Chapter 1
Introduction

Nico Keilman and Stefano Mazzuco

1.1 Demographic Forecasting

Future trends in population size, age structure, regional distribution, and other demographic variables are of paramount importance for a wide range of planning situations. Government policy for old-age pensions and long-term care depends on the number of elderly in the future. An assessment of future trends in population variables also is an important prerequisite for exploring environmental issues and the demand of resources in the future. Other things remaining the same, a larger population implies more use of water, electricity, fuel, food etc. in a certain region. Stronger needs for transportation are another effect of growing populations. Local planners have to decide on investments in hospitals and schools. Retailers of certain products (such as baby food) are interested in the size of particular age groups in the future.

Demographic projections and forecasts rely on assumptions of the future developments for components of change for population size, that is, births and fertility, deaths and mortality, and international migration when the interest is in the population for a country as a whole. In case one considers the future state of a certain population sub-group (e.g. persons who live in a specific region or those who are currently divorced), additional components are relevant (regional migration, marriage and marriage dissolution).

N. Keilman
Department of Economics, University of Oslo, Oslo, Norway
e-mail: nico.keilman@econ.uio.no

S. Mazzuco (✉)
Department of Statistical Sciences, University of Padova, Padova, Italy
e-mail: stefano.mazzuco@unipd.it

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_1

Given the importance of insight in future demographic trends, many statistical agencies routinely compute national population forecasts. They do so by means of the so-called cohort-component model, which has become the standard approach in population forecasting (National Research Council – NRC 2000; UNECE 2018). This model requires assumptions on future trends of fertility, mortality, and international migration. We will discuss this approach further in Sect. 1.2.

To make accurate demographic forecasts is both an art and a science, similar to predictions in other fields (Tetlock and Gardner 2016). The scientific part is in the model, and in the fine mathematical and statistical details of the computations. However, to formulate reliable assumptions for the future course of fertility, mortality and migration is an art, largely. Most of the research on demographic forecasting aims at increasing the scientific part, and reducing the impact of selecting the right assumptions – the "art part" in population forecasting. "The quest for knowledge about the future has moved from the supernatural towards the scientific" (Willekens 1990, 9). One way to achieve this aim is to formulate explicit models for fertility, mortality and migration. In that case, one attempts to find a model that describes the historical development of these components of change accurately enough. The model may be an explanatory model with exogenous variables, or a purely statistical (e.g. time series) model. In either case, the model is used to extrapolate the components into the future, and next their future values are used as inputs for the cohort-component model.

The primary aim of this book is to sketch new developments in the scientific part of demographic forecasting. It does not give an extensive review of the field. Such reviews have appeared regularly; see, for example, Hajnal 1955; Keyfitz 1972; Land 1985; Willekens 1990; Keilman and Cruijsen 1992; National Research Council – NRC 2000; Wilson and Rees 2005; Booth 2006; Alho and Spencer 2005; Alho 2015. In contrast, with this book we wanted to show the readers examples of promising new research on demographic forecasting.

In the remainder of this chapter, we discuss selected topics in demographic forecasting, thus sketching the wider context of many of the problems that our authors address. We start with a brief overview of the cohort-component tradition (Sect. 1.2). Next, we describe in Sect. 1.3 how population forecasters account for the inherent uncertainty in their results. Common approaches are to use various deterministic scenarios or, alternatively, a probabilistic model. An important distinction in the statistical modelling of the components of change is that between a Bayesian and a frequentist perspective. We discuss both approaches in Sect. 1.4. Population forecasters often rely on the opinions of experts, when they formulate their assumptions on the future trajectories of demographic components. However, in some cases these trajectories are purely data-driven. This is the topic of Sect. 1.5. Another issue, taken up in Sect. 1.6, is whether one should use data from the country of interest only, or also include trends in other countries. A recent development in demographic forecasting is the evaluation of probabilistic forecasts. Forecasts of this type have been computed since the mid-1980s, and some statistical agencies, too, produce them regularly. After a few decades, one knows the actual development of the variables of interest. Hence, one may want to know how well the forecast, published in terms of a predictive distribution, has performed. This is taken up in Sect. 1.7. Once the forecast has been computed, an important question concerns the best way to communicate its results to the user. Demographers could learn from forecasters in other disciplines, where this question has been analysed. We summarize the most important findings in Sect. 1.8. Section 1.9 gives a brief presentation of the chapters that follow.

1.2 The Cohort-Component Approach[1]

The main idea of the cohort-component model (CCM) is to update a table with known numbers for the population pyramid, taken from a recent census or from a population register, to a new table 1 year later. The update requires assumptions on mortality (the share of persons of a given age who survive 1 year later), fertility (the mean number of children per woman born during the year), and migration (for instance, age- and sex-specific numbers of net-migration). Based on assumptions of this kind for many years in the future, the process can be repeated, resulting in a population forecast for many years in the future. See demographic textbooks such as those by Preston et al. (2001) or Rowland (2003) for technical details, and O'Neill and Balk (2001) for a non-technical introduction.

Edwin Cannan first developed this forecasting approach in 1895. By the 1930s, it had become generally accepted by the statistical agencies of many countries (De Gans 1999). In the 1970s, the model was extended to include a regional breakdown of the population (multiregional model; see, e.g., Rogers 1995), or an extra dimension in general, such as educational level, labour market status, or household status (multistate/multidimensional model). Chapter 10 by Zhang and Bryant and Chap. 11 by Raymer, Bai, and Smith focus on inter-regional migration and follow the tradition of multiregional models.

The CCM is by its nature a pure accounting approach. The new population equals the old one minus deaths and emigrants, plus surviving births and surviving immigrants. This process is repeated for each age group, men and women separately. Assumptions for the three components of change are in terms of age-and sex-specific rates, used as inputs by the CCM. Chapter 4 by Castiglioni, Dalla-Zuanna and Tanturri, and Chap. 9 by Keilman and Kristoffersen discuss certain aspects of this approach.

