Chapter 2

Stochastic Population Forecasting: A Bayesian Approach Based on Evaluation by Experts

Rebecca Graziani

2.1 Introduction

Probabilistic population forecasting has recently received growing attention from researchers and, to a lesser extent, from official agencies, which traditionally derive population projections deterministically. As discussed in Keilman et al. (2002) and Keilman (2018), there are three main approaches to stochastic population forecasting. The first approach relies on the theory of time series, with models suggested both by the frequentist and the Bayesian approaches. The best-known time series approach in a classical framework is due to Lee and Carter (1992), originally proposed to forecast mortality, and later modified to address fertility forecasting, see Lee (1993) and Lee and Tuljapurkar (1994). Many extensions, generalizations and modifications have been proposed: see, among others, Booth et al. (2002), Booth and Tickle (2008), Booth (2006) Cairns et al. (2006, 2011), Hyndman and Ullah (2007), Hyndman and Booth (2008), and Hyndman et al. (2013). Using the Bayesian approach, Alkema et al. (2011) suggest a Bayesian hierarchical time series model for fertility forecasting, Raftery et al. (2013) for mortality forecasting, and Bijak and Wiśniowski (2010) and Bijak and Bryant (2016) for migration forecasting. As a sign that probabilistic approaches are entering

Electronic Supplementary Material The online version of this chapter (https://doi.org/10.1007/978-3-030-42472-5_2) contains supplementary material, which is available to authorized users.

R. Graziani (✉)

Department of Social and Political Sciences, Bocconi University, Milan, Italy

Dondena Centre for Research on Social Dynamics and Public Policy, Milan, Italy

Bocconi Institute for Data Science and Analytics, Milan, Italy

e-mail: rebecca.graziani@unibocconi.it

© The Author(s) 2020

S. Mazzuco, N. Keilman (eds.), *Developments in Demographic Forecasting*,
The Springer Series on Demographic Methods and Population Analysis 49,
https://doi.org/10.1007/978-3-030-42472-5_2

the mainstream of demographic forecasting, in 2014 the United Nations for the first time issued official probabilistic population projections for all countries up to 2100, see Alkema et al. (2015).

The second approach derives population forecasts based on the extrapolation of empirical errors. The observed errors from historical forecasts are used in the assessment of the uncertainty, see, among others, Stoto (1983). Following this approach, Alho and Spencer (1990) proposed the Scaled Model of Error, which was used to obtain stochastic population forecasts within the UPE (Uncertainty Population of Europe) project, see Alders et al. (2007).

The third approach, known as random scenarios or expert based approach, derives the forecast distribution of demographic components based on suitably elicited expert evaluations on their future trend, see, among others, Lutz et al. (1998). This is the approach that we follow in the present paper. The advantages and disadvantages of methods that rely on expert evaluations have been widely discussed in the literature. Goldstein (2004) and Lutz and Goldstein (2004), among others, stress how the random-scenario approach might be appealing to official agencies, due to its simplicity, the fact that its framework is based on scenarios, and the direct involvement of experts. The use of expert opinions allows taking into account behavioural theories on the future of the population (as argued by Lutz 2013) and allows incorporating in the forecasting exercise the knowledge of trends (such as policy changes and environmental changes) that might have an impact on the population dynamics. A further important advantage of expert based forecasting is that it does not require data on the past and therefore can be especially useful for developing countries, for which past data are usually poor. The main criticism of the expert based approach is related to the well-known and widely observed tendency of experts to underestimate the uncertainty. Keilman (1990) observed that, particularly when recent trends have been stable, the overconfidence of experts results in overly narrow prediction intervals. Among others, Alho and Spencer (1990) stress the conservativeness of expert opinions with respect to the decline in mortality, while Lee (1993) and Booth (2006) express concerns with respect to the accuracy of fertility forecasts. A further recognized drawback is that a forecast approach based on expert evaluations needs to focus on summary indicators of the demographic changes, and therefore turns out to be inflexible in forecasting age-schedules. Moreover, existing random-scenario methods, being generally based on trajectories that are obtained by the interpolation of a starting known and a final random value, are characterized by a variance and covariance structure which is not particularly flexible. Finally it is commonly emphasized that it is not easy to elicit from experts opinions on the across-time correlations for a single indicator and correlations between indicators.

Our method derives probabilistic population forecasts based on expert opinions in such a way as to take into account relations both between the demographic components and between the expert evaluations. As for the first kind of relation, between demographic components, there is a certain debate about the advisability and/or need to model such dependence. Indeed, if for some pairs of indicators the dependence is not under scrutiny, as for male and female life expectancies at

birth, in other cases, as for immigration and fertility, it is more questionable. It is common practice in population forecasting to assume independence between the three components of demographic change, fertility, mortality and migration, for which separate forecasts are provided. Our method lets expert evaluations on the future trends of the demographic components drive the detection of the presence and assessment of the strength of any relations between them. Indeed, in principle the method we suggest can take into account any type of dependence between any pair of indicators. Our method does not exclude independence between the demographic components, an independence that can be the result of expert opinions.

Our method also takes into account any dependencies between expert evaluations. Indeed, we expect that experts who have been trained and work in the same field would share a certain amount of knowledge and information, which could induce associations between their opinions. In an expert based approach, an important and delicate issue to face is how to combine the opinions provided by several experts. A wide literature is available on the problem of the aggregation of expert opinions, see, among others, Genest et al. (1986) for a review. Popular pooling methods suggest combining expert opinions by working out averages. For instance, the linear rule derives the collective assignment through their (possibly weighted) average. Similarly, one can define geometric or logarithmic pooling rules. Such pooling techniques take into account the variability of the expert evaluations, but do not take into account their potential associations, associations that we think cannot be neglected. Here we suggest a method for combining expert opinions that allows modelling both the associations between them and their diversity, taking into account several sources of uncertainty.

