Draft, March 24, 2004

Population Scenarios Based on Probabilistic Projections: An Application for the Millennium Ecosystem Assessment

Brian O'Neill
International Institute for Applied Systems Analysis

Presented at the 2004 Annual Meeting
of the Population Association of America,
Boston, Mass., USA
April 1-3, 2004

0. Abstract

Probabilistic population forecasts offer a number of advantages to users. However, in some cases population is one component of a larger analysis that may take a different approach to uncertainty. For example, integrated assessments of environmental issues such as climate change or ecosystem degradation have typically used a small number of alternative scenarios to explore uncertainty in future environmental outcomes. In such cases, population projections that are provided only as probability distributions are difficult to use. I present a method of employing probabilistic population projections to derive individual, deterministic projections that can be used within scenarios for integrated assessments. The principal advantages of this approach are that (1) it provides a less ad hoc way of defining deterministic projections intended to be consistent with more comprehensive scenarios that describe, among other things, future socio-economic developments; (2) it provides more flexibility in specifying input assumptions for deterministic projections as compared to choosing off-the-shelf projections; and (3) it provides a quantitative assessment of the uncertainty associated with any given deterministic projection. I describe the application of the method to the development of population projections used in integrated scenarios for the Millennium Ecosystem Assessment, an international scientific effort to assess the current conditions of and future outlook for global ecosystem goods and services. Results show that the MA scenarios are each consistent with a relatively wide range of demographic outcomes. For some scenarios, ranges of plausible outcomes in particular regions overlap substantially, indicating that particular population projections could be consistent with more than one scenario. In other cases, uncertainty ranges for different scenarios are distinct, indicating that a projection consistent with one scenario is unlikely to be also consistent with another. Comparing variances of the conditional projections also provides insight into how much different storylines constrain future demographic developments. The development of the MA projections also points to important areas of future research in correlations involving demographic rates and uncertainty across scales.

1. Background

The development and analysis of alternative scenarios as an approach to uncertainty has a history of at least 50 years, beginning with post-World War II military planning, extending to business strategy development for major corporations, and more recently to planning for sustainable development (Schwartz, 1991). Scenarios are stories about the way the future might unfold. They can be qualitative, quantitative, or both, and they can be used in a variety of ways, including for educating participants in the scenario development process about the issues at hand, for communicating key insights to the intended audience, or in decision analyses which search for response options robust in the face of many possible future development paths. Scenarios tend to be used as an approach to uncertainty when problems are complex and uncertainties are very large, precluding meaningful estimates of the likelihood of various future outcomes.

Prominent examples of the scenario approach within environmental assessments include the range of alternative scenarios for future greenhouse gas emissions developed by the Intergovernmental Panel on Climate Change (Nakicenovic et al, 2000), an assessment of the outlook for the global environment by the UN Environment Programme (GEO-3), and scenarios for ecosystem goods and services and human well-being developed for the Millennium Ecosystem Assessment (MA). Each of these examples used an approach that combined qualitative and quantitative elements. First, a small number of different qualitative “storylines” were developed, describing in narrative form broad socio-economic and technological development patterns that could unfold over the 21st century. Next, particular quantitative paths for fundamental driving forces of the environmental issues under consideration (e.g., population, economic growth, rates of technological change) were selected that were judged to be consistent with each storyline. Finally, a number of different modeling teams produced quantitative interpretations of the storylines (in terms of, for example, quantities of greenhouse gases emitted, changes in land use, supply of and demand for fresh water, etc.), using the agreed upon paths for driving forces as inputs. Thus, producing quantitative scenarios for environmental outcomes requires quantitative scenarios for driving forces, including population.

Available population projections typically do not meet the needs of integrated assessments very well. Sets of alternative population projections are certainly available: they have been produced by many institutions for many years, most visibly by the United Nations. However, such “off the shelf” projections have a number of shortcomings for use in integrated assessment, many of them related to the fact that there are typically only a small number of projections available to choose from, and the assumptions they are based on may not match very well the storyline being used in the integrated scenarios. For example, the U.S. Census Bureau and the World Bank both produce only a single, best guess population projection that assumes moderate levels of fertility, life expectancy, and migration in the future. These projections are therefore unsuitable for any integrated scenario that calls for relatively high or relatively low future paths for any of the components of population change. The UN scenarios (UN, 2003) use three alternative fertility variants (high, medium, or low) but only one variant for mortality and migration. Thus while they provide more choices in terms of future fertility trends, it is not possible to use the UN projections for integrated scenarios that call for mortality or migration trends that deviate from their best guess pathways. Furthermore, even the three alternative fertility paths used in the UN projections are somewhat constraining. Each assumes that fertility is high, medium, or low in all countries at the same time, and over the entire projection period. It is impossible then to use these projections in an integrated assessment scenario calling for, for example, low fertility (i.e., fertility lower than in a best guess case) in industrialized countries, but high fertility in developing country regions, or low fertility early in the century but high fertility later on.

Previous IIASA projections (Lutz et al., 1996) provided a large number of scenarios, including options for mixing assumptions across regions and across components of change (although not over time). Largely for this reason, two of these scenarios were used in the IPCC emissions scenario work. However, this practice was not continued in the most recent IIASA projections (Lutz et al., 2001), which provide output only in terms of distributions, and not as specific scenarios.

A second reason that off-the-shelf population projections fall short for use in integrated scenarios is that they lack an assessment of uncertainty associated with a given scenario. For example, for an integrated scenario that assumes that future socio-economic conditions will unfold in a manner that would be consistent with relatively high fertility in developing countries, it is possible to find an off-the-shelf population projection that meets this need. But such a projection would not indicate what range of population outcomes (size and age structure) would be consistent with relatively high fertility. Since one of the main reasons for carrying out integrated scenario analysis is to explore the implications of uncertainty, this is an important shortcoming. When only a single projection is used to represent all high fertility futures, for example, much of the range and diversity of plausible outcomes is lost.