The mechanical nature of the CCM-approach has been criticized, since it ignores possible feedback mechanisms. Rapid population growth, resulting from high rates for fertility and immigration leads to increasing population density. However, the

[1]Parts of this section and the next one are based upon the paper "Uncertainty in population forecasts for the twenty-first century" by the first author, forthcoming in Annual Review of Resource Economics 2020. Permission to reuse this material is gratefully acknowledged. https://www.annualreviews.org/page/authors/author-instructions/distributing/permissions

CCM does not account for a direct effect of population density on future fertility, mortality, or emigration. More generally, in the long run, one should take into account that the scarcity of resources and degradation of the environment may have an impact on human behaviour and population dynamics (Cohen 2010; De Sherbinin et al. 2007). Also, if the crude birth rate is lower than the crude death rate for a number of decades, population size will fall (assuming that migration has little or no effect), and authorities will likely attempt to promote a pro-natalist policy. Feedbacks of this kind are not included in demographic forecasts that follow the CCM-tradition, at least not explicitly. In some cases, this is reasonable, because demographic variables are much less important than non-demographic variables. For instance, Raftery et al. (2017) combine country-specific probabilistic population forecasts with forecasts of CO_2-emissions and temperature change to 2100. They find that population growth is not a major factor that contributes to global warming, and a feedback mechanism is not necessary. Other studies do include an explicit feedback. For instance, in its population forecast, Statistics Norway assumes that future immigration numbers for various immigrant sub-groups depend, among others, of the stock of migrants already present in the country (Cappelen et al. 2014). See also National Research council – NRC (2000, pp. 31–32) and O'Neill and Balk (2001) for explanations and discussions of the feedback problem, and Sanderson (1998) for the lack of explanatory factors in population forecasts. Burch (2003) gives a more general critical review of the CCM.

1.3 Deterministic Scenarios and Probabilistic Forecasts

Most statistical agencies in the world that compute population forecasts do so using a deterministic approach (NRC 2000). They analyse historical trends in fertility, mortality, and migration, and extrapolate those trends into the future, using expert opinion and statistical techniques. The extrapolations reflect their best guesses. In addition to computing a likely development of population size and structure, many agencies also compute a high and a low variant of future population growth, in order to tell forecast users that future demographic developments are uncertain. For example, the previous official population forecast for Norway, published in 2018, indicates 6.5 million inhabitants in 2060, if current trends continue (see https://www.ssb.no/en/statbank/list/folkfram). However, population growth to 2060 might be weaker or stronger than what current trends suggest, leading to population sizes between 5.8 and 7.8 million persons. The forecasters assumed low and high trajectories for future fertility (leading to 1.6 or 1.9 children per woman on average in 2060), life expectancy of men (between 86.0 and 90.4 years in 2060) and women (between 88.1 and 92.1 years), and international migration (a migration surplus between 10,700 and 41,400 persons annually).

Different projections or scenarios can be produced by systematically combining different assumptions. Collectively, those different scenarios can give some impression of the degree of uncertainty, but not in any quantified way. The probability that an outcome will be within a certain range is unknown (Dunstan and Ball 2016). We do not know if chances are 30, 60, or 90% that Norway in 2060 will have between 5.8 and 7.8 million inhabitants. Yet in many planning situations, it is important for the users to know how much confidence they should have in the predicted numbers. How robust should the pension system be with respect to fast or slow increases in life expectancies? Should we plan for extra capacity in primary schools, in case future births turn out to be much higher than expected? Indeed, as Keyfitz (1981) wrote almost 40 years ago: “Demographers can no more be held responsible for inaccuracy in forecasting population 20 years ahead, than geologists, meteorologists, or economists when they fail to announce earthquakes, cold winters, or depressions 20 years ahead. What we can be held responsible for is warning one another and our public what the error of our estimates is likely to be”.

Indeed, the statistical agencies of some countries have started to publish their forecasts in the form of probability distributions, following common practice in, for example, meteorology and economics. A key use of probabilistic demographic forecasts is in modelling the long-term fiscal impact of an ageing population by policy agencies (Tuljapurkar 1992; Lee and Tuljapurkar 2000; Alho et al. 2008; Dunstan and Ball 2016).

Various methods for probabilistic population forecasting have been developed since the 1960s, although Törnquist (1949) was probably the first to integrate probabilistic thinking into population forecasting. In this approach, the fertility and mortality rates, as well as the migration parameters are random variables. This means that the predicted population becomes random. Early contributions were made by Pollard (1966), Sykes (1969), Schweder (1971), Alho and Spencer (1985), and Cohen (1986). The initial aim was to find analytical solutions for the predictive distributions of the variables of interest. Due to correlations between components, between ages and between men and women, as well as autocorrelations in all variables, approximations were necessary (Tuljapurkar 1992). Later work (e.g., Keyfitz 1985; Kuijsten 1988; Lee and Tuljapurkar 1994) used Monte Carlo simulation.

Statistical agencies of some countries have started to publish the results of probabilistic forecasts, following common practice in, for example, meteorology and economics. Statistics Netherlands pioneered the field; see Alders and De Beer (1998). Statistics New Zealand (2011) and Statistics Italy (ISTAT 2018) are the other two known examples. In Chap. 3, Dion, Galbraith, and Sirag suggest that Statistics Canada is likely to follow soon. In addition, we should mention the Population Division of the United Nations, which is responsible for regular updates of population forecasts for all countries of the world. In 2014, the Population Division issued the first official probabilistic population forecasts for all countries, using the methodology developed by Raftery et al. (2012). See also Gerland et al. (2014) and Ševčíková et al. (2016). These probabilistic forecasts do not replace the traditional deterministic UN population forecasts, but supply additional information to the user. After 2014, the UN updated the probabilistic forecasts a few times. The most recent revision is from 2019; see http://esa.un.org/unpd/wpp/Graphs/Probabilistic/POP/TOT/900. The aim of a probabilistic forecast is not to present estimates of future trends that are more accurate than a deterministic forecast, but rather to give the user a more complete picture of prediction uncertainty.