We suggest combining the expert evaluations by resorting to the so-called Supra-Bayesian method of pooling, introduced by Morris (1974) and then developed by many authors, see, among others, French (1980, 1981), Winkler (1981), Lindley (1983, 1985), Gelfand et al. (1995), and Roback and Givens (2001). In this approach, expert opinions on unknown quantities are treated as observations and combined based on the theoretical framework provided by the Bayesian approach to statistics. The analyst specifies a likelihood function, to be parametrized in terms of the unknown objects, and a prior distribution for the parameters. The posterior distribution, obtained by applying Bayes's theorem, updates the analyst's prior opinion, on the basis of the evaluations provided by the experts, and can then be used as a forecast distribution for the unknown quantities of interest. This approach takes into account and exploits the variability of the expert evaluations. Hence, the larger the number of experts, the more informative the procedure is.

In the next section, we provide a description of the method that was first suggested in Billari et al. (2014). We discuss in detail the elicitation procedure, the model, and the Markov Chain Monte Carlo algorithm. In Sect. 2.3 we describe the results of applying the model to forecasting of the Italian population from 2010 to 2065. In Sect. 2.4 some concluding remarks are provided.

2.2 The Supra-Bayesian Forecasting Method

As is common practice in the expert based approach, our method focusses on summary indicators of the three components of demographic change: fertility, mortality and migration. Population forecasts by age and sex are then obtained, relying on the commonly used cohort-component model, with age-schedules derived from the corresponding summary indicators, based on suitable models. In the following we describe the method, considering the case of two summary indicators R_1 and R_2 to be jointly forecast from time t_0 to time T . The inputs of the method are the expert opinions, which we presume to have been elicited according to the conditional elicitation procedure suggested in Billari et al. (2012).

The elicitation procedure works as follows. Split the forecast interval $[t_0, T]$ into two subintervals, considering a time point t_1 in it. In the first stage, the expert is asked to provide a forecast for each indicator at time t_1 and at time T , and an upper quantile for one of the two indicators at time t_1 , say for instance R_1 , as a value such that R_1 takes on a greater value with a predetermined probability. In the implementation of the method, this probability is set equal to 10%. In the second stage, the expert is asked to provide the following conditional forecasts:

- A forecast and an upper quantile at t_1 for the second indicator R_2 presuming that R_1 takes at t_1 a value equal to the elicited upper quantile and the forecast respectively;
- A forecast and an upper quantile at T for R_1 presuming that it takes at t_1 a value equal to the elicited upper quantile and the forecast respectively;
- Three different forecasts at T for R_2 presuming three different combinations of values for R_1 at t_1 and T and R_2 at t_1 .

In order to understand how the indicators' mean and variance, along with their correlations, can be derived from the elicited values, consider the case of one single indicator. In the case of the forecast of one single indicator, the expert should provide at the first stage forecasts for times t_1 and T , say m_{t_1} and m_T , and an upper quantile at time t_1 , say q_{t_1} as a value such that there is a probability equal to α that the indicator takes on a value greater than q_{t_1} . We assume Gaussian distributions for the indicator at the two time points, with means m_{t_1} , respectively, m_T . Under the Gaussian assumption, the variance $\sigma_{t_1}^2$ of the indicator at time t_1 can be easily derived from m_{t_1} and q_{t_1} as follows:

$$\sigma_{t_1}^2 = \left(\frac{q_{t_1} - m_{t_1}}{z_{1-\alpha}} \right)^2$$

with $z_{1-\alpha}$ being the quantile of order $1 - \alpha$ of a standard Gaussian random variable.

At the second stage, the expert is asked to provide a forecast, say $m_{T|t_1}$, of the indicator at time T presuming that it takes at time t_1 a value equal to the elicited quantile q_{t_1} and an upper quantile of the indicator at time T , say $q_{T|t_1}$, presuming that at time t_1 the indicator is m_{t_1} . Under the assumption of Gaussian distributions

at the two time points, the conditional distribution of the indicator at T given that it is equal to q_{t_1} at t_1 is Gaussian, has mean $m_{T|t_1}$ and variance $\sigma_{T|t_1}^2$ that can be derived as before from $m_{T|t_1}$ and $q_{T|t_1}$. The conditional distribution of the indicator is in this way completely specified, so that the correlation between the indicator at the two time points can be derived from standard results on Gaussian distributions.

This method can be easily generalized to the case of two indicators to be jointly forecast at t_1 and T . Therefore, the elicitation procedure allows indirectly eliciting across-time correlations for a single indicator, as the correlation between the rates at the two considered time points t_1 and T , and correlations both at the same time and across time for a pair of indicators, by asking for conditional forecasts.

This elicitation procedure yields vectors of forecasts of the two indicators at the two time points and their covariance matrix, one vector for each expert. In the method we suggest, the forecasts and the covariance matrices are used in a different way. We follow the Supra-Bayesian approach and suggest treating as data the forecasts provided by each expert at the two time points. In a Bayesian approach to inference, the analyst should, then, specify both the likelihood function, describing the random mechanism generating the evaluations and therefore to be parametrized in terms of the demographic summary indicators, and a prior distribution of these parameters, incorporating any information the expert has on them.

The likelihood function shapes the dependences between the expert evaluations. In Lindley (1983, 1985) a multivariate Gaussian distribution is used. Such a choice is motivated primarily by mathematical convenience, since it simplifies all computations related to the derivation of the posterior distributions. Nevertheless, the construction of a likelihood function of this kind is cumbersome, due to the large number of terms to be specified. Indeed, in the case of opinions elicited on several indicators at different time points, the choice of a multivariate Gaussian distribution requires the specification of all marginal means and variances and covariances.

