Research Article

Stefan Heinz

Uncertainty quantification of world population growth: A self-similar PDF model

Abstract: The uncertainty of world population growth represents a serious global problem. Existing methods for quantifying this uncertainty face a variety of questions. An essential problem of these methods is the lack of direct evidence for their validity, for example by means of comparisons with independent observations like measurements. A way to support the validity of such forecast methods is to validate these models with reference models, which play the role of independent observations. Desired properties of such a reference model are formulated here. A new reference world population model is formulated by a probabilistic extension of recent deterministic UN projections. This model is validated in terms of theory and observations: it is shown that the model has all desired properties of a reference model, and its predictions are very well supported by the known world population development from 1980 till 2010. Applications of this model as a reference model demonstrate the advantages of the stochastic world population model presented here.

Keywords: World population growth, UN world population forecasts, uncertainty quantification, stochastic world population model

MSC 2010: 60H25, 60H30

Stefan Heinz: Department of Mathematics, University of Wyoming, 1000 East University Avenue, Laramie, WY 82071, USA, e-mail: heinz@uwyo.edu

1 Introduction

Continuing world population growth will increase the global demand for food for the foreseeable future. There is a growing competition for land, water, and energy, which will be augmented by climate change effects due to global warming [15]. Hence, the growth of world population represents a fundamental problem: see the recent review of Lee [21] and the discussions of a variety of related questions in the July 2011 issue of Science. Population growth forecasts are also used to address many other relevant questions, including predicting the demand for education and medical services, as well as the future impact on labor markets, pension systems, and the environment [7].

There are three typical deterministic approaches to address world population forecasts [23]. The first approach is the consideration of one 'best-guess' forecast, e.g., the medium variant of the United Nations (UN) world population forecasts [37]. This approach provides valuable information, but it does not provide any insight into the uncertainty of this complex development, which is seen to represent a fundamental problem of populations forecasts [21]. A second approach is the consideration of alternative scenarios, which means typical deterministic realizations of the evolution of the global economic system involving population growth as one of the ingredients [29–32]. However, the validity of underlying assumptions of such projections is the concern of ongoing debates [8, 13, 18, 35], and this approach does only illustrate potential developments without explaining their likelihood. A third approach is the consideration of high and low variants of population growth in addition to a medium variant [37]. Such high and low variants indicate the range of variations of projections, but they do not explain the uncertainty of forecasts (the UN state that they do not interpret their high and low projection variants in a probability sense).

There are three typical probabilistic approaches used to quantify the uncertainty of population forecasts [14, 20]: the (classical and Bayesian) time series approach [5, 6, 9, 12, 19, 22], the extrapolation of empirical errors approach [1–4], and the random scenario approach [7, 24–28, 33]. A recent review of these approaches was given by Billari, Graziani and Melilli [10]. Previously, only the World Population Programme of the International Institute for Applied Systems Analysis (IIASA) published such forecasts at the global level: Lutz, Sanderson and Scherbov presented such expert-based probabilistic population projections based on a stochastic difference equation for the world population which is driven by normally distributed fertility, mortality, and migration rates [24–28, 33]. The relevant problem of this approach is the empirical expert-based determination of means, variances, and correlations of the fertility, mortality, and migration rates as functions of time [10, 23]. Such projections are helpful to illustrate the probability of potential developments and advantages of the probabilistic modeling approach, but the accuracy of such predictions is undetermined because of the uncertain knowledge of many adjustable model parameters. A promising new probabilistic method was recently presented by Alkema et al. [6, 7] for the modeling of the fertility rate of single countries. The difference to previously applied probabilistic modeling approaches is that this approach is based on modeling assumptions which reflect observed structural changes of the total fertility rate. Currently, influence factors like child mortality and economic development, which are difficult to predict, are not accounted for [6]. By including assumptions on mortality and migration rates, this approach can be extended for obtaining probabilistic projections of the population development of single countries. However, this approach faces two relevant problems. First, the purpose of developing such an approach should be a projection with reduced uncertainty, i.e., a projection that applies knowledge about structural developments for the reduction of the uncertainty predicted by simple approaches. However, this goal could not be consistently achieved: for most of the African countries the probabilistic UN projections [6, 7] have a very high uncertainty, which is, e.g., far outside of the high and low UN variants. Second, another significant problem is the use of such projections for probabilistic forecasts of world population growth [16]. There are several ways to combine country trajectories for obtaining one stochastic realization of the world population: the combination of randomly chosen country trajectories (which introduces errors), the consideration of all possible combinations of country trajectories (which is computationally extremely expensive), or the consideration of weighted or sorted trajectories (which can be done in different ways). Any of these approaches implies an overprediction of the world population uncertainty, i.e., a reduced uncertainty (see the first problem described above) is not obtained. The reason for the uncertainty overprediction is the neglect of correlations of the developments of countries in certain regions (Europe or Africa [27]). Such country correlations may imply that the population development uncertainties of several countries balance each other, i.e., the total population development uncertainty of these countries is smaller than the sum of all country uncertainties. There are several empirical ways of accounting for such correlations. Thus one has the choice between several results.