1.4 Bayesian vs Frequentist

What emerges from the pages of this book is consistent with the most recent literature: the Bayesian approach to population forecasting (we would better say to population studies, in general) is rapidly gaining ground. This has already been noticed by Bijak and Bryant (2016) who also highlight that such an increase has certainly been accelerated – if not triggered, by work at the United Nations, when the World Population Prospects in 2014 for the first time were based on a Bayesian hierarchical model (Gerland et al. 2014). The model is "hierarchical" as it considers a single model for all countries, but there are country-specific parameters, leading to a "hierarchy" in the model structure. However, other explanations of the increasing number of Bayesian population forecasts can be given, as Bayesian analysis is on the rise in general, due also to increasingly fast algorithms which make the computational burden of Bayesian inference lighter and lighter. Indeed, in past years, the main obstacle to Bayesian inference was its intractability: apart from specific cases, deriving posterior distributions cannot be done analytically, thus approximation should be used. Nowadays, several algorithms (MCMC, Hamiltonian MonteCarlo, Variational Bayes, to mention some of them) allow fast solutions. Moreover, it should also be noted that forecasting is a natural product of Bayesian inference. As Geweke and Whiteman (2006) note, forecasting means that one uses the information at hand to make statements about the likely course of events, or said differently, to predict future outcomes, conditionally on what we know. Bayesian inference implies conditioning on what we know (data) to predict what is unknown (the so-called posterior distribution), so one might say that Bayesian forecasting is actually Bayesian inference with missing data: missing data is future value of the outcome considered, which posterior distribution is derived based on the prior information, represented by the past values of data. Therefore, it is quite natural that if the number of Bayesian inference applications increases in population studies, also Bayesian forecasting follow the same trend. However, this does not clarify whether the Bayesian approach provides something more (or something different) than the frequentist one. One suggested advantage is that within a Bayesian framework, information from previous studies or from experts' opinions can easily and transparently be incorporated into the forecasting model through a proper informative prior. This is certainly attractive for the field of population forecasting, where experts' opinions have been used in a non-systematic manner, and where it has been already proposed to use such opinions even in the framework of probabilistic population forecasting (Lutz et al. 1996). However, we do not always see the use of priors as elicitation of experts' opinions, neither in Chaps. 2, 5, or 10 proposing a Bayesian approach, nor in the UN methodology (Gerland et al. 2014). Aliverti, Durante, and Scarpa (Chap. 5) use prior distributions to specify the structure of temporal dependence, but experts' opinions are not considered. Graziani (Chap. 2) uses experts' opinion, but she treats them as observed data, while non-informative priors are used. Finally, Chap. 10 uses a "weakly informative" prior for several parameters and a generalized random walk with drift to parameters with a time trend. Even UN methods use priors as a way to include statistical noise and temporal dependence to the forecasting model, and experts' opinions are not included. In facts, if we consider, for example, the UN-model for fertility forecasting (see Alkema et al. 2011), we could say that experts' opinions are rather incorporated in the statistical model while priors are essentially non informative. Thus, while in theory it might look appealing to mix in a formal and transparent way expert's opinions and information coming from observed data, in practice this is rarely done.

Dunson (2001) provides a much more pragmatical answer on why it could be advantageous using a Bayesian approach in some cases: the class of statistical models that can be estimated via Bayesian inference is much broader than would be possible with other approaches. In some cases, this involves retrieving the full conditional distribution of parameters and this might be far from straightforward. For example, Aliverti, Durante, and Scarpa (Chap. 5) use a particular result on skew normal distributions (a posterior distribution from a Gaussian prior combined with a skew normal likelihood gives a unified skew normal distribution, see Canale et al. 2016), while Zhang and Bryant (Chap. 10) combine the Gibbs Sampling algorithm with a Metropolis-Hastings step. In other cases, this is not necessary: for example, the software STAN (see Carpenter et al. 2017) use a Hamiltonian Monte Carlo method for which retrieving the full conditional distribution of parameters is not necessary. Another practical advantage of the Bayesian approach is that once that computation has been implemented and posterior distributions of parameters have been obtained, also the posterior of any function of model parameters can be easily achieved. For instance, Zhang and Bryant (Chap. 10) after the estimation step, provide the predicted posterior distribution of migration rates, which are functions of the estimated parameters.

However, complex models can be estimated using a frequentist approach, too: Basellini and Camarda (Chap. 6), for example, decompose mortality age patterns into three components, with a specific model for each of them, and the parameters of these models are jointly estimated with maximum likelihood. In their case, given the complexity of the model, prediction intervals can neither analytically nor numerically be obtained, so a bootstrap procedure (Efron and Tibshirani 1993) is implemented. Such a procedure involves resampling data for K times, which in some cases can be computationally intensive, but not necessarily more intensive than MCMC algorithms that are needed to obtain posterior distribution (from which credibility intervals can be calculated) of parameters.

Summarizing, we believe that an increasing understanding and application of the Bayesian approach in the field of population forecasting is certainly beneficial to this research area, as it enlarges the forecaster's possibilities by broadening the class of models that can be used. At the same time, frequentist approaches remain a valid opportunity, not necessarily a second choice. Our suggestion is to choose the inference approach to be used after having determined what the most appropriate forecast model is. Once this has been decided, inference method can be determined more easily, on the base of the considerations exposed above.

1.5 Expert Opinions vs Data Driven

In the past, much discussion on population forecasting methods was devoted to experts' opinions based forecasts as opposed to probabilistic ones. Experts' opinions based methods are usually referred to as synonym of deterministic methods, until Lutz and Scherbov (1998) proposed an expert judgment based probabilistic method. Thus the discussion in the literature involves the issue of uncertainty quantification: probabilistic models provide a natural assessment of forecast uncertainty, while experts-based methods specify some scenarios (usually three, labeled "high", "medium" and "low") with no possibility of variation (Booth 2004). However, even in case one uses a probabilistic approach (see Sect. 1.3 for examples), experts-based method are still used but integrated in statistical models that ensure random variability and uncertainty quantification. United Nations population forecasts, for instance, are still strongly based on experts' opinion, and, based on demographic transition theory, the UN World Population Prospects (United Nations 2017) suggest all countries' mortality, fertility, and migration rates will converge, eventually, to the same patterns. Castiglioni, Dalla-Zuanna, and Tanturri demonstrate in Chap. 4 that such a convergence is not confirmed by observed data and that it might be useful to reconsider this assumption. This example shows that, while the opposition between experts' opinions based forecasts and probabilistic forecasts no longer has any reason to exist, forecasters have to decide whether, and to what extent, they can rely on experts' judgments or whether they let the data speak for themselves, by using a data-driven method. In this book, you can find two examples of purely experts' opinion based forecasting method (exposed by Graziani in Chap. 2 and by Dion, Galbraith, and Sirag in Chap. 3) and an example of a strongly data-driven method (exposed by Aliverti, Durante, and Scarpa in Chap. 5). Graziani (Chap. 2) does not consider observed data but embeds experts' judgments into a statistical model, so that prediction intervals can be derived with a proper uncertainty quantification. Of course, we have to bear in mind that experts shape their opinions on observed data, so basing forecasts on their judgment does not mean disregarding evidence coming from data. What is paramount when using experts' opinion is the elicitation process of their judgements. Dion, Galbraith, and Sirag (Chap. 3) show an extremely detailed expert elicitation protocol that allows experts to have an appropriate feedback of their judgments, forcing them to reflect more on the likelihood of their opinion. On the data-driven side, Aliverti, Durante, and Scarpa (Chap. 5) propose an extremely flexible statistical model to make forecast, so that no specific pattern is imposed to the data or to forecasts of the fertility age schedule. Interestingly, Graziani (Chap. 2) reports that the experts-based forecast of the Total Fertility Rate in Italy in 2018 on average is above the actual estimates of the Italian national institute of statistics. The explanation is that "experts did not perceive the persistence of the great recession", and this means that if you want to rely on experts' opinion, you have to bear in mind that such opinions are not necessarily right. On the other hand, Aliverti, Durante and Scarpa (Chap. 5) expect that the mean, variability, and skewness of Italian fertility age schedule will remain approximately constant in the future. The simple explanation of this prediction is that the most recent observed age schedules have remained stable, and this trend has been extrapolated to future years. However, an expert could have objected that also between 1995 and 1999 age schedules have remained stable, but the mean age at childbearing has increased after 1999, while the age pattern became less skewed. These two cases help us to understand that the choice on whether relying more on data or on experts' judgment is a delicate one, as both experts and data can be misleading, and a forecaster needs needs to consider very carefully, for each specific case, which of the two sources is reliable. For example, Bergeron-Boucher, Kjærgaard, Pascariu, Aburto, Alvarez, Basellini, Rizzi, and Vaupel (Chap. 7) show that in the case of Denmark, mortality forecasting is difficult due to broken trends generated by a life expectancy stagnation starting in 1980. Therefore, they compare different extrapolative methods and find a high sensitivity of forecasts to model selection. Actually if we use a cohort perspective, Denmark's life expectancy trend is much smoother than what period life expectancies suggest. This is another example showing that data, even though seemingly "neutral", can also misguide forecasts.

