{od} for variation between each of the origins-destinations pairs. The coefficient β_{od} measures how fast this origin-destination interaction changes over time, in response to changes in time-specific effect κ_t.

Model $\mathcal{M}_2$ differs from $\mathcal{M}_1$ in the modelling origin and destination interactions with time. Instead of a two-way changes of the origin-destination interaction, we introduce origin-specific changes β_o and destination-specific changes β_d, with their own time-specific progression in time captured by κ_{1t} and κ_{2t}, respectively.

Model $\mathcal{M}_3$ is a direct extension of $\mathcal{M}_1$ that allows the age profile to change over time. The tempo of this change is captured by β_a and it corresponds to the time effect κ_{2t} in Eq. (5). This extension is similar to the decomposition of the age structure of mortality in the Lee and Carter (1992) model.

To forecast the migration flows, we assume that the time-effect parameters κ follow a random walk without drift:

$$\kappa_t \sim \text{Normal}\left(\kappa_{t-1}, \tau_T\right). \tag{9}$$

This implies that the expectation of the change of the given profile, whether it is origin, destination, origin and destination, or age, is the same as observed in the previous period. This specification ensures stability of the forecast levels of migration but also leads to increasing uncertainty over time. Other specifications, such as univariate and multivariate autoregressive processes are possible. In this application on the Italian data, the number of observations over time is seven. We believe that this series is too short for any meaningful inference with a larger number of parameters, and with the absence of any subjective information that could be fed into the model in form of the prior distributions. Finally, to ensure identification of the bilinear terms in the models (i.e. terms $\beta\kappa$), we assume that parameters representing the tempo of change for a given component β sum to one.

2.2 Prior distributions

For the coefficients capturing constant profiles by origin (α_o), destination (α_d), and both origin and destination (α_{od}), we assume normal prior distributions:

$$\alpha_x \sim \text{Normal}\left(0, \tau_x^{\alpha}\right), \tag{10}$$

where x denotes the interaction o, d or od.

For the tempo parameters β we assume a prior in form of a normal distribution, conditional on the constraint that all elements of the vector sum to one:

$$\beta_x^{(1:z-1)} \sim \text{Normal}_{(z-1)}\left(1/z, \tau_x^{\beta}\Psi\right), \quad \beta_x^{(z)} = 1 - \sum_{i=1}^{z-1} \beta_x^{(i)}, \tag{11}$$

where (z) denotes the last element of the given vector and precision matrix Ψ can be derived by using equations for conditional normal distributions and has the value of two on the diagonal elements and one on the off-diagonal elements.

We assume weakly informative prior distributions for the precision parameters in form of a left-truncated normal distributions, as suggested by Gelman (2006):

$$\tau \sim \text{Normal}\left(0, 10^{-3}\right) \mathbb{1}(\tau \geq 0), \tag{12}$$

where $\mathbb{1}(f)$ denotes an indicator function taking one if f holds and zero otherwise. These prior distributions are practically non-informative in a sense that they allow the data to drive the estimation process.

3 Results

The posterior distributions for model parameters and forecasts of migration were obtained from the MCMC simulations implemented in the OpenBUGS software (Lunn et al., 2009). The example results for models $\mathcal{M}_1$, $\mathcal{M}_2$ and $\mathcal{M}_3$ for the flow from the Northwest to the South are presented in Figure 1. The entire matrix containing observed migration flows for the year 2001 and probabilistic forecasts from Model $\mathcal{M}_3$ is presented in Figure 2.

The goodness of fit analysis reveals that the models are slightly over-fitting the data. The coverages of the data that are fed into the model are presented in Table 1. The over-fitting likely results from the fact that the data for which the fit is measured are used to compute the coverage.

Model	50% interval	80% interval	95% interval
$\mathcal{M}_1$	59.6%	84.5%	94.6%
$\mathcal{M}_2$	61.2%	83.4%	93.9%
$\mathcal{M}_3$	59.8%	86.1%	95.1%

Table 1: Coverage of the data by the posterior predictive distributions for the data

4 Summary

In this paper, we have developed a probabilistic approach for forecasting age-specific interregional migration. The approach focuses on the underlying structures found in multidimensional tables of migration flows. This work is still in development. Our initial analyses demonstrate the potential of this work. In the future, we plan to extend this work by applying the modelling approach to a complete annual time series of flows and to data obtained from other countries. We also plan to apply other model specifications and carry out in-sample forecasts to test their accuracy.