Albert et al. (2012) suggest relying on a hierarchical random effects model, as a more parsimonious approach. At the beginning of the analysis, the experts are grouped by the analyst into a fixed number of homogeneity classes, corresponding to similar backgrounds or similar schools of thought. At the first level of the hierarchy, the opinions provided by the experts belonging to the same group are assumed to have the same distribution, indexed by parameters varying across groups. Then the different groups are assumed to have a common knowledge that is linked through a common distribution assigned to the group parameters and indexed by the parameter that represents the object of the expert evaluations. Finally, at the last level, a prior is assigned to this parameter, representing the overall uncertainty of the elicitation.

We suggest choosing a mixture model for the likelihood. Through this choice, we assume, as in Albert et al. (2012), that there are several different random mechanisms generating the expert evaluations, but we do not know which is the random mechanism generating the evaluations provided by each expert. Again, we presume that the experts can be grouped into a given number of classes, based on their shared knowledge and information, but for each expert we do not know which is the class the expert belongs to. We let the opinions provided by the

Likelihood
$x_i \mu_1, \dots, \mu_J, \Sigma_1, \dots, \Sigma_J, p_1, \dots, p_J \text{ ind} \sim \sum_{j=1}^J p_j N_4(\mu_j, \Sigma_j) \quad i = 1, \dots, K$
Priors
$\begin{aligned} \mu_j \Sigma_j \text{ ind} &\sim N_4(R, \frac{1}{k_0} \Sigma_j) \quad j = 1, \dots, J \\ \Sigma_j \text{ iid} &\sim IW(\Sigma_0, n_0) \quad i = 1, \dots, J \\ p_1, \dots, p_J &\sim Dir(\alpha_1, \dots, \alpha_J) \\ R &\sim N_4(\mu_R, \Sigma_R) \end{aligned}$
Fixed Hyperparameters
$J, \Sigma_0, k_0, n_0, \alpha, \mu_R, \Sigma_R$

Fig. 2.1 The mixture model

experts determine their group membership, so as to implicitly derive the dependence structure of the expert evaluations.

On the side of the prior distributions, as in Albert et al. (2012), the group centres are assumed to be independent and to have the same distribution, centred at the vector of summary indicators. In this way, we take into account the heterogeneity of the expert evaluations due to their possessing different pieces of information. Finally, we use the elicited covariance matrices to specify the prior distribution of the unknown clusters covariance matrix.

The resulting hierarchical model can be schematized as in Fig. 2.1, for the case of K experts where x_i is the vector of forecasts provided by expert i on the two indicators at two time points and $R = (R_{1t_1}, R_{1T}, R_{2t_1}, R_{2T})$ with R_{jt} being the random variable associated with indicator j at time t . The evaluations of the two summary indicators at the two time points are assumed to be conditionally independent and drawn from a mixture of J multivariate Gaussian distributions of dimension 4, each denoted by $N_4(\mu_j, \Sigma_j)$, for $j = 1, \dots, J$ and with J fixed by the analyst, being the number of groups of experts, with weights $p_1, \dots, p_J$. We assume in this way that each expert evaluation is distributed according to $N_4(\mu_j, \Sigma_j)$ with probability p_j . As for the prior distributions, the group means μ_j are assumed to be independent conditional on the covariance matrix Σ_j and distributed according to a multivariate Gaussian distribution centred at the vector of summary indicators at the two time points R , and with covariance matrix equal to Σ_j scaled by k_0 so as to end up with a diffuse prior, as discussed below. The covariance matrices Σ_j are assumed to be independent and identically distributed according to an inverse-Wishart distribution with scale matrix Σ_0 and n_0 degrees of freedom. The group probabilities $p_1, \dots, p_J$ are assumed to have a Dirichlet distribution with

parameters $(\alpha_1, \dots, \alpha_J)$. The vector of summary indicators at the two time points R is assumed to have a multivariate Gaussian distribution. It is worth emphasizing that this choice of prior distributions ensures conditional conjugacy (see, among others, Lavine and West 1992), which is something we draw on in the design of the Markov Chain Monte Carlo algorithm needed for the simulation of the posterior distribution of the vector of summary indicators R , described later.

The analyst needs then to specify the number J of components of the mixture and the parameters of the priors $\Sigma_0, k_0, n_0, \alpha, \mu_R, \Sigma_R$. The number of components can be chosen by fitting models with different J and then comparing them on the basis of indexes such as the Bayesian Information Criterion (BIC) or the Akaike Information Criterion (AIC). Since Σ_0 is the centre of the prior on the groups covariance matrix, we suggest specifying it based on the elicited covariance matrices. In our implementation of the model, we set Σ_0 equal to the arithmetic average of these covariance matrices, scaled so as to increase the variance of the elicited indicators. In this way we take into consideration and can correct the over-confidence of the experts, who tend to underestimate the variability of their forecasts. Since μ_R is the centre of the prior assigned to vector R , it represents a prior guess of the future values of the indicators and can then be specified using all available information. For instance, it can be fixed based on the central scenarios provided by national and international statistical agencies.

As for the remaining hyper-parameters, we suggest specifying them so as to end up with very diffuse priors. In this way, the posterior distribution can be mainly determined by the data, the expert elicited forecasts. Indeed, k_0 and n_0 affect the spread of the prior distributions on the group means and on the group covariances, respectively: the smaller they are, the larger is the spread. We suggest setting them as small as possible in order to increase the variability of the priors. Due to the properties of the Dirichlet distribution, the smaller is the value of α_j , the larger is the variability. Moreover α_j is the probability for an expert to belong to group j . A standard choice to depict no prior information on the group membership is $\alpha_j = \frac{1}{J}$. Σ_R is the covariance matrix of the prior distribution on R . We suggest choosing rather high variances so as to end up with a diffuse prior, and setting the covariances equal to 0, which corresponds to assuming the a priori independence of the indicators.