1.6 Coherent Modelling

Another dilemma that a forecaster has to deal with is whether the forecasts of a demographic component of different populations should be considered jointly or separately. Should, for example, male and female mortality be forecast by one common model, or by two distinct models? Recently, coherent forecasting models have gained ground in the field of mortality (not so much in fertility forecasting), since Li and Lee (2005) have proposed a coherent model, assuming that future trends of mortality in similar (or neighbor) countries are mutually dependent. One may use the same argument for mortality of male and female populations of a given country, assuming that they follow similar trends. However, is pooling countries or population sub-groups always beneficial? Raftery et al. (2013) propose information pooling for mortality forecasting, and actually this might be a good idea in case of scarce or bad-quality data for some specific population. However, in other cases, things might go in the opposite direction. Booth (Chap. 8) shows that implementing a coherent model not necessarily improves the forecasts: for example, a sex-coherent model improves forecasts for the mortality of males, but not for females, compared to independent forecasts. Interestingly, she also shows that much of the performance depends on the standard used in coherent modeling and a same-sex low-mortality-standard is optimal. However, the point that we stress here is that pooling information from other countries (or other populations) has not necessarily a positive effect.

Perhaps the concept of exchangeability – well known in Bayesian statistics literature – can help to explain this issue. Exchangeability means that one can re-order and re-label data while the joint density remains the same. In our context, assuming exchangeability with data from multiple countries means that inference (and prediction) can be equivalently done exchanging data even across countries, i.e. data can be pooled together irrespective of the country where the data come from. This is clearly not realistic at a country level (however it is normally done at the regional level, as national forecasts normally pool regional data), even for very similar countries. A "partial pooling" solution might be more acceptable, and this solution is naturally achieved with a hierarchical model: all data are used for inference and forecast but when forecasting for a specific country, data from other countries are differently weighted. In terms of exchangeability, data can be exchanged across countries, but a "penalty" (or a lower weight) is given to other countries' data, the penalty depending on countries' sample sizes and variabilities (see Jackman 2009, Section 7.1.2 on exchangeability in connection with hierarchical models). Thus, whether pooling (completely or partially) countries together or not depends on the extent to which their data are exchangeable. However, in many cases modelling and forecasting without pooling data is not a viable choice. Raymer, Bai, and Smith (Chap. 11) and Zhang and Bryant (Chap. 10), for example, face the challenge of predicting internal migration. Their main issue is that flows from one region to another may have very low sample sizes (see, for instance, Figure 8 in Zhang and Bryant, where it turns out that migration rates from East to North West are estimated only for one age group). Therefore they are obliged to borrow information for these flows, which leads to some kind of data pooling. Raymer, Bai, and Smith use a multiplicative model for that purpose, while Zhang and Bryant implement a hierarchical model. Are data from different region, at least partially, exchangeable? The answer can be given only by experts, not directly by data.

1.7 Evaluating Probabilistic Forecasts

Once a forecast has been published, some 10–20 years later its accuracy can be evaluated, when *ex-post facto* observed data for population size and age structure have become available. Evaluating deterministic forecasts against empirical data has a long history, which goes back at least to the work by Myers (1954). The topic received systematic attention in the 1980s, by Keyfitz (1981), Ahlburg (1982), and Stoto (1983); see NRC (2000) for a review. However, to assess the accuracy of a probabilistic forecast is difficult, because it requires that one compare a forecaster's predicted probabilities with the actual but unknown probabilities of the events under study. For that reason, statisticians have developed "scoring rules": distance measures between the predicted distribution of the variable in question, and the empirical value it actually turns out to have. Gneiting and Raftery (2007) and Gneiting and Katzfuss (2014) review the field. The score that one finds for a certain variable has no intrinsic meaning. Only in a comparative perspective, one can interpret the scores in a useful manner. Indeed, scoring rules are frequently used in comparing two or more competing probabilistic forecasts. A second type of application is to study how fast the quality of the forecast deteriorates with increasing lead-time (forecast horizon).

The results of a probabilistic forecast can be made available in different ways. For demographic forecasting, we distinguish between forecast results in the form of simulation samples or as prediction intervals. Each category has its own scoring rules.

Although the methodology around evaluation of probabilistic forecasts and scoring rules has been known for some time, there are very few applications of scoring rules to population forecasting. Shang et al. (2016) analyse the accuracy of probabilistic cohort-component forecasts for the UK, and compare two forecasting methods. They use a scoring rule for prediction intervals. Shang (2015) and Shang and Hyndman (2017) evaluate interval forecasts for age-specific mortality rates in various countries, and use interval scores to select the best among several methods of forecasting mortality. Alexopoulos et al. (2018) employ interval scores to prediction intervals of age-specific mortality of England & Wales and New Zealand, and evaluate the predictive performance of five different mortality prediction models. All four papers use holdout samples to evaluate the probabilistic demographic forecasts. We are aware of only one example of genuine out-of-sample evaluation of probabilistic demographic forecasts (Keilman 2020).

As an alternative to using scoring functions to prediction intervals, one could check how large the share of actual data is that fall within the intervals. An example is the work by Raftery et al. (2012), who validate their Bayesian method of forecasting populations for 159 countries by estimating the model based on data for the 40-year period 1950–1990. Next, they use the model to generate a predictive distribution of the full age- and sex-structured population for the 20-year period 1990–2010. They compare the resulting 80% and 95% prediction interval distributions with a test data set of actual observations, and check the proportion of the validation sample that falls within their intervals. These proportions are close to the nominal values of 80% and 95%, and the authors conclude that their approach is satisfactory.