Sanderson et al. (2003) suggest that conditional probabilistic projections are a means of using probabilistic projections in a manner similar to scenarios. Conditional probabilistic projections produce distributions of output based on input distributions to which some constraints have been applied. For example, Sanderson et al. (2003) use this approach to measure the sensitivity of population size and age structure to both fertility and mortality assumptions within a probabilistic context. O’Neill (2003) applies the conditional probabilistic projection approach to an integrated environmental assessment (greenhouse gas emissions), producing probabilistic emissions projections conditional on a number of scenario storylines. However in both these cases outcomes are still probability distributions (albeit conditional ones), not single deterministic projections needed for traditional scenario analysis.

Here I build on these previous approaches by developing conditional probabilistic projections consistent with environmental storylines for the Millennium Ecosystem Assessment scenarios, and also defining a single deterministic projection that is representative of each conditional probabilistic projection. The first advantage of this methodology is that the representative, deterministic projection can be tailored to the assumptions underlying the MA scenarios, therefore avoiding the primary problem associated with using off-the-shelf population projections. The second advantage is that the conditional probabilistic projections provide a quantitative assessment of the uncertainty associated with each representative deterministic projection.

2. Methodology

The IIASA probabilistic projections (Lutz et al., 2001) serve as the basis for developing deterministic population scenarios. These projections are made at the level of 13 world regions, and were produced by defining subjective probability distributions for the components of population change (fertility, mortality, and migration) for each region, and by specifying correlations among components, across regions, and over time. Probability distributions for each component were arrived at through expert opinion, with consideration of errors in past projections and especially of a recent assessment of projections carried out by the U.S. NAS (Bongaarts and Bulatao, 2000). Uncertainty distributions for outcomes were derived by carrying out thousands of individual projections,[1] each one driven by randomly selected paths for fertility, mortality, and migration, using the probability distributions for inputs and the correlation matrix as a constraint, and compiling the results into uncertainty distributions for population by age and sex and by region of the world over time.

There are several possible ways to derive single population projections consistent with a given storyline from probabilistic projections generated in this manner. The approach that may appear to be the simplest and most direct would be to search the database of individual simulations that underlies probabilistic projections, and select one that matches the trends assumed to be consistent with the storyline. However in practice this turns out to be quite difficult. First, single projections that meet even relatively broad criteria may be rare. Since these criteria generally involve specifying trends in three input variables (fertility, mortality, and migration), typically by requiring them to fall into one of at least three categories (high, medium, or low), for many regions of the world, and potentially with variations over time (e.g. fertility high over the first half of the century, then low over the second half), the size of the database required to ensure that it contains at least one scenario meeting all criteria is very large. For example, assuming that criteria for a scenario specify the three components of change as falling into one of three equally probable categories for four separate regions of the world, and assuming no correlations among components or regions, one would need $(3^3)^4 = 531{,}441$ scenarios to generate examples for all possible permutations. Non-zero correlations can greatly reduce the size of the database needed to represent the more likely combinations of assumptions, but nonetheless plausible scenarios may be entirely absent from a database of a few thousand or even tens of thousands of simulations.

In addition, even if a single scenario that did meet the criteria for a storyline could be identified, it may still be problematic because it may have idiosyncrasies. The IIASA projection methodology generates realistic fluctuations in fertility, mortality and migration over time, which is important to generating realistic uncertainty ranges for outcomes. However in the context of an integrated scenario that is not taking a probabilistic approach, idiosyncratic scenarios can present problems. The aim in scenario analysis is to have a representative scenario in which the outcome is being driven by the criteria used to develop it. A population projection whose outcomes is driven not only by the criteria used to select it (e.g., that fertility should be relatively high), but also by, for example, a temporary baby boom that occurs in a particularly important region, complicates scenario analysis. In order to draw useful lessons from the scenario exercise as a whole, the effect on outcomes of the demographic idiosyncrasy will have to be disentangled from the effect of all other driving forces.

A second possible approach to deriving a single scenario would be to select a particular percentile of the unconditional distribution to use as a deterministic projection. This could be

---

[1] The projections presented in Lutz et al. (2001) were based on 2,000 simulations. For the current study, S. Scherbov provided 10,000 simulations generated in an identical manner.

done based on the output distribution (e.g., selecting a percentile of population size for each region), or based on the input distributions (selecting particular percentiles of fertility, mortality, and migration for each region). This method could be a useful way of generating a single projection, but it would not help in characterizing uncertainty. In addition, the selection of a particular percentile would not be straightforward, because actual paths underlying uncertainty distributions (for outputs or inputs) do not follow percentiles. Thus, for example, there is substantially less than a 20% chance that population will remain above the 80th percentile of the projected global population size in the IIASA projections; due to imperfect autocorrelation in vital rates, population paths frequently cross percentile lines over time.

Third, a conditional probability distribution for each input could be defined based on the storyline, and then a percentile of these conditional distributions could be selected as the single representative scenario. I pursue this option here. By defining single projections relative to conditional probability distributions, a framework is put in place for quantifying the uncertainty associated with the representative scenario. It allows communicating the range of possible demographic outcomes that would be consistent with the storyline, while at the same time providing a single scenario for use in the analysis.

The development of the population projections begins with defining qualitative criteria based on the MA storylines; next, conditional probabilistic projections are defined, and finally single deterministic projections are specified.

2.1 Qualitative assumptions based on the MA storylines

The MA scenarios are based on four different qualitative storylines (Alcamo et al., in review). While the scenarios are differentiated along a number of dimensions, there are two dominant axes: the approach taken to environmental management (proactive vs. reactive), and the degree of connectedness of institutions across and within regions (globally connected vs. regional focus). Within this broad framework, storylines also differ on the form of capital with which society is most concerned (natural, human, manufactured, or social), the presumed substitutability of these forms of capital, and the favored types of macroeconomic policies and approaches to global pollutants and property rights.

Two scenarios describe globally connected worlds. The *Global Orchestration* (GO) scenario describes a globally connected world that takes a reactive approach to environmental management. There is a focus on social policy, and investments in human, manufactured, and social capital are high. Investment in natural capital is low, unless it is seen to matter for development. The *Technogarden* (TG) scenario is also globally connected, but takes a proactive approach to environmental management with a heavy emphasis on technological solutions to ecological problems. Investments in all types of capital are reasonably balanced. Two additional scenarios are regionally, rather than globally, focused. The *Order from Strength* (OS) scenario is driven by a concern for security that turns regions inwards toward their own concerns, and responses to environmental problems are reactive. Inequality is high within and among regions, with investments in human, manufactured and social capital high among elites but low when directed toward non-elites. Finally, the *Adapting Mosaic* (AM) scenario describes a regionally oriented world that makes strong investments in social capital through encouragement of local institutions, and that responds proactively to environmental issues. Investments in human and natural capital are high, there is a lot of experimentation across regions with alternative approaches to managing environmental issues, and lessons learned in one part of the world are communicated and adopted elsewhere.