The joint posterior distribution of the indicators $(R_{1t_1}, R_{1T}, R_{2t_1}, R_{2T})$ can then be used as their forecast distribution at the two considered time points. Since this cannot be expressed in closed form, we suggest a Markov Chain Monte Carlo algorithm to draw samples from it. More precisely, we develop an auxiliary variables Gibbs-sampler, with full-conditionals that are all available in closed form due to the conditional conjugacy ensured by the choice of the prior distributions. For each observation, we introduce at each iteration an auxiliary variable Z_i taking values in $\{1, 2, \dots, J\}$, which flags its group membership and is updated each iteration. At each iteration of the algorithm, the group means and covariance matrices are updated by drawing them from a multivariate Gaussian distribution and an inverse-Wishart distribution respectively, the vector of latent variables is updated by drawing each component from a discrete distribution on $\{1, 2, \dots, J\}$, the vector of group

probabilities $(p_1, \dots, p_J)$ is updated by drawing it from a Dirichlet distribution and the vector R of summary indicators is updated by sampling it from a multivariate Gaussian distribution. The draws of the summary indicators from the joint posterior distribution are used as forecasts of the two summary indicators at the two time points, while forecasts for all points of the interval are obtained by resorting to suitable interpolation methods. In the application discussed in the next section, standard elementary quadratic interpolation techniques are used. As a by-product, the draws of the latent variables $Z_1, \dots, Z_K$ can be used for the estimation of the composition of the groups, that is, the clustering of experts in the J groups.

The Matlab package `supraBayesian_popproj`, downloadable from the web site of the publication, provides the codes implementing the Gibbs-sampler along with the codes for the derivation of the population forecasts by age and sex based on the simulations from the posterior distribution of the summary indicators.

2.3 An Application: Forecasting the Italian Population

In this section we illustrate an application of our forecasting method. The experts opinion used as inputs of the model were elicited according to the described procedure, through a questionnaire administered in 2012 in collaboration with the Italian Statistical Office (ISTAT). Experts were provided with information on the latest scenarios depicted by Eurostat and by the United Nations on the Italian summary indicators of demographic change. In 2015 the first official probabilistic population forecasts of the Italian population were issued by ISTAT starting from such elicited opinions. The Italian Statistical Office followed the method suggested in Billari et al. (2012) for the derivation of expert-based forecasts of the summary indicators. In the ISTAT forecasting exercise, the indicators were treated as independent and a multivariate Gaussian distribution was taken as the forecast distribution, with mean and covariance matrix obtained by averaging across the experts' elicitations. In 2017, ISTAT provided an update of the population projections of 2015, based on the same elicited opinions; a detailed description of the implemented methodology is provided in ISTAT (2017).

The forecasting period was 2010–2065 and was split into two sub-intervals, employing 2030 as the midpoint. The opinions were elicited on the following summary indicators: Total Fertility Rate, Mean Age at Birth, Male and Female Life Expectancies at Birth, Total Number of Immigrants and of Emigrants. The opinions on Total Fertility Rate and Total Number of Immigrants were jointly elicited, as were the opinions on Male and Female Life Expectancies at birth. Figure 2.2 displays the forecasts of the Total Fertility Rate and of the Total Number of Immigrants at 2030 and 2065 provided by 14 experts, while Fig. 2.3 depicts the corresponding correlations indirectly elicited.

With the Total Fertility Rate, there was low variability across expert evaluations: almost all the experts foresee a moderate increase in the rate from 2030 to 2065. With the Total Number of Immigrants, the evaluations show a higher variability,

Fig. 2.2 Expert forecasts of Total Fertility Rate and Total Number of Immigrants

especially for 2065; the majority of experts forecast a decrease in the Total Number of Immigrants. As to the correlations, there is a general agreement on a positive high correlation between Total Number of Immigrants at 2030 and Total Number of Immigrants at 2065 and on a positive moderate/high correlation between Total Fertility Rate at 2030 and at 2065. For the majority of experts there is a positive correlation between Total Number of Immigrants at 2030 and Total Fertility Rate at 2030 and no correlation between the two rates at 2065. With regard to the correlation between Total Number of Immigrants and Total Fertility Rate at two different time points, for one-half of the experts there is no correlation and for the other half a moderate/high negative correlation between Total Number of Immigrants at 2030 and Total Fertility Rate at 2065, while all experts agree on there being no correlation between Total Fertility Rate at 2030 and Total Number of Immigrants at 2065.

Fig. 2.3 Elicited correlations: Total Fertility Rate (average number of children per woman) and Total Number of Immigrants (in thousands)

Fig. 2.4 Expert forecasts of Male and Female Life Expectancies (years)

Figure 2.4 presents the forecasts of the Life Expectancies for males and females and Fig. 2.5 presents the corresponding correlations. Note that the forecasts were provided by 16 experts, but only nine provided all inputs needed for the derivation of the correlations. All forecasts show a low variability, both at 2030 and 2065. With regard to the correlations, there is agreement among the experts on the correlation of Male Life Expectancy at the two time points and on the correlation between Male and Female Life Expectancies at 2030, all experts forecasting a positive high correlation. Similarly, almost all experts forecast a positive high correlation between Female Life Expectancy at 2030 and Male Life Expectancy at 2065. Regarding the correlation between Female Life Expectancy at the two time points, we observe for three experts a positive high correlation, for one expert a negative high correlation, and for all other experts, correlations almost equal to zero. In the case of the correlations between Male and Female Life Expectancies at 2065, three experts

Fig. 2.5 Elicited correlations: Male and Female Life Expectancies

forecast a negative high correlation, one expert a very low positive correlation and the remaining five experts a positive high correlation.