Similarly, in Chap. 10, Zhang and Bryant present Bayesian forecasts for internal migration in Iceland. They consider two models: a baseline model that does not include region-time interactions, and a revised model that does. Both models are estimated with data for the period 1999–2008, and 80% prediction intervals ("credible intervals" in the Bayesian perspective) for the migration rates predicted for the years 2009–2018 are checked against a test dataset with empirical rates for these years. The authors inspect the proportion of values of the test dataset that lie within the 80% credible intervals and find that the revised model is much better calibrated than the baseline model, in that actual coverage (71–73% for the revised model) comes quite close to the nominal coverage (80%). Therefore, the authors base their forecasts on the revised model.

Also Raymer, Bai, and Smith (Chap. 11) forecast inter-state migration for Australia for two 5-year periods: 2006–2011 and 2011–2016. The model they use for forecasting was fitted to observed values for 5-year periods from 1981–1986 to 2001–2006. The authors investigate two versions of the model, and note that the proportions of observed data for the two recent periods that lie within 80% or 95% prediction intervals agree quite well with the nominal values of 80% and 95%.

One important drawback of this approach is the fact that the values from the test dataset are not necessarily independent of each other. They may be generated by correlated variables. This means that one has less information in the test data than the sheer number of values suggests. Thus, a comparison between observed proportions and nominal values is not valid, strictly speaking (Alho and Spencer 2005, 248; Gneiting et al. 2007, 253).

1.8 Communicating Forecast Results

As noted in Sect. 1.3, forecasters use deterministic scenarios and probabilistic forecasts to express the inherent uncertainty in statements about the future size and structure of populations. Since many population forecasts and projections are general-purpose calculations that serve the needs of many different users, often there is no frequent systematic contact between users and the producer of the forecast. However, to communicate forecast results in an appropriate way to the users is of paramount importance. Various questions arise in this context. Does the forecast produce predictions of the type of variables (age groups, components of change, regional disaggregation, persons or households, etc.) that satisfy user needs? Is there enough detail in the predicted variables (one-year age groups, short-term versus long-term forecasts)? Are the results available as data files or in print only? Is the forecast updated when new information on current demographic trends becomes available?

Following up on points made by the National Research Council – NRC (2000) on the presentation of demographic forecasts, a Task Force on Population Projections working for the United Nations Economic Commission of Europe (UNECE) recently formulated a large number of recommendations on communicating population forecasts and projections (UNECE 2018). The task force based its work on information from three distinct sources: a survey among users of national and international population forecasts and projections, a survey among national statistical agencies of UNECE member countries, and a consultation round among a small group of experts in the field of population projections. Finally, a literature review using insights from demography, psychology, and science communication complemented the analysis of perspectives from users, statistical agencies, and experts. The task force addressed a number of issues, including ways to provide pertinent and accessible results, the need for transparency, accounting explicitly for uncertainty, and ways to foster relationships with users. Many of the 26 recommendations for good practice seem obvious, although they are not always followed by statistical agencies. Examples are to communicate results in clear and simple language, to disseminate results by single age and calendar year whenever possible, to make electronic dissemination materials accessible, to provide clear descriptions of data, methods and assumptions, or to clearly define key terms used in dissemination products. However, other recommendations are more novel for official forecasts and projections, such as developing an explicit strategy for characterizing and communicating the uncertainty of population forecasts and projections, identifying and characterizing the major sources of uncertainty, providing both sensitivity and uncertainty analyses, and engaging directly with users in a substantive manner, for instance by using new media.

A number of chapters in this book address demographic uncertainty explicitly (e.g. Graziani in Chap. 2, Dion, Galbraith, and Sirag in Chap. 3, Aliverti, Durante, and Scarpa in Chap. 5, Basellini and Camarda in Chap. 6, and Zhang and Bryant in Chap. 10, Scherbov and Sanderson in Chap. 12). Therefore, an important question is whether forecast users want to know about the uncertainty of the forecast. There is little evidence from systematic investigations, but available data suggest that the answer is yes. The survey organized by the UNECE task force mentioned above showed that 69% of the users who answered the relevant question (N = 148) considered quantification of uncertainty of the projections important or very important for their own work. Of 119 users who gave their opinion about the way uncertainty was stated in the projection they use, 42% noticed it was stated, but that it could be stated more clearly, whereas 29% found uncertainty not clearly stated. On the other hand, about one-third of the statistical agencies were of the opinion that the lack of knowledge about uncertainty among users is a challenge in communicating uncertainty. At the same time, one-third of the agencies noted that users are interested in one single scenario.

The interest of population forecast users in forecast uncertainty noted above is similar to the findings by Wilson and Shalley (2019) for Australia. Using data from a small online survey and subsequent focus groups of subnational population forecast users, the authors find that 90% of users who responded were in favour of receiving information on forecast uncertainty. Reasons selected from a list of options for wanting information on uncertainty consist of the need to emphasise the fact that forecasts are not exact (73%), to aid decision-making based on a range of projected population numbers (58%), and to allow consideration of risk or contingency strategies (57%).

In this connection, it is also worth to report the user and producer experiences of Dunstan and Ball (2016) of Statistics New Zealand after they had implemented a probabilistic approach in 2012; see also Dunstan (2019). They stress that the change from a deterministic to a probabilistic forecast was less difficult to make than one might expect. Uncertainty in fertility, mortality and migration can be modelled simply or with more complexity, and progressively applied to different types of forecasts (national forecasts first, followed by derived forecasts: regional, labour market, ethnic groups, level of education, households etc.). A close contact between forecaster producers and main users is essential in the process of preparing the probabilistic forecast. Many users will be interested in deterministic results only and do not need prediction intervals or the full set of sample paths. They can still employ the probabilistic results, for instance the median forecast, possibly supplemented by the upper and lower bounds of the 80% prediction interval. Moreover, probabilistic forecasts, with their quantified measures of uncertainty, can help statistical agencies to define an appropriate horizon for the forecast. Prediction intervals for demographic variables far into the future become progressively wider and flatter, and hence they do not contain much useful information. However, the situation is very different for different demographic variables. Forecasts for subpopulations are often more uncertain than those for aggregate populations. This means that users of probabilistic forecasts can inspect prediction intervals to make their own informed decisions about the usefulness of different forecasts across any projection period.

Closely related is the notion of the "shelf life" of a forecast, recently developed by Wilson and colleagues; e.g. Wilson (2018), Wilson and Shalley (2019), Wilson et al. (2018), Simpson et al. (2019). The concept is borrowed from perishable food labelling to describe the number of years into the future a population forecast is likely to remain of reasonable quality. In practice, 'reasonable quality' could be defined as the period in which the 80% prediction interval (for a probabilistic forecast), or 80% of past errors (for a deterministic forecast for which past errors are available) remain within $\pm 10\%$ error. When the forecast horizon exceeds the shelf life, the forecast is no longer of reasonable quality. In an illustration using Australian data, Wilson et al. (2018) suggest a shelf life of 8 years for forecasts of a population of 10,000 persons, 12½ years for a forecast population of 50,000, and 14 years when population size is 150,000. Using data for a number of past official subnational English forecasts, Simpson et al. (2019) find shelf lives of 6 years for London Boroughs forecasts, and 21 years for Metropolitan Districts. While the choice of 10% for the absolute error is rather arbitrary, the respondents to the Australian user survey found the concept of shelf life very useful (Wilson and Shalley 2019).