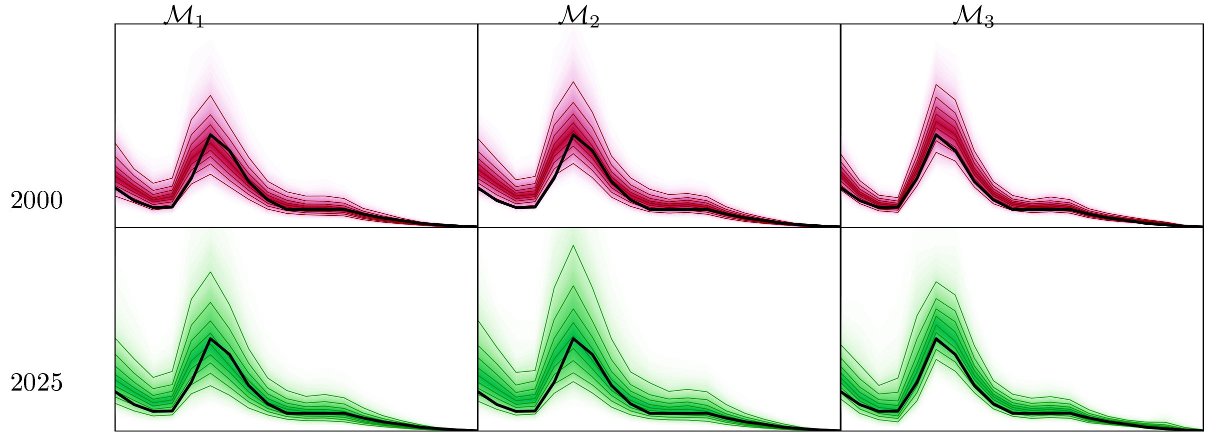

Figure 1: Internal migration flows from the Northwest to the South.
Note: The Y axis ranges from 0 to 7000 migrants; the X axis are 20 five-year age groups. The black curve is 2000 data, fit to these data in red, 2025 forecast is green; green lines denote the percentiles of the posterior distribution.

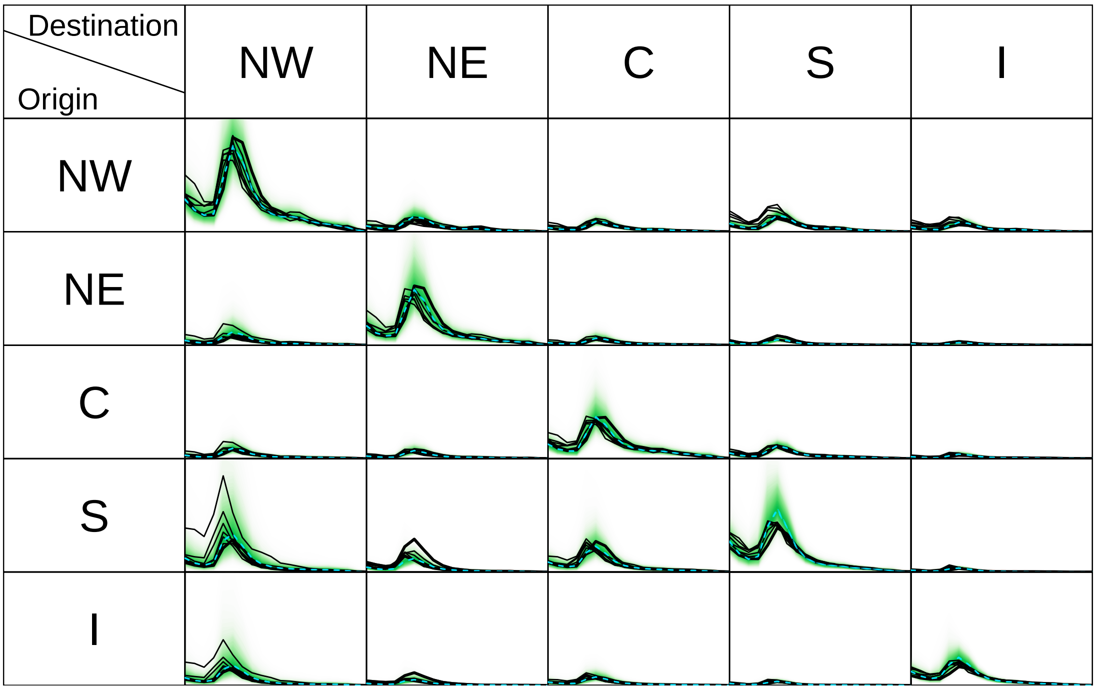

Figure 2: Internal migration matrix in Italy – Model $\mathcal{M}_3$.
Note: The Y axis on the diagonal ranges from 0 to 70000 migrants; off-diagonal plots have 0 to 25000. The X axis are 20 five-year age groups. 1970-2000 data (black), 2025 forecast (green) with medians (blue). NW denotes Northwest region, NE – Northeast, C – Centre, S – South, and I – Islands.

References

Bryant, J. R., Graham, P. J., et al. (2013). Bayesian demographic accounts: Subnational population estimation using multiple data sources. *Bayesian Analysis*, 8(3):591–622.

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by browne and draper). *Bayesian analysis*, 1(3):515–534.

Gullickson, A. and Moen, J. (2001). The use of stochastic methods in local area population forecasts. In *Annual Meeting of the Population Association of America, Washington DC*.

Lee, R. D. and Carter, L. R. (1992). Modeling and forecasting US mortality. *Journal of the American Statistical Association*, 87(419):659–671.

Lunn, D., Spiegelhalter, D., Thomas, A., and Best, N. (2009). The BUGS project: Evolution, critique and future directions. *Statistics in medicine*, 28(25):3049–3067.

Raymer, J., Bonaguidi, A., and Valentini, A. (2006). Describing and projecting the age and spatial structures of interregional migration in Italy. *Population, Space and Place*, 12(5):371–388.

Smith, P. W., Raymer, J., and Giulietti, C. (2010). Combining available migration data in england to study economic activity flows over time. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 173(4):733–753.

Sweeney, S. H. and Konty, K. J. (2002). Population forecasting with nonstationary multiregional growth matrices. *Geographical Analysis*, 34(4):289–312.

Wilson, T. and Bell, M. (2007). Probabilistic regional population forecasts: The example of Queensland, Australia. *Geographical Analysis*, 39(1):1–25.

Wiśniowski, A., Smith, P. W., Bijak, J., Raymer, J., and Forster, J. J. (2015). Bayesian population forecasting: Extending the lee-carter method. *Demography*, 52:1035–1059.