Table 1 lists the qualitative assumptions about fertility, mortality and migration for each storyline, decided upon by the MA Scenarios Working Group with input from IIASA demographers.[2] These assumptions are expressed in terms of categories High/Medium/Low (H/M/L), defined not in absolute but in relative terms. That is, a high fertility assumption for a given region means that fertility is assumed to be high relative to the median of the probability distribution for future fertility in the IIASA projections. Since the storylines describe events unfolding through 2050, the demographic assumptions specified here apply through 2050 as well. For the period 2050-2100 assumptions were assumed to remain the same, in order to gauge the consequences of trends through 2050 for the longer term. This is not intended to reflect any judgment regarding the plausibility of trends beyond 2050.

Fertility and mortality in currently high fertility countries: Trends in these rates were selected based on demographic transition reasoning. In Global Orchestration (GO), higher investments in human capital (especially education and health) and greater economic growth rates are assumed to be associated with a relatively fast transition, implying that lower fertility and mortality would be likely in this scenario, relative to a central or best-guess outlook. In Order from Strength (OS), it is assumed that lower investments in human capital and slower economic growth are more likely to be associated with a slower transition (i.e., higher fertility and mortality). Technogarden (TG), with more moderate investments and economic growth assumptions, is assumed to undergo a moderate pace of change in both fertility and mortality. The Adapting Mosaic (AM) storyline begins similarly to Order from Strength and then diverges later as learning occurs and successful development strategies are replicated in other regions. Demographic trends are therefore specified to follow Order from Strength for 10 years, and then diverge to medium assumptions by mid-century.

Fertility in currently low fertility countries: The knowledge of determinants of long-term trends in fertility in low fertility countries is limited, and therefore there is little basis for preferring one set of assumptions to another for a given storyline. In the face of this uncertainty the overarching rationale for specifying trends for given storylines was chosen to be the scope of convergence in fertility across low fertility countries. Since Order from Strength describes a regionalized, divergent world, and Global Orchestration a globalizing, convergent world, these characteristics were applied to future fertility. Thus the low fertility countries were divided into two groups (one with Very Low Fertility, one with Low Fertility, see note to Table 1), and fertility assumptions made such that fertility in these two groups would tend to converge in the GO scenario and diverge in the OS scenario. In the Adapting Mosaic scenario, fertility initially follows the OS assumptions, and then diverges toward medium levels. In the Technogarden scenario, medium fertility is assumed.

Mortality in Industrialized country regions: Mortality was assumed to be lowest in the Global Orchestration scenario, consistent with its high economic growth rates, technological progress that is assumed to occur in the health sector as well, and reduction in inequality within the region. In contrast, Order from Strength, which assumes growing inequality within the industrialized countries and even the potential for re-emergence of some diseases, is assumed to have the highest mortality. Technogarden assumes a medium pace of mortality change, and Adapting Mosaic follows the OS assumptions for 10 years before diverging to medium levels in 2050.

---

[2] I thank Wolfgang Lutz and Sergei Scherbov for their valuable input on this task.

Migration: Net migration rates are assumed to be low in the regionally oriented scenarios (Adapting Mosaic and Order from Strength), consistent with higher barriers between regions in OS and strong regional ties in AM. In Global Orchestration, permeable borders and high rates of exchange of capital, technology, and ideas are assumed to be associated with high migration. Technogarden assumes a more moderate migration level.

2.2 Defining conditional probability distributions for projection inputs

To define conditional distributions for inputs for each region, individual simulations from a database of 10,000 simulations were selected that matched the qualitative criteria defined in Table 1 for three of the MA scenarios.[3] Clearly, taking the step from qualitative descriptions of demographic trends to quantitative ranges of plausible values involves many subjective judgments. For developing projections for the MA scenarios, our aim was to be as simple and transparent as possible, keeping in mind that the overall goal was to provide a traceable account of the definition of the deterministic projections, as well as a sense of the range of possible outcomes that might be consistent with a given set of rather broad, qualitative characteristics of future demographic developments. Even with a conscious focus on simplicity, the process inevitably involved specification of a number of methodological details (described below), beginning with the construction of an index for each demographic variable that could be used to compare different simulations.

2.2.1 Index

High, medium, and low categories were defined such that they contained those simulations that fell into three evenly divided quantiles of the unconditional distribution for each variable. The basic metric for each variable was a time average. The rationale for using a time average is that the qualitative demographic characteristics judged to be consistent with the storylines describe general tendencies over time, but do not rule out fluctuations. In addition, time averages are a good indicator of long-term population outcomes for variables such as population size that integrate over input variables. For example, a large population size is more likely to result from a fertility path with a high average level than from one with a high level only at the end of the period, and the average level will be a better predictor of population size.

Selecting an averaging period requires several considerations, including the timescale of the trends described in the storyline, and the persistence of the variable (i.e., the degree of autocorrelation). The relationship between these two timescales is a determinant of the variance of the conditional distribution. A variable with a persistence timescale that is short compared to the averaging period will have a large variance. As a result, for much of the averaging period a path that is categorized as “high” may actually have a relatively low value (during its many temporary excursions to lower levels). The proportion of the averaging period that a variable should spend within a particular sub-range of the unconditional distribution is a user choice: does the storyline allow for possible fluctuations, and if so of what size and duration?