Demographers may learn from experiences in other fields when the interest is in communicating the results of a probabilistic forecast. Bijak et al. (2015) remind us of meteorology and climatology, aviation, macroeconomics, as well as cognitive sciences. See also Raftery (2014) and Spiegelhalter et al. (2011). Building upon experiences from weather forecasting, Fundel et al. (2019) highlight several recommendations for communicating probabilistic forecasts. It is important to explain probabilities as relative frequencies to a lay audience, for instance when presenting an 80% prediction interval: "In 80 out of 100 situations with a forecast like this …". At the same time, we should be aware of the limitations of probabilistic forecasts: people may misinterpret them, there may be a mismatch between the information one needs and the prediction interval, and too wide prediction intervals may cast doubt on the competence of the forecaster who produced them (Goodwin 2014). In any case, it is useful to remember the words of Bank of England's Governor Mervyn King. He said, in his Annual Lecture for the British Academy on 1 December 2004: " … in a wide range of collective decisions, it is vital to think in terms of probabilities … (W)e must accept the need to analyse the uncertainty that inevitably surrounds these decisions … (I)n order that public discussion can be framed in terms of risks, the public needs to receive accurate and objective information about the risks. Transparency and honesty about risks should be an essential part of both the decision-making process and the explanation of decisions." (King 2004). For demographic forecasting more specifically, we refer to UN Population Division Director John Wilmoth, who said in 2013: "... I expect that demographers will continue to be surprised by trends that do not follow our prior expectations. It is for this reason that the Population Division has worked hard in recent years to be more explicit and precise about the degree of uncertainty affecting projections of future population trends." (http://www.un.org/en/development/desa/news/population/population-division-director.html as of 6 December 2019).

1.9 A Brief Presentation of Chapters 2–12

In Chap. 2, Graziani proposes a procedure for deriving expert based stochastic population forecasts within the Bayesian approach. The joint distributions of all summary indicators are obtained based on evaluations by experts, elicited according to a conditional procedure that makes it possible to derive information on the centres of the indicators, their variability, their across-time correlations, and the correlations between the indicators. The forecasting method is based on a mixture model within the Supra-Bayesian approach that treats the evaluations by experts as data and the summary indicators as parameters. The derived posterior distributions are used as forecast distributions of the summary indicators of interest.

Chapter 3 by Dion, Galbraith, and Sirag also focuses on modeling experts opinions. Particular care is given to experts' opinions elicitation and their uncertainty quantification: experts are asked to provide estimates of 'most likely' values for a series of demographic indicators, along with corresponding 80% prediction intervals. A flexible distribution (metalog) is used to estimate experts' forecasts uncertainty for all components of population growth.

In Chap. 4, Castiglioni, Dalla-Zuanna and Tanturri evaluate the "convergence" hypothesis that is assumed by UNPD in several population revisions. They find out that in fact, such a convergence does not find empirical support, especially for life expectancy.

Chapter 5 by Aliverti, Durante, and Scarpa provides a data-driven model to forecast age-specific fertility rates (ASFRs). The model is based on a Gaussian process applied to a model of ASFRs. The latter is based on the skew normal distribution, a generalization of the normal distribution that allows for skewed shapes. The Gaussian process allows including model time dependent parameters, used to forecast future values of ASFRs. Forecasting ASFRs might be useful as in many cases forecasts of the TFR are available, but the age schedule is also needed to forecast the number of births.

Basellini and Camarda propose in Chap. 6 to analyse and forecast mortality developments over age and time by introducing a nonparametric decomposition of the mortality age pattern into three independent components corresponding to Childhood, Early-Adulthood and Senescence, respectively. Each component-specific death density is modeled with a relational model that associates a time-invariant standard to a series of observed distributions by means of a transformation of the age axis. This approach allows to capture mortality developments over age and time, and forecasts can be derived from parameters' extrapolation using standard time series models.

Chapter 7 by Bergeron-Boucher, Kjærgaard, Pascariu, Aburto, Alvarez, Rizzi, and Vaupel questions the assumption of linear (or log-linear) development of mortality indicators, such as death rates or life expectancy. This assumption can be problematic in countries where mortality development has been non-linear, such as in Denmark: the country experienced a stagnation of longevity improvement from the 1980s until the mid-1990s. The forecast performance of 11 models for Danish females and males and for period and cohort data are evaluated.

Chapter 8 by Booth focuses on coherent models, where a standard mortality pattern has to be defined. The chapter investigates the impact of different standards used in sex-coherent forecasts and standard-coherent ones. The analysis confirms that low mortality standards usually bring about lower bias, even though some exceptions, especially for males are found.

Chapter 9 by Keilman and Kristoffersen considers the uncertainty in mortality forecasts and analyses the extent to which life expectancy predictions for 2030 and 2050 were revised in subsequent rounds of population forecasts published by statistical agencies in selected countries. In a previous study, the conclusion was that life expectancy forecasts for some European countries for the year 2050 had been revised upwards systematically. Here they show that the period of upward revisions seems to have ended for some European countries.

Zhang and Bryant construct in Chap. 10 a forecasting model for internal migration, with an application to Iceland. The model proposed is a Bayesian hierarchical one. The motivation of using a hierarchical model stems from sparsity of data, which requires information borrowing, especially for flows characterized by low numbers.

Chapter 11 by Raymer, Bai, and Smith also considers internal migration, but the authors propose a log-linear model, which they apply to Australian regions. In particular, they show that multiplicative components can be used to capture the structure of migration flow tables. They combine the model with time series models to produce a hold-out sample of forecasts of interstate migration with measures of uncertainty. Goodness-of-fit statistics and calibration are then used to identify the best fitting models.

Scherbov and Sanderson consider in Chap. 12 a quite different matter: provided that demographic components are evolving over time (especially mortality), ageing could also be defined as an evolving concept. A prospective measure of ageing is considered. This measure could be based on remaining life expectancy or on mortality rates.

References

Ahlburg, D. (1982). How accurate are the U.S. Bureau of the Census projection of total live births? *Journal of Forecasting, 1*, 365–374.

Alders, M., & De Beer, J. (1998). Kansverdeling van de bevolkingsprognose. *Maandstatistiek van de Bevolking, 46*, 8–11.

Alexopoulos, A., Dellaportas, P., & Forster, J. J. (2018). Bayesian forecasting of mortality rates by using latent Gaussian models. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 182*(2), 689–711.