Similar disagreement is expressed about the correlation between Male Life Expectancy at 2030 and Female Life Expectancy at 2065: Two experts forecast a negative high correlation, three experts a zero correlation and four experts a positive high correlation.

Figure 2.6 displays the forecasts and correlations for Total Number of Emigrants provided by 16 experts. There is a high variability in the forecasts, both for 2030 and 2065. In particular, we can notice that six experts provided the same forecasts at 2030 and 2065, this is the reason why in the top panel of Fig. 2.6 only red asterisk is displayed for these experts. Regarding the across-time correlations, almost all experts forecast a positive high correlation. We could work out correlations only for 14 experts, since two of them did not provide the needed conditional forecasts.

Based on the results of the elicitation procedure, the forecasting method explained in the previous section was then used to simulate the joint forecast distribution of Total Fertility Rate and Total Number of Immigrants at 2030 and 2065 and of the joint forecast distribution of Male and Female Life Expectancies at 2030 and 2065. The same method was applied to the separate simulation of the forecast distributions of the Total Number of Emigrants and of the Mean Age at Birth at 2030 and 2065. The prior parameters were specified as described in the previous section. In particular, the means and variances of the priors for the summary indicators were specified based on the ISTAT scenarios available in 2012: μ_R was set equal to the vector of central scenarios and the variances for Σ_R were derived from the high–low ISTAT scenarios available in 2012. The covariances were all fixed to 0. The mixture model was fit for different choices of the number J of components of the mixture, ranging from two to five. The model with two components was selected, since it had the smallest BIC.

The results shown in Tables 2.1, 2.2, 2.3, and 2.4 were obtained through a long run of the MCMC algorithm that provided 20,000 samples from the joint

Fig. 2.6 Expert forecasts (in thousands) and correlations, Total Number of Emigrants

Table 2.1 2030 and 2065 Demographic Indicators: Prior, expert evaluations and posterior means and standard deviations, TFR as average number of children per woman, Life Expectancies as number of years, Total Number of Immigrants and of Emigrants in thousands

Demographic Indicator	2010	Prior Mean		Prior SD		Opinions mean		Opinions SD		Posterior Mean		Posterior SD	
		2030	2065	2030	2065	2030	2065	2030	2065	2030	2065	2030	2065
Total Fertility Rate	1.42	1.50	1.50	0.50	0.50	1.55	1.65	0.12	0.14	1.53	1.64	0.10	0.16
Mean Maternal Age	31.40	31.80	32	1	1	31.95	32.52	1.50	2.17	31.80	31.80	0.81	0.93
Male Life Expectancy	79.50	82.80	86.60	3	3	83.01	86.96	1.84	2.94	82.93	86.89	1.76	2.04
Female Life Expectancy	84.60	87.70	91.50	3	3	87.24	90.88	1.21	2.37	87.21	91.02	1.54	2.88
Number of Immigrants	408.66	321	304	100	100	253.57	212.14	94.99	145.68	280.20	262.55	66.28	88.22
Number of Emigrants	83.81	101	128	30	30	70.00	62.81	34.69	35.76	91.48	91.01	23.13	24.23

Table 2.2 Prior and posterior correlations, Total Number of Immigrants and Total Fertility Rate

	Prior correlation	Posterior correlation
Tot. Num. Immigrants in 2030 and 2065	0.4473	0.4381
TFR at 2030 and 2065	0.2084	0.3233
Tot. Num. Immigrants in 2030 and TFR at 2030	0.2448	0.1288
Tot. Num. Immigrants in 2030 and TFR at 2065	−0.2175	−0.0814
Tot. Num. Immigrants in 2065 and TFR at 2030	0.0003	−0.0233
Tot. Num. Immigrants in 2065 and TFR at 2065	−0.476	−0.0712

Table 2.3 Prior and posterior correlations, Male and Female Life Expectancies at birth

	Prior correlation	Posterior correlation
Male Life Expectancies at 2030 and 2065	0.8569	0.6702
Female Life Expectancies at 2030 and 2065	0.1413	0.082
Male Life Expectancies at 2030 and Female Life Expectancies at 2030	0.9636	0.9066
Male Life Expectancies at 2030 and Female Life Expectancies at 2065	0.1017	0.018
Male Life Expectancies at 2065 and Female Life Expectancies at 2030	0.8418	0.6427
Male Life Expectancies at 2065 and Female Life Expectancies at 2065	0.0496	−0.0065

Table 2.4 Total population forecasts and ISTAT estimates (in millions)

Year	Forecast	%85 forecasts interval	ISTAT estimates
2011	60,484	(60,479 60,490)	60,626
2012	60,659	(60,637 60,679)	59,394
2013	60,814	(60,767 60,862)	59,685
2014	60,952	(60,870 61,045)	61,035
2015	61,073	(60,948 61,199)	60,796
2016	61,180	(61,004 61,357)	60,666
2017	61,275	(61,041 61,512)	60,589
2018	61,361	(61,061 61,666)	60,484

posterior distribution of the indicators at the two time points, 2030 and 2065; the first 10,000 were discarded, as burn-in. The convergence of the algorithm was assessed through different techniques, the trace plots of the chains run for Total Fertility Rate and Total Number of Immigrants and discarding the first 10,000 draws are depicted in Fig. 2.7. The analysis can be replicated using the Matlab code “supraBayesian_popproj” available in the online material of this book.