For developing the projections for the MA, 50-year averaging periods were used. The 50-year timescale matches the timescale of the MA storylines. The storylines do not include

---

[3] The projection for the AM scenario was developed in a different manner, not using conditional probabilistic projections; it was defined as following the assumptions for the OS scenario for 10 years, then deviating afterwards to medium levels of all input variables (i.e., the TG scenario) by 2050.

substantial fluctuations in socio-economic conditions over timescales less than 50 years, making a shorter averaging period unnecessary for selecting simulations consistent with the storylines. A longer averaging period would also be unsuitable because it would allow some simulations to be categorized as, for example, high fertility pathways even if fertility followed a path more consistent with a medium or even low category before 2050, and did not become high until after 2050. Since the MA storylines are primarily defined over the period 2000-2050, this would be an undesirable outcome; high fertility should mean generally high over the first half of the century.

The qualitative demographic assumptions are extended to 2100 in order to assess the implications of the first 50 years for the outlook for the end of the century. Therefore, to maintain the quantitative conditional distributions over the entire century, I assume that any path for a vital rate that is categorized as, say, "high" should fall in the top one-third of all simulations when ranked by 50 year average values, for any 50 year period one might select within the period 2000-2100. For practical purposes, I use three overlapping periods (2000-2050, 2025-2075, 2050-2100) as an approximation.

A further consideration is that time averages introduce boundary effects: near boundaries of averaging periods, conditional distributions will tend to revert toward the unconditional distribution. These effects could have large consequences by, for example, diminishing the differentiation across different conditional distributions near the end of the averaging period. The use of overlapping averaging periods prevents this from happening at the end of the 2000-2050 averaging period. To address boundary effects near 2000 and 2100, I apply additional criteria for averages over shorter periods (2000-2015, 2085-2100).[4] These are applied, however, not to match storyline assumptions, but simply to maintain the variance of the conditional distribution near the boundaries at a level similar to its value over the middle of the century.

2.2.2 Regional aggregation

The selection process also requires adopting an appropriate level of regional aggregation. Vital rates could be averaged not only over time, but also across multiple regions. For example, rates could be averaged across all currently low-fertility regions, or all developing country regions. Again, this is a user choice. A scenario with mortality that is low averaged over all developing country regions could have relatively high mortality in a (population-weighted) minority of subregions. An alternative would be to apply criteria to each subregion individually, which would ensure that mortality is low everywhere. The choice of approach depends on the application.

Here I apply criteria to each IIASA region individually. This approach also addresses the issue of correlation across regions. By treating projections for each region independently, I define conditional distributions for each region assuming no inter-regional correlations in fertility, mortality, or migration.

The selection process therefore consists of producing an average fertility, mortality, and migration rate for each of 10,000 simulations in the unconditional projection, with averaging performed over the period 2000-2050, 2025-2075, 2050-2100, as well as 2000-2015, and 2085-

---

[4] In particular, simulations must also fall above the median (for "high" paths), below the median (for "low" paths), or within the interquartile range (for "medium" paths) over the periods 2000-2015 and 2085-2100. These ranges and time periods were selected by trial and error in order to produce a constant variance for the conditional distributions of fertility mortality, and migration.

2100, for each region. Simulations are then ranked according to these indexes in each region, and assigned to high, medium, and low categories by dividing each ranking into three sets of outcomes containing the first 3333, next 3334, and last 3333 simulations. Those simulations that fall in the high category for each of the averaging periods are classified as high overall (and likewise for medium and low), and form the conditional distributions of the input rates. Finally, simulations that match the particular combination of assumptions for all three input rates for each MA storyline as defined in Table 1 become part of the conditional probabilistic projection for the population outcomes consistent with that storyline. This approach produces what might be called input-constrained conditional probabilistic projections, since the conditions (storylines) impose constraints on the uncertainty distributions of the demographic rates that serve as inputs to the projection.

2.3 Specifying representative deterministic projections

Deterministic projections are specified by selecting the median of the conditional distributions for each input variable, for each region, associated with the storylines. For example, the GO storyline assumes low fertility, low mortality, and high migration for developing country regions. For each developing country region in the IIASA projections, the median of the "low" conditional distributions for fertility and mortality, and the "high" conditional distribution for migration, is defined as the deterministic scenario for the single population projection for the GO scenario. This process is repeated for all regions and all MA scenarios. These deterministic input assumptions are then used to drive individual deterministic population projections.

Note that it is not necessary that deterministic projections be based on the medians of the conditional input distributions. One could just as well take a higher, or lower, percentile to use as the deterministic path. This is a user choice. Here, the aim was to develop deterministic projections that were most representative of the conditional distributions, and therefore the median seems a logical choice. However one might wish to choose a percentile above or below the median to satisfy other possible criteria for scenario development. For example, if one aim of the scenario analysis was to span a wide range of plausible population outcomes, it might be preferable to choose percentiles above or below the median that would produce more extreme outcomes.

3. Results

Table 2 shows the number of simulations (out of a total of 10,000) that satisfy the criteria for fertility, mortality, and migration, individually and in various combinations, for selected regions and scenarios. The table illustrates several distinctive features of the conditional distributions. On average, 10% or less of simulations satisfy the definition of "high" for any single demographic rate, and similar percentages fall into "medium" or "low" categories. The reason that such a small percentage of simulations satisfy these criteria is the requirement that each rate be sustained at a particular level over time (achieved by using overlapping averaging periods as selection criteria). For example, while by definition one third of all simulations fall into the high category for fertility averaged over the period 2000-2050, many fewer fall into the high category in succeeding 50-year periods as well. Based on the autocorrelation of rates assumed in the IIASA projections, most simulations produce fertility paths that drift up and down substantially, and do not remain within a given tercile of the unconditional distribution over time.

Furthermore, many fewer simulations meet the criteria for individual rates in the TG scenario relative to other scenarios, and also for fertility in Western Europe in the GO scenario. This occurs because the TG scenario assumes medium levels for each demographic rate for all regions, and the GO scenario assumes medium fertility for Western Europe (and three other regions not shown in the table). Because the unconditional uncertainty for each rate is assumed to be normally distributed in the IIASA projections, dividing the distribution of 50-year averages into three equally probable quantiles does not produce quantiles with equal absolute ranges. The middle quantile, used to define ranges for "medium" projections, covers a smaller absolute range (where the probability density function is more concentrated). For example, in 2050, the 80% uncertainty interval for fertility in North Africa has a range of 0.58 births per woman in the medium conditional distribution, while the range is 0.77 and 0.71 births per woman in the high and low distributions, respectively. As a result, even though it is equally likely that a simulation will fall into any one of the three categories for a given 50-year period, it is less likely that a simulation will remain in the medium category over successive 50-year periods than it will remain in a high or low category over successive periods.