Alho, J. M. (2015). Population forecasting. In J. D. Wright (Ed.), *International encyclopedia of the social and behavioral sciences* (2nd ed., pp. 593–596). Amsterdam: Elsevier.

Alho, J. M., & Spencer, B. (1985). Uncertain population forecasting. *Journal of the American Statistical Association, 85*, 306–314.

Alho, J. M., & Spencer, B. (2005). *Statistical demography and forecasting*. New York: Springer.

Alho, J. M., Hougaard Jensen, S. E., & Lassila, J. (Eds.). (2008). *Uncertain demographics and fiscal sustainability*. Cambridge: Cambridge University Press.

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*, 815–839.

Bijak, J., & Bryant, J. (2016). Bayesian demography 250 years after Bayes. *Population Studies, 70*(1), 1–19. https://doi.org/10.1080/00324728.2015.1122826.

Bijak, J., Alberts, I., Alho, J., Bryant, J., Buettner, T., Falkingham, J., Forster, J. J., Gerland, P., King, T., Onorante, L., Keilman, N., O'Hagan, A., Owens, D., Raftery, A., Ševcíková, H., & Smith, P. W. F. (2015). Letter to the editor: Probabilistic population forecasts for informed decision-making. *Journal of Official Statistics, 31*(4), 537–544.

Booth, H. (2004). On the importance of being uncertain: Forecasting population futures for Australia. *People and Place, 12*, 1–12.

Booth, H. (2006). Demographic forecasting: 1980 to 2005 in review. *International Journal of Forecasting, 22*(3), 547–581.

Burch, T. K. (2003). The cohort-component population projection: A strange attractor for demographers. In J. Fleischhacker, H. A. de Gans, & T. K. Burch (Eds.), *Populations, projections and politics* (pp. 39–57). Amsterdam: Rozenberg Publishers.

Canale, A., Kenne Pagui, C. E., & Scarpa, B. (2016). Bayesian modeling of university first-year students' grades after placement test. *Journal of Applied Statistics, 43*(16), 3015–3029.

Cappelen, Å., Skjerpen, T., & Tønnessen, M. (2014). Forecasting immigration in official population projections using an econometric model. *International Migration Review, 49*(4), 945–980.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software, 76*(1). https://doi.org/10.18637/jss.v076.i01.

Cohen, J. (1986). Population forecasts and confidence intervals for Sweden: A comparison of model-based and empirical approaches. *Demography, 23*(1), 105–126.

Cohen, J. (2010). Population and climate change. *Proceedings of the American Philosophical Society, 154*(2), 158–182.

De Gans, H. (1999). *Population forecasting 1895–1945*. Dordrecht: Kluwer.

De Sherbinin, A., Carr, D., Cassels, S., & Jiang, L. (2007). Population and environment. *Annual Review of Environment and Resources, 32*, 345–373.

Dunson, D. B. (2001). Commentary: Practical advantages of Bayesian analysis of epidemiologic data. *American Journal of Epidemiology, 153*(12), 1222–1226. https://doi.org/10.1093/aje/153.12.1222.

Dunstan, K. (2019). *Stochastic projections: The New Zealand experience*. Paper presented at the Eurostat/UNECE Work Session on Demographic Projections, Belgrade, 25–27 November 2019.

Dunstan, K., & Ball, C. (2016). Demographic projections: User and producer experiences of adopting a stochastic approach. *Journal of Official Statistics, 32*(6), 947–962.

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. Chapman and Hall/CRC Press.

Fundel, V. J., Fleischhut, N., Herzog, S. M., Göber, M., & Hagedorn, R. (2019). Promoting the use of probabilistic weather forecasts through a dialogue between scientists, developers and end-users. *Quarterly Journal of the Royal Meorological Society, 145*(Supplement 1), 210–231.

Gerland, P., Raftery, A., Ševčíková, H., Li, N., Gu, D., Spoorenberg, T., Alkema, L., Fosdick, B., Chunn, J., Lalic, N., Bay, G., Buettner, T., Heilig, G., & Wilmoth, J. (2014). World population stabilization unlikely this century. *Science, 346*, 234–237.

Geweke, J., & Whiteman, C. (2006). Bayesian forecasting. In *Handbook of economic forecasting* (Vol. 1, pp. 3–80). Boston: Elsevier.

Gneiting, T., & Katzfuss, M. (2014). Probabilistic forecasting. *Annual Review of Statistics and Its Applications, 1*, 125–151.

Gneiting, T., & Raftery, A. (2007). Strictly proper scoring rules, prediction and estimation. *Journal of the American Statistical Association, 102*(477), 359–378.

Gneiting, T., Balabdaoui, F., & Raftery, A. E. (2007). Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society, Series B, 69*, 243–268.

Goodwin, P. (2014). Getting real about uncertainty. *Foresight – The International Journal of Applied Forecasting, 33*, 4–7.

Hajnal, J. (1955). The prospects for population forecasts. *Journal of the American Statistical Association, 50*(270), 309–322.

ISTAT. (2018). Il futuro demografico del paese: Previsioni regionali della popolazione residente al 2065 (base 1.1.2017). Report Statistische 3 maggio 2018. Roma: ISTAT.

Jackman, S. (2009). *Bayesian analysis for the social sciences*. Eastbourne: Wiley.

Keilman, N. (2020). Evaluating probabilistic population forecasts. Économie et Statistique/Economics and Statistics (to appear).

Keilman, N., & Cruijsen, H. (Eds.). (1992). *National population forecasting in industrialized countries* (NIDI-CBGS Publications 24). Amsterdam: Swets & Zeitlinger.

Keyfitz, N. (1972). On future population. *Journal of the American Statistical Association, 67*(338), 347–363.

Keyfitz, N. (1981). The limits of population forecasting. *Population and Development Review, 7*(4), 579–593.

Keyfitz, N. (1985). A probability representation of future population. *Zeitschrift für Bevölkerungswissenschaft, 11*(2), 179–191.

King, M. (2004). *What fates impose: Facing up to uncertainty*. Speech given by Mervyn King, Governor of the Bank of England at the Eighth British Academy Annual Lecture, London 1 December 2004. https://www.bankofengland.co.uk/speech/2004/what-fastes-impose-facing-up-to-uncertainty. (As of 6 December 2019).

Kuijsten, A. C. (1988). Demografische Toekomstbeelden van Nederland. Verkenningen no. 49 Planologisch en Demografisch Instituut, Universiteit van Amsterdam.

Land, K. (1985). Methods of national population forecasts: A review. *Journal of the American Statistical Association, 81*(396), 888–901.

Lee, R., & Tuljapurkar, S. (1994). Stochastic population forecasts for the United States: Beyond high, medium and low. *Journal of the American Statistical Association, 89*, 1175–1189.