Table 2.1 shows the prior and posterior means and standard deviations for the summary indicators at 2030 and 2065, along with the arithmetic average and standard deviations of the corresponding expert opinions. For all indicators, as expected the posterior standard deviation at 2030 is smaller than the one at 2065, and both posterior standard deviations are smaller than the prior ones, since noninformative priors are used. Our forecasts show a lower variability compared

Fig. 2.7 Trace plots, TFR as average number of children per woman, Total number of Immigrants in thousands

against the one induced by ISTAT scenarios. The ISTAT central scenario, used as prior mean, predicts a Total Fertility rate equal to 1.5 both for 2030 and 2065, the arithmetic average of the expert opinions is 1.55 at 2030 and 1.65 at 2065, and our model predicts, as posterior mean, 1.53 at 2030 and 1.64 at 2065. The same kind of pattern can be observed for the Total Number of Immigrants, for which the ISTAT central scenario, used as prior mean, predicts for 2030 321,000 and for 2065 304,000; the arithmetic average of the expert elicitations is around 254,000 for 2030 and 212,000 for 2065; while the model forecasts, as posterior mean of the indicator, are 280,000 for 2030 and 262,000 for 2065. Regarding the Life Expectancies, the ISTAT central scenario, used as prior mean, predicts a Male Life Expectancy equal to 82.80 at 2030 and equal to 86.60 at 2065, a Female Life Expectancy equal to 87.70 at 2030 and equal to 91.50 at 2065, while the arithmetic averages of the expert opinions are 83.01 for 2030 and 86.96 for 2065 for males and 87.24 and 90.88 for females. The mixture model predicts posterior means of Male Life Expectancy equal to 82.93 at 2030 and to 86.89 at 2065, and a Female Life Expectancy equal to 87.21 at 2030 and to 91.02 at 2065. The ISTAT central scenario on Total Number of Emigrants predicts 101,000 emigrants in 2030 and 128,000 in 2065; the arithmetic average of the expert evaluations is 70,000 and 62,810 for 2030 and 2065 respectively; and the model predicts a Total Number of Emigrants equal to 91,480 in 2030 and 91,010 in 2065.

Table 2.2 provides the prior and posterior correlations at the same time (2030 and 2065) and across time for the Total Fertility Rate and the Total Number of Immigrants, and the correlations at the same time and across time between the two summary indicators. It is worth emphasizing that the prior correlations are derived from Σ_0 , which was obtained as the scaled arithmetic average of the covariance matrices elicited from each expert, while the posterior correlations are obtained from the 10,000 draws of the two rates at the two time points. The model predicts a moderate positive posterior across-time correlation for the Total Number of Immigrants and a moderate/low positive across-time correlation for Total Fertility Rate. All posterior correlations between the two rates are around zero, apart from the correlation between Total Number of Immigrants at 2030 and Total Fertility Rate at 2030, equal to 0.1288. The forecast of this positive, even though weak, correlation is in concordance with Sobotka (2003), Sobotka et al. (2008), Haug et al. (2002), Coleman (2006), and Goldstein et al. (2009), who argue that fertility rates in many European countries may have been increased by the compositional effect of the rising share of higher-fertility immigrants. The fact that the correlation between the two rates is almost zero at 2065 is due, in our opinion, to the difficulty for the experts to express, even indirectly, opinions on the long term associations.

Table 2.3 presents the prior and posterior correlations at 2030 and 2065 and across-time for the Male and Female Life Expectancy. Based on the elicited opinions, our model predicts a moderate/high correlation between Male Life Expectancy at 2030 and 2065, between Male and Female Life Expectancy at 2030, and between Female Life Expectancy at 2030 and Male Life Expectancy at 2065. All other correlations are predicted to be around zero.

For each of the summary indicators, from the 10,000 values obtained as draws from the corresponding posterior distribution, 10,000 trajectories over the time interval from 2010 to 2065 are obtained by relying on standard quadratic interpolation techniques. The forecast of the Italian Population from 2010 to 2065 was then derived based on the cohort-component model. The inputs of the model are the age- and time-specific fertility rates, age- and time-specific male and female survival rates, and age- and time-specific net migration rates, obtained from the corresponding summary indicators by applying standard smoothing techniques. In particular, the matrices of male and the matrices of female age- and time-specific mortality rates are obtained from the corresponding life expectancies at birth on the basis of the extended model life tables provided by the United Nations. The matrices of age- and time-specific fertility rates are derived from the vectors of total fertility rates and the vectors of mean maternal ages at birth, using a rescaled normal model. For migration, the matrices of male and female age-specific net migration flows are derived from the corresponding vectors of total net flows, applying a rescaled gamma model. This is a simplifying assumption that assumes the absence of pre-school, retirement, and post-retirement peaks in the age profile of migrations, with the only peak being related to labour migration.

Starting from an estimated total population at 2010 of 60,343 million, our model predicts a slight increase at 2030, with the total population forecast to be 61,795 million with an 85% forecast interval ranging from 60,137 million to 63,475 million. After 2030, the total population is predicted to decrease, reaching 57,146 million, with an 85% forecast interval from 50,135 to 64,503 million. As expected, the latter forecasts have a higher variability.

Table 2.4 presents the Italian population forecasts and prediction intervals obtained through our method and the values estimated by ISTAT from 2011 to 2018. Overall, our forecasts are above the ISTAT estimates, with differences in absolute value ranging from 142,000 in 2011 to 1,265,000 in 2018. One explanation of this over-prediction might be found in Table 2.1, where we see that on average, expert opinions at 2030 and especially at 2065 on Total Fertility are well above what is expected by the ISTAT central scenarios, and the same for Male Life Expectancy. It is as well plausible that the experts did not perceive the persistence of the great recession, which was linked to lower fertility (see Goldstein et al. 2013, Comolli and Bernardi 2015, Comolli 2017 and Matysiak et al. 2018) and to lower levels of net migration (see Anelli and Peri 2017), leading to smaller population sizes. The failure of our method to capture the decrease in the total population estimated by ISTAT from 2014 to 2018 might be due as well to the interpolation techniques used for the derivation of the forecast indicators between the starting time 2010 and 2030 and between 2030 and 2065.