Finally, typically less than 10 simulations out of 10000 meet all three criteria (for fertility, mortality, and migration) simultaneously. This is a simple consequence of the independence across vital rates assumed in the IIASA projections combined with the relative scarcity of simulations meeting the criteria for each rate individually. When these simulations are particularly scarce (as in the case of the TG scenario), there can even be no simulations at all that meet all three criteria simultaneously. Since it is precisely those simulations that meet several criteria simultaneously that constitute the distribution of outcomes in the conditional probabilistic projections, this feature of the results presents a particular challenge discussed further below.

Figure 1 shows selected percentiles of the unconditional and conditional distributions for the total fertility rate for the North Africa and Western Europe regions. Note that the distributions for the three MA scenarios considered together span a very wide range of the unconditional distribution, despite the fact that combined they consist of less than 30% of the total number of simulations. Roughly speaking, in North Africa the distribution for the OS scenario (high fertility) has an 80% uncertainty interval extending from the median to well above the 90th percentile of the unconditional distribution. The TG scenario (medium fertility) has a conditional distribution with an 80% interval extending from about the 30th to about the 70th percentile of the unconditional distribution. The GO scenario (low fertility) has an 80% interval that extends from the median to well below the 10th percentile of the unconditional distribution. In Western Europe, there is no MA scenario that assumes relatively high fertility, so the upper end of the unconditional distribution is not represented, but the scenarios thoroughly cover the middle and lower end of the unconditional range.

Conditional distributions based on different assumptions about demographic trends overlap substantially, even though they consist of completely distinct sets of simulations. Overlap occurs because the selection of simulations, as described above, was based on indexes of fertility averaged over time, while the distributions in Fig. 1 are of fertility at each point in time. The fertility fluctuations in individual simulations lead to the substantial overlaps of the conditional distributions. As a consequence, at any given time a particular fertility rate in a given region may be consistent with more than one storyline, but time-averages of fertility become more and more consistent with one storyline versus others as the averaging period approaches 50 years.

Fig. 2 shows the conditional and unconditional distributions of population size for North Africa and Western Europe, along with the deterministic projections for each region used in the MA scenarios. The results for North Africa are based on simulations that meet fertility and mortality criteria for each scenario, but migration criteria are not used in order to ensure that there are enough simulations to characterize the uncertainty distribution (see Table 2). Eliminating migration as a selection criterion does not affect results too much, given that migration has a relatively small effect on population size in this region. One way to see this is to note that the deterministic projections (which take each scenario's migration assumptions into account by using the median of the conditional migration distributions as an input to the projection) differ very little from the median of the population size distributions (which is based on simulations that do not constrain migration assumptions).

Results show that the deterministic projections are typically associated with a substantial range of plausible outcomes that would still be consistent with the storylines. For example, in the OS scenario (high fertility and mortality), the deterministic projection reaches 445 million in the year 2100. The 80% uncertainty interval for this scenario extends from 365 to 553 million, or an uncertainty of about +/- 20%. In the GO scenario (low fertility and mortality), the 80% interval is smaller in absolute terms (202 to 327 million in 2100) but somewhat larger in percentage terms (closer to +/- 25%). The two distributions have little overlap: their 80% uncertainty intervals do not intersect over most of the century. This indicates that the demographic assumptions underlying them, even accounting for uncertainty in their quantification, lead to distinctly different demographic outcomes. In a sense, comparing their conditional distributions demonstrates that the two deterministic projections developed for these scenarios are significantly different from each other. This result can be used to conclude that it is unlikely that a single population outcome would be consistent with both the OS and the GO scenario.

In contrast, in the TG scenario (medium fertility and mortality), population outcomes overlap substantially with the GO scenario. The deterministic projection for the TG scenario lies at about the 90th percentile of the GO conditional distribution. It is thus not implausible that the deterministic projection used in the TG scenario could occur even in the GO scenario. The overlap in the conditional distributions indicates that population outcomes that lie between the deterministic TG and GO projections could plausibly occur under the assumptions in either scenario.

The variance of the TG distribution is substantially smaller than for the other two scenarios (310 to 361 million 80% uncertainty interval, or about +/- 8% around the deterministic projection). This is a result of the narrower (in absolute terms) conditional distributions of fertility and mortality when they are assumed to fall in the "medium" category, as discussed above.

The conditional probabilistic projections for W. Europe are based on simulations that meet the criteria for mortality and migration, but not for fertility. Fertility is not used as a selection criterion because we assume that the storyline does not constrain possible future fertility paths in the currently low fertility regions (see section 2.1 above). We selected particular high, medium or low paths for fertility to use within each scenario in these regions based on a logic related to matching convergence or divergence of demographic rates with the degree of openness across borders within the storylines. However, this approach was intended to be an organizing principle for developing the population projections, not a judgment on the likelihood of the matching between fertility trends and openness of borders. Rather, it was judged that the storylines did not lead to a preference for any particular range of fertility within the unconditional distribution. Therefore, we use the full unconditional range of fertility to define the conditional distribution of population outcomes.

Results for this region have similar features to the results for North Africa, but show a few novel characteristics. Most strikingly, the deterministic projection for the OS scenario does not lie near the median of its conditional distribution, but rather lies closer to the 10th percentile. This occurs because the deterministic projection assumes low fertility, but the conditional population size distribution assumes the full range of uncertainty in fertility, thus producing a wide range of generally higher population outcomes. (It does not occur in the other scenarios because they both assume medium fertility.) This result helps put the projection used in the OS scenario in context: a choice was made to specify this scenario as low fertility (and therefore relatively low population size) independent of any judgment about the degree to which the storyline would be more or less likely to be associated with such an outcome. Based on the storyline alone, the uncertainty in outcomes is rather large, and we should keep in mind that a substantially larger population size than the one used in the MA scenario analysis could well be consistent with the scenario.

Given the generally large variances of the conditional distributions of population size in this region (due to the unconstrained fertility assumptions), there is more overlap across scenarios. As a consequence, any particular population outcome could plausibly occur under more than one of the scenarios.