Lee, R., & Tuljapurkar, S. (2000). Population forecasting for fiscal planning: Issues and innovations. In A. J. Auerbach & R. D. Lee (Eds.), *Demography and fiscal policy* (pp. 7–57). Cambridge: Cambridge University Press.

Li, N., & Lee, R. (2005). Coherent mortality forecasts for a group of populations: An extension of the Lee-Carter method. *Demography, 42*(3), 575–594. https://doi.org/10.1353/dem.2005.0021.

Lutz, W., & Scherbov, S. (1998). An expert-based framework for probabilistic national population projections: The example of Austria. *European Journal of Population, 14*, 1–17.

Lutz, W., Sanderson W. C., & Scherbov, S. (1996). *Probabilistic world population projections based on expert opinion* (IIASA Working Paper). IIASA, Laxenburg: WP-96-017.

Myers, R. J. (1954). Comparison of population projections with actual data. In *Proceedings world population conference 1954* (Vol. III, pp. 101–111). New York: United Nations.
National Research Council – NRC. (2000). Beyond six billion: Forecasting the World's population. Panel on population projections. In J. Bongaarts & R. Bulatao (Eds.), *Committee on population, commission on behavioral and social sciences and education*. Washington, DC: National Academy Press.
O'Neill, B., & Balk, D. (2001). World population futures. *Population Bulletin, 56*(3), 3–40. Washington: Population Reference Bureau.
Pollard, J. H. (1966). On the use of the direct matrix product in analysing certain stochastic population models. *Biometrika, 53*, 397–415.
Preston, S. H., Heuveline, P., & Guillot, M. (2001). *Demography: Measuring and modeling population processes*. Oxford: Blackwell Publishers.
Raftery, A. E. (2014). *Use and communication of probabilistic forecasts*. Mimeo, University of Washington. http://arxiv.org/abs/1408.4812. (As of 6 December 2019).
Raftery, A. E., Li, N., Ševčíková, H., Gerland, P., & Heilig, G. K. (2012). Bayesian probabilistic population projections for all countries. *PNAS, 109*, 13915–13921.
Raftery, A. E., Chunn, J. L., Gerland, P., et al. (2013). Bayesian probabilistic projections of life expectancy for all countries. *Demography, 50*(3), 777–801.
Raftery, A. E., Zimmer, A., Frierson, D. M. W., Startz, R., & Liu, P. (2017). Less than 2° C warming by 2100 unlikely. *Nature Climate Change, 7*, 637–641.
Rogers, A. (1995). *Multiregional Demography*. Chichester: Wiley.
Rowland, D. (2003). *Demographic methods and concepts*. Oxford: Oxford University Press.
Sanderson, W. (1998). Knowledge can improve forecasts: A review of selected socioeconomic population projection models. *Population and Development Review, 24*(Supplement), 88–117.
Schweder, T. (1971). The precision of population projections studied by multiple prediction methods. *Demography, 8*(4), 441–450.
Ševčíková, H., Li, N., Kantorová, V., Gerland, P., & Raftery, A. (2016). Age-specific mortality and fertility rates for probabilistic population projections. In R. Schoen (Ed.), *Dynamic demographic analysis* (pp. 285–310). Cham: Springer.
Shang, H. L. (2015). Statistically tested comparisons of the accuracy of forecasting methods for age-specific and sex-specific mortality and life expectancy. *Population Studies, 69*(3), 317–335.
Shang, H. L., & Hyndman, R. (2017). Grouped functional time series forecasting: An application to age-specific mortality rates. *Journal of Computational and Graphical Statistics, 26*(2), 330–343.
Shang, H. L., Smith, P., Bijak, J., & Wisniowski, A. (2016). A multilevel functional data method for forecasting population, with an application to the United Kingdom. *International Journal of Forecasting, 32*, 629–649.
Simpson, L., Wilson, T., & Shalley, F. (2019). The shelf life of official sub-national population forecasts in England. *Applied Spatial Analysis and Policy*. Published online 21 November 2019. https://doi.org/10.1007/s12061-019-09325-3. (As of 6 December 2019).
Spiegelhalter, D., Pearson, M., & Short, I. (2011). Visualising uncertainty about the future. *Science, 333*, 1393–1400.
Statistics New Zealand. (2011). National Population Projections: 2011(base)–2061. Bulletin published 19 July 2012, ISSN 1178-0584. http://archive.stats.govt.nz/browse_for_stats/population/estimates_and_projections/NationalPopulationProjections_HOTP2011.aspx. (As of 6 December).
Stoto, M. A. (1983). The accuracy of population projections. *Journal of the American Statistical Association, 78*, 13–20.
Sykes, Z. M. (1969). Some stochastic versions of the matrix model for population dynamics. *Journal of the American Statistical Association, 64*, 111–130.
Tetlock, P., & Gardner, D. (2016). *Superforecasting: The art and science of prediction*. New York: Crown.
Törnquist, L. (1949). Om de synspunkter, som bestämt valet av de primära prognos-antaganda. In J. Hyppolä, A. Tunkelo, & L. Törnquist (Eds.), *Beräkningar rörande Finlands Befolkning, dess*

Reproduktion och Framtida Utveckling (Statistiska Meddalanden 38) (pp. 69–75). Helsinki: Statistiska Centralbyrån. (in Finnish, Swedish and French).
Tuljapurkar, S. (1992). Stochastic population forecasts and their uses. *International Journal of Forecasting, 8*(3), 385–391.
United Nations. (2017). Department of Economic and Social Affairs, Population Division. *World population prospects: The 2017 revision, methodology of the United Nations population estimates and projections* (Working Paper No. ESA/P/WP.250). New York: United Nations.
United Nations Economic Commission for Europe – UNECE. (2018). *Recommendations on communicating population projections* (Report no. ECE/CES/STAT/2018/1). Geneva: United Nations.
Willekens, F. J. (1990). Demographic forecasting: State-of-the-art and research needs. In C. Hazeu & G. Frinking (Eds.), *Emerging issues in demographic research* (pp. 9–66). Amsterdam: Elsevier.
Wilson, T. (2018). *Communicating population forecast uncertainty using perishable food terminology*. Darwin: Northern Institute, Charles Darwin University (Research Brief RB03/2018). http://www.cdu.edu.au/sites/default/files/research-brief-2018-03_0.pdf. (As of 6 December 2019).
Wilson, T., & Rees, P. (2005). Recent developments in population projection methodology: A review. *Population, Space and Place, 11*, 337–360.
Wilson, T., & Shalley, F. (2019). Subnational population forecasts: Do users want to know about uncertainty? *Demographic Research, 41*(13), 367–392.
Wilson, T., Brokensha, H., Rowe, F., & Simpson, L. (2018). Insights from the evaluation of past local area population forecasts. *Population Research and Policy Review, 37*(1), 137–155.