2.4 Concluding Remarks

The method we have suggested makes explicit use of expert evaluations to derive probabilistic forecasts of the future trends in the population by age and sex. Our method makes use of expert opinions not only about the expected future behaviour of the demographic components but also about the across-time correlations of single indicators and about the correlations between the indicators. The expert evaluations are then combined in such a way as to take into account their associations. The advantages and limitations of an expert based approach have been discussed in the Introduction. Here, it is worth emphasizing the fact that experts are always involved in the population forecast at different levels of the forecasting procedure and to different degrees. In the time series approach, experts contribute to the choice of the model and the specification of the prior distributions. In the extrapolation from past errors approach, experts provide the central trajectories and contribute to the evaluation of the forecasts. Furthermore, we do not neglect information on past trends when considering expert evaluations as the main source for deriving the population forecasts. Indeed, expert evaluations should be based as well on such information. Our method allows taking into account the overconfidence of experts in their opinions, which might produce an undervaluation of the uncertainty of the forecasts. The entire process is treated within the formal framework provided by the Bayesian paradigm.

Our modelling strategy has some specific limitations. The main limitation is that we have focussed on summary indicators of the demographic changes, which are then converted into age schedules based on parametric models. An extension of the method is in principle feasible, the main difficulty being related to the elicitation of opinions on curves, depicting age patterns. Moreover, our method does not take into account the uncertainty in the initial distributions of the population by age and sex, this being particularly problematic in the case of inconsistencies between the census-based and register-based population records. Experts could be asked to express their opinion on the initial structure by age and sex of the population as well. Lastly, our method exploits expert opinions to derive the forecast distribution of two summary indicators at two time points, while forecasts for the years between the starting one and the midpoint, t_1 and between t_1 and the final time T , are obtained relying on standard interpolation techniques. In principle, our method can be generalized to the case of more than two indicators at more than two time points. The main limitation is on the side of the inputs of the forecasting procedure for the indicators. The indirect elicitation of the correlations requires, as seen in Sect. 2.2, questions on conditional forecasts that in the case of more than two time points and more than two indicators can be extremely cumbersome. More work should then be devoted to the selection of suitable interpolation techniques and experts could be involved in this choice as well, by asking them to express their opinion on the expected trends between the considered time points.

As a general consideration, the performance of the forecasting procedure relies on the number of experts and their commitment. The application of the method

discussed in the previous section was based on the results of the first round of the questionnaire, when at most 16 experts contributed. A new round of the questionnaire is currently running, the results of which are not yet available. However, almost 100 experts have contributed, and we expect a better performance of the method here suggested.

Acknowledgements The author would like to thank Francesco Billari, Eugenio Melilli and an anonymous referee for extremely useful comments and suggestions.

References

- Albert, I., Donnet, S., Guihenneuc-Jouyaux, C., Low-Choy, S., Mengersen, K., Rousseau, J., et al. (2012). Combining expert opinions in prior elicitation. *Bayesian Analysis*, 7(3), 503–532.
- Alders, M., Keilman, N., & Cruijsen, H. (2007). Assumptions for long-term stochastic population forecasts in 18 European countries. *European Journal of Population/Revue Européenne de Démographie*, 23(1), 33–69.
- Alho, J. M., & Spencer, B. D. (1990). Error models for official mortality forecasts. *Journal of the American Statistical Association*, 85(411), 609–616.
- Alkema, L., Raftery, A. E., Gerland, P., Clark, S. J., Pelletier, F., Buettner, T., & Heilig, G. K. (2011). Probabilistic projections of the total fertility rate for all countries. *Demography*, 48(3), 815–839.
- Alkema, L., Gerland, P., Raftery, A., & Wilmoth, J. (2015). The United Nations probabilistic population projections: An introduction to demographic forecasting with uncertainty. *Foresight (Colchester, Vt.)*, 2015(37), 19.
- Anelli, M., & Peri, G. (2017). Does emigration delay political change? Evidence from Italy during the Great Recession. *Economic Policy*, 32(91), 551–596.
- Bijak, J., & Bryant, J. (2016). Bayesian demography 250 years after Bayes. *Population Studies*, 70(1), 1–19.
- Bijak, J., & Wiśniowski, A. (2010). Bayesian forecasting of immigration to selected European countries by using expert knowledge. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 173(4), 775–796.
- Billari, F. C., Graziani, R., & Melilli, E. (2012). Stochastic population forecasts based on conditional expert opinions. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 175(2), 491–511.
- Billari, F. C., Graziani, R., & Melilli, E. (2014). Stochastic population forecasting based on combinations of expert evaluations within the Bayesian paradigm. *Demography*, 51(5), 1933–1954.
- Booth, H. (2006). Demographic forecasting: 1980 to 2005 in review. *International Journal of Forecasting*, 22(3), 547–581.
- Booth, H., & Tickle, L. (2008). Mortality modelling and forecasting: A review of methods. *Annals of Actuarial Science*, 3(1–2), 3–43.
- Booth, H., Maindonald, J., & Smith, L. (2002). Applying lee-carter under conditions of variable mortality decline. *Population Studies*, 56(3), 325–336.
- Cairns, A. J. G., Blake, D., & Dowd, K. (2006). A two-factor model for stochastic mortality with parameter uncertainty: Theory and calibration. *Journal of Risk and Insurance*, 73(4), 687–718.
- Cairns, A. J. G., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., & Khalaf-Allah, M. (2011). Mortality density forecasts: An analysis of six stochastic mortality models. *Insurance: Mathematics and Economics*, 48(3), 355–367.
- Coleman, D. (2006). Immigration and ethnic change in low-fertility countries: A third demographic transition. *Population and Development Review*, 32(3), 401–446.