Figure 3 shows the deterministic and conditional probabilistic projections of population size for the world [to come; currently figure 1 shows only the deterministic MA projections]. The uncertainty distributions were created by randomly combining the simulations making up the conditional distributions for each region. This assumes no correlation in rates across regions. Since there is likely a positive correlation across regions (as assumed, e.g., in the IIASA global projections), these uncertainty distributions must be considered a lower bound. A second reason they are a lower bound is that some regions do not have many simulations available that meet the criteria for particular scenarios, and therefore the variance of the distribution of population size may be underestimated by this small sample size.

Discussion and conclusions

This analysis demonstrates the approach of defining individual, deterministic population projections based on conditional probabilistic projections. The principal advantages of this approach are that (1) it provides a less ad hoc way of defining deterministic projections intended to be consistent with more comprehensive scenarios that describe, among other things, future socio-economic developments; (2) it provides more flexibility in specifying input assumptions for deterministic projections as compared to choosing off-the-shelf projections; and (3) it provides a quantitative assessment of the uncertainty associated with any given deterministic projection. An application to the development of projections for use in the Millennium Ecosystem Assessment scenario illustrates the approach in the context of a major international scientific assessment of an interdisciplinary issue: the impact of human activity on ecosystem goods and services and, in turn, the possible consequences of those impacts for human well being. The projections developed for the MA demonstrate that within any given storyline, a range of demographic outcomes is possible. Reporting these conditional distributions of outcomes, as well as single projections, can communicate in quantitative terms the uncertainty inherent in such an exercise. For example, the larger uncertainty ranges for population outcomes in currently low fertility regions can reflect more transparently the fact that less is currently known about how future socioeconomic conditions may influence demographic developments in these regions, relative to knowledge in high fertility countries where demographic transition theory provides a more substantial basis for making judgments.

The development of the projections for use in the MA scenario analysis also demonstrates that the conditional probabilistic approach requires making a number of subjective choices regarding the methodology. This is a strength of the approach in that it allows several degrees of freedom in matching criteria for the design of the population projections to the judgment of experts regarding the demographic consequences of the MA storylines. However it presents challenges in that methodological choices can influence the results in ways that may not be obvious *a priori*. The principal choices discussed in preceding sections are the index to use to categorize a projection as consistent, or inconsistent, with a socio-economic scenario; the averaging period (if this index involves a time average of demographic rates) and its relation to the storyline; the quantitative boundaries to separate index values into categories; and how to treat correlations across regions. The definition of high, medium, and low categories for each rate deserves special attention. Here, the conditional distributions of the input variables were designed by assuming that these three categories should be equally likely when averaged over the primary period of interest (2000-2050). As discussed above, this leads relatively small variance in conditional probabilistic projections that assume medium levels of demographic rates. Should the categories be designed instead so that they are equally likely over the entire period (2000-2100)? Or so that they have equal absolute variances, regardless of their relative likelihood? Should the variance or the relative likelihood of the output distributions be taken into consideration in designing the projections? Arguments could be made for each of these approaches. It should be kept in mind that the results for the MA projections are dependent on the methodological choices made.

Also, although questions regarding methodological details are important, the development of projections for the MA provides a basis for asking broader questions as well. In general, what is the best way to develop conditional probabilistic projections? Here, they have been constructed as sub-distributions of an overall unconditional projection. Perhaps it would be better to construct each conditional distribution independently, without regard for the unconditional distribution, by directly specifying probability density functions for the input variables based on expert opinion given particular storylines. Should each storyline be considered equally likely, at least in terms of the likelihood of their demographic consequences? Perhaps more extreme scenarios (e.g., high or low fertility) should be considered less likely than more moderate ones (e.g., medium fertility).

The development of the MA projections also points to useful areas of future research in demographic projections. Many of the issues raised in this analysis are a result of the assumptions regarding correlations of vital rates – over time, across regions, and between different rates – that underlie the IIASA population projections. These correlations affect the variance of the conditional distributions for both inputs and outputs, and the choice of averaging period and aggregation level for comparing demographic trends across simulations. The fact that the IIASA projections assume a single, deterministic correlation across regions introduces particular difficulties since the MA storylines differ markedly in their assumptions about cross-regional correlations. Strengthening the empirical basis for these assumptions should be a high priority. A second issue is scale. The storyline/quantification methodology used in the scenario analysis raises the question of how to match storylines describing broad development trends for large world regions over long time periods to projections made on smaller geographic units with rates that vary over shorter timescales. For example, given a particular scenario for socio-economic trends at the scale of a continent, what should be the demographic uncertainty for the continent? How should this differ from the uncertainty for sub-regions, or for individual countries?

Exploring these questions will provide a basis for the development of improved approaches to integrated scenario analysis that can explicitly account for the uncertainty inherent in future demographic developments. They may also lead to new ideas on how to address joint uncertainty in demographic as well as social, economic, and technological developments.

References

Alcamo, J., et al. Scenario 1 through 4. Chapter 7 of *Scenarios Working Group Assessment Report*, Millennium Ecosystem Assessment. In review.

Bongaarts, J. and Bulatao, R.A. (eds.) *Beyond Six Billion: Forecasting the World's Population* (Washington, DC: National Academy Press, 2000).

Lutz, W., Sanderson, W., Scherbov, S., and Goujon, A. 1996. World population scenarios for the 21st century. In Lutz, W. (ed) *The Future Population of the World: What can we assume today*. Earthscan, London, UK.

Lutz, W., Sanderson, W. and Scherbov, S. 2001. The end of world population growth. Nature 412, 54545.

Nakicenovic, N. et al. 2001. *Special Report on Emissions Scenarios*. Intergovernmental Panel on Climate Change. (Cambridge University Press, 2001).

O'Neill, B.C. Conditional probabilistic population projections: an application to climate change (2003) IIASA Interim Report IR-03-051. In press at *International Statistical Review*.

Sanderson, W., Scherbov, S., O'Neill, B.C., and W. Lutz. (2003) Conditional probabilistic population forecasting, IIASA Interim Report IR-03-052. In press at *International Statistical Review*.

Schwartz, P. 1991. *The Art of the Long View*. Doubleday, New York. 258 pp.

United Nations, 2003. World Population Prospects: The 2002 Revision Highlights. ESA/P/WP.180, United Nations, New York. 22 pp.

Table 1: Qualitative fertility, mortality, and migration assumptions for the Millennium Assessment scenarios.

Variable	Global Orchestration	Technogarden	Order From Strength	Adapting Mosaic
Fertility	HF: low LF: low VLF: medium	HF: medium LF: medium VLF: medium	HF: high LF: high VLF: low	Fortress until 2010, deviate to medium by 2050
Mortality	D: low I: low	D: medium I: medium	D: high I: high	Fortress until 2010, deviate to medium by 2050
Migration	High	medium	low	low

Notes:

[1] I = Industrialized country regions; D = Developing country regions; HF = High Fertility regions (TFR>2.1 in year 2000); LF = Low Fertility regions (1.7<TFR<2.1); VLF = Very Low Fertility regions (TFR<1.7, includes W. Europe, E. Europe, Sov. Europe, Pac. OECD).

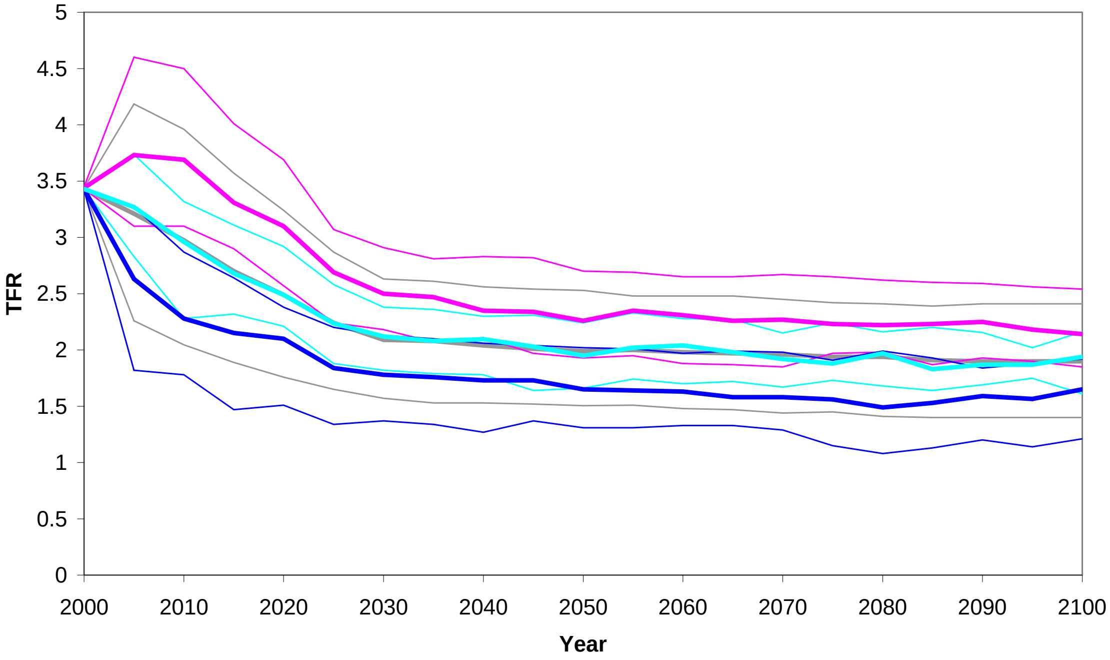
[2] In the IIASA projections, migration is assumed to be zero beyond 2070, so all scenarios have zero migration in the long run.

Table 2: Numbers of simulations (out of 10,000 total) meeting criteria for components of population change, for three MA scenarios and two selected regions.

Variable	Global Orchestration		Technogarden		Order From Strength	
	N. Afr.	W. Eur.	N. Afr.	W. Eur.	N. Afr.	W. Eur.
Fertility	794	367	310	367	759	743
Mortality	756	735	446	370	799	755
Migration	1026	1026	650	650	971	971
Fert.+Mort.	57	19	16	20	60	61
Mort.+Migr.	76	74	28	21	81	87
Fert.+Mort.+Migr.	6	1	0	3	5	4

Figure 1: Percentiles of unconditional and conditional distributions for TFR in the (a) North Africa region and (b) Western Europe region. Thick solid lines show medians, thin solid lines show 10th and 90th percentiles. Unconditional distribution in gray, conditional distribution for OS scenario in pink, TG scenario in light blue, and GO scenario in dark blue (for Western Europe, TG and GO scenarios have same fertility assumptions).

(a)



(b)