- Comolli, C. L. (2017). The fertility response to the great recession in Europe and the United States: Structural economic conditions and perceived economic uncertainty. *Demographic Research*, 36, 1549–1600.
- Comolli, C. L., & Bernardi, F. (2015). The causal effect of the great recession on childlessness of white American women. *IZA Journal of Labor Economics*, 4(1), 21.
- French, S. (1980). Updating of belief in the light of someone else's opinion. *Journal of the Royal Statistical Society: Series A (General)*, 143(1), 43–48.
- French, S. (1981). Consensus of opinion. *European Journal of Operational Research*, 7(4), 332–340.
- Gelfand, A. E., Mallick, B. K., & Dey, D. K. (1995). Modeling expert opinion arising as a partial probabilistic specification. *Journal of the American Statistical Association*, 90(430), 598–604.
- Genest, C., Zidek, J. V., et al. (1986). Combining probability distributions: A critique and an annotated bibliography. *Statistical Science*, 1(1), 114–135.
- Goldstein, J. R. (2004). Simpler probabilistic population forecasts: Making scenarios work. *International Statistical Review*, 72(1), 93–106.
- Goldstein, J. R., Sobotka, T., & Jasiloniene, A. (2009). The end of “lowest-low” fertility? *Population and Development Review*, 35(4), 663–699.
- Goldstein, J., Kreyenfeld, M., Jasiloniene, A., & Örsal, D. D. K. (2013). Fertility reactions to the “Great Recession” in Europe: Recent evidence from order-specific data. *Demographic Research*, 29, 85–104.
- Haug, W., Compton, P., Courbage, Y., et al. (2002). *The demographic characteristics of immigrant populations*, Vol. 38. Strasbourg: Council of Europe Publishing.
- Hyndman, R. J., & Booth, H. (2008). Stochastic population forecasts using functional data models for mortality, fertility and migration. *International Journal of Forecasting*, 24(3), 323–342.
- Hyndman, R. J., & Ullah, M. S. (2007). Robust forecasting of mortality and fertility rates: A functional data approach. *Computational Statistics and Data Analysis*, 51(10), 4942–4956.
- Hyndman, R. J., Booth, H., & Yasmeen, F. (2013). Coherent mortality forecasting: The product-ratio method with functional time series models. *Demography*, 50(1), 261–283.
- ISTAT. (2017). Il futuro demografico del paese. Nota metodologica. Technical report.
- Keilman, N. (1990). *Uncertainty in national population forecasting: Issues, backgrounds, analyses, recommendations*, Vol. 20. Boca Raton: CRC.
- Keilman, N. (2018). Probabilistic demographic forecasts. *Vienna Yearbook of Population Research*, 16, 1–11.
- Keilman, N., Pham, D. Q., & Hetland, A. (2002). Why population forecasts should be probabilistic—illustrated by the case of Norway. *Demographic Research*, 6, 409–454.
- Lavine, M., & West, M. (1992). A Bayesian method for classification and discrimination. *Canadian Journal of Statistics*, 20(4), 451–461.
- Lee, R. D. (1993). Modeling and forecasting the time series of us fertility: Age distribution, range, and ultimate level. *International Journal of Forecasting*, 9(2), 187–202.
- Lee, R. D., & Carter, L. R. (1992). Modeling and forecasting us mortality. *Journal of the American Statistical Association*, 87(419), 659–671.
- Lee, R. D., & Tuljapurkar, S. (1994). Stochastic population forecasts for the united states: Beyond high, medium, and low. *Journal of the American Statistical Association*, 89(428), 1175–1189.
- Lindley, D. (1983). Reconciliation of probability distributions. *Operations Research*, 31(5), 866–880.
- Lindley, D. V. (1985). Reconciliation of discrete probability distributions. *Bayesian Statistics*, 2(375–390), 375–390.
- Lutz, W. (2013). *The future population of the world: What can we assume today*. London: Routledge.
- Lutz, W., & Goldstein, J. R. (2004). Introduction: How to deal with uncertainty in population forecasting? *International Statistical Review*, 72(1), 1–4.
- Lutz, W., Sanderson, W. C., & Scherbov, S. (1998). Expert-based probabilistic population projections. *Population and Development Review*, 24, 139–155.

- Matysiak, A., Vignoli, D., & Sobotka, T. (2018). The great recession and fertility in Europe: A sub-national analysis. Technical report, Vienna Institute of Demography Working Papers.
- Morris, P. A. (1974). Decision analysis expert use. *Management Science*, 20(9), 1233–1241.
- Raftery, A. E., Chunn, J. L., Gerland, P., & Ševčíková, H. (2013). Bayesian probabilistic projections of life expectancy for all countries. *Demography*, 50(3), 777–801.
- Roback, P. J., & Givens, G.H. (2001). Supra-Bayesian pooling of priors linked by a deterministic simulation model. *Communications in Statistics-Simulation and Computation*, 30(3), 447–476.
- Sobotka, T. (2003). Tempo-quantum and period-cohort interplay in fertility changes in Europe: Evidence from the Czech Republic, Italy, the Netherlands and Sweden. *Demographic Research*, 8, 151–214.
- Sobotka, T., et al. (2008). Overview chapter 7: The rising importance of migrants for childbearing in Europe. *Demographic Research*, 19(9), 225–248.
- Stoto, M. A. (1983). The accuracy of population projections. *Journal of the American Statistical Association*, 78(381), 13–20.
- Winkler, R. L. (1981). Combining probability distributions from dependent information sources. *Management Science*, 27(4), 479–488.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.