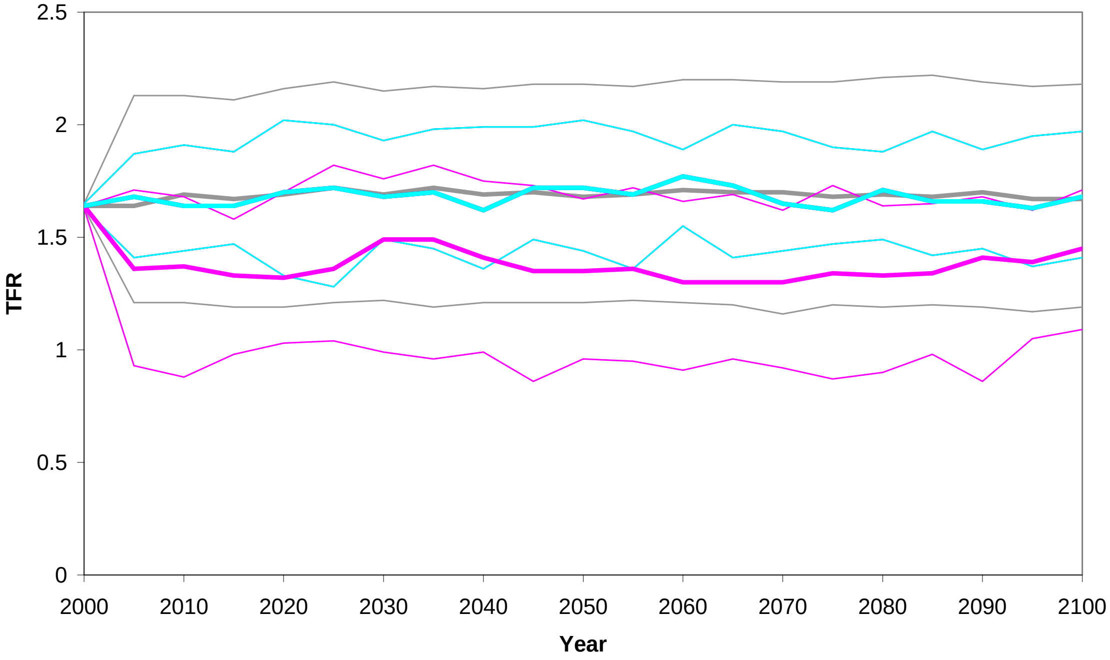
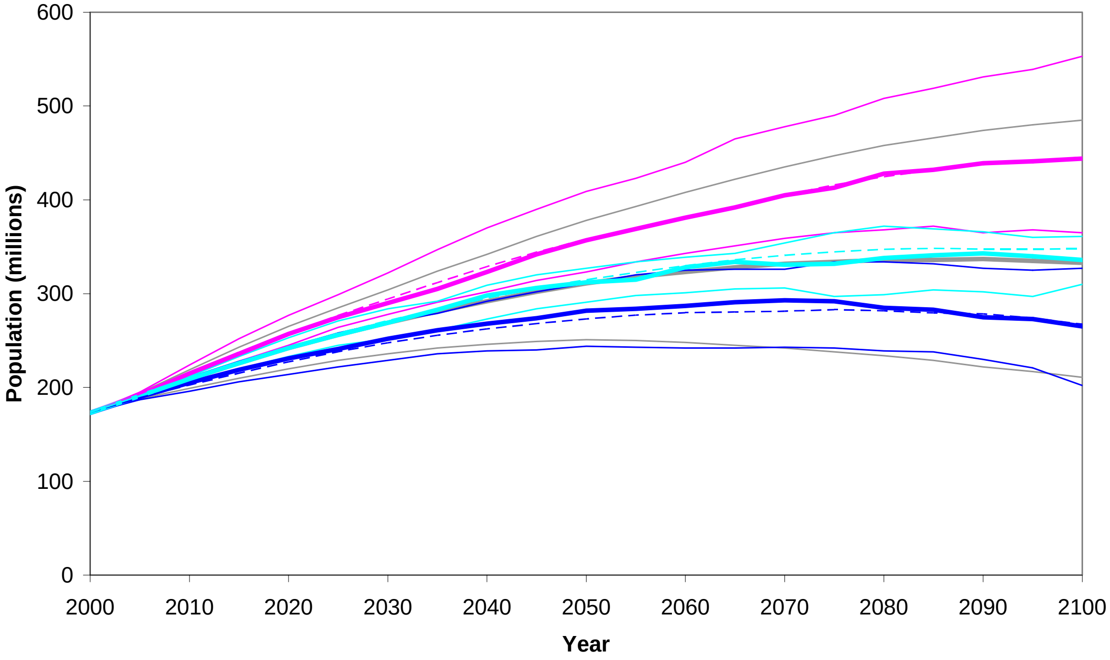


Figure 2: Percentiles of unconditional and conditional distributions of population size in the (a) North Africa region and (b) Western Europe region. Thick solid lines show medians, thin solid lines show 10th and 90th percentiles. Unconditional distribution in gray, conditional distribution for OS scenario in pink, TG scenario in light blue, and GO scenario in dark blue. Long dashed lines show deterministic projections used in MA scenarios.

(a)



(b)

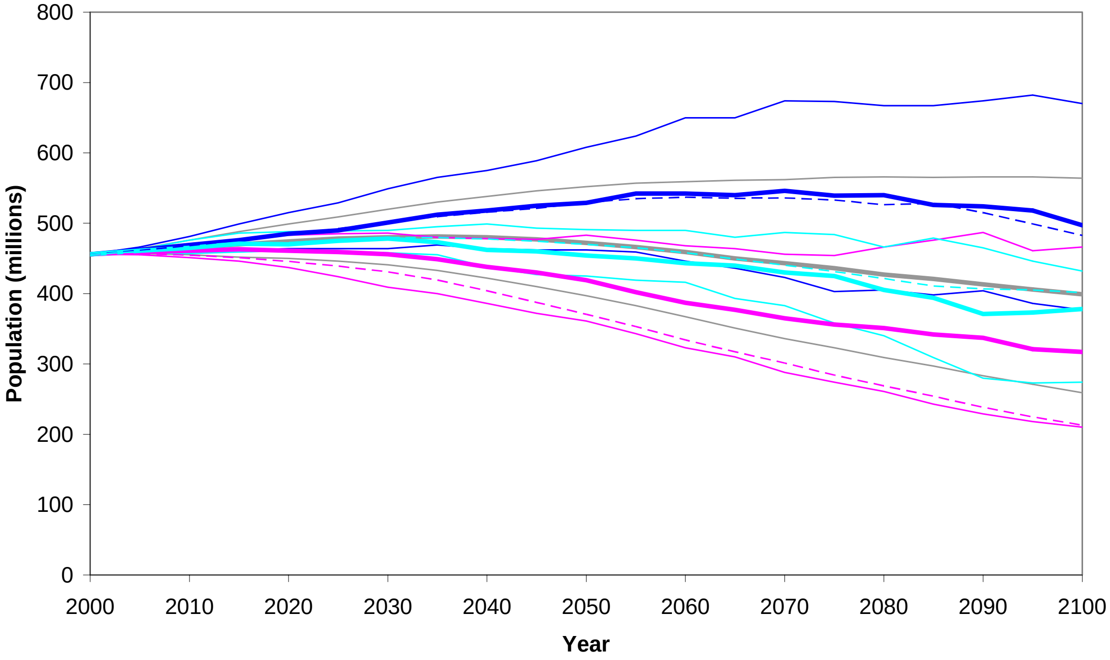


Figure 3: MA world population projections. For reference, the 10th and 90th percentile of the IIASA 2001 projections (Lutz et al., 2001) are shown, as well as the UN 2002 projections (UN, 2003) for 2050.