Among the approaches described above, only the probabilistic projection approach has the potential to fully determine the uncertainty of future population developments, i.e., the population probability density function (PDF). Nevertheless, existing probabilistic projection methods face relevant questions:

(i) existing models are based on country projections which results in unsolved questions about the consequences for the world population development (see the discussion in the preceding paragraph),
(ii) these models involve many adjustable model parameters that have to be guessed,
(iii) the models provide non-analytical results for the world population PDF, which hampers applications,
(iv) the relationship between different modeling approaches (like probabilistic projection methods and the deterministic high-low UN scenarios) is unclear,
(v) these models are not supported by theory applied to describe the evolution of complex systems.

These problems lead to the interesting theoretical question of how it is possible to develop a probabilistic projection method for which accurate proof (for example, via comparisons with independent observations like measurements) can never be obtained. The only possible approach to address this question is to validate models with other models based on different assumptions (which have to play the role of independent observations). What should be the properties of a model that can be used as a reference model for existing probabilistic projection approaches? Such a reference model should be characterized by the following properties:

(P1) It should be a global model for the world population that is not based on country projections.
(P2) It should be based on a minimum of assumptions involving usually applied model parameters.

(P3) It should be an analytical PDF model.
(P4) It should be a bridging model that relates the world population PDF to deterministic scenarios.
(P5) It should be supported by theory for the random evolution of complex systems.
The properties (P1)–(P5) address the problems (i)–(v) mentioned in the preceding paragraph. The advantage of property (P1) is that the problems related to the combination of country projections are avoided. The advantage of property (P2) is the reduction of uncertainty related to the involvement of many adjustable parameters. Property (P3) has a variety of advantages. Analytical PDF models are computationally very efficient and easy to use. A specific advantage is the clarity of conclusions, for example, regarding the bridging of modeling approaches, relevance and effect of model parameters, formulation of measures used for the model validation, and the variability of means and variances (see Sections 4–5). The advantage of property (P4) is the possibility to relate and compare conclusions obtained by deterministic and stochastic modeling approaches. Property (P5) may be seen to be the most relevant property. It provides significant support for the validity of the model formulation, which is, in particular, relevant to the case considered here where accurate proof for the model (for example, via comparisons with independent observations like measurements) can never be obtained. A detailed discussion of the relevance of property (P5) will be provided in Section 4.

A model that has the properties (P1)–(P5) will be presented in the following. The paper is organized in the following way. Section 2 prepares the following developments by an analysis of the 2010 UN world population forecasts released in May 2011 [37]. These UN projections will be used in Section 3 for the development of a new stochastic world population model. A theoretical validation of the model obtained will be performed in Section 4: it will be shown that the population model obtained has the reference model properties (P1)–(P5). The model validation by observations is the concern of Section 5. Section 6 illustrates the benefits of the new stochastic world population model. A summary of the results obtained here will be provided in Section 7.

2 The 2010 UN world population projections

To prepare the development of a stochastic model for the world population PDF, let us describe the 2010 UN world population forecasts released in May 2011 (see [37]) by means of analytical functions. Regarding the results reported in this Section 2 it is worth noting the following. First, the results reported in this section differ from the UN data only by the fact that the results reported here provide accurate analytical curve fits for the UN median, high, and low variants for $1950 \le t \le 2100$. Second, the latter analytical curve fits were not published by the UN.

To derive analytical curve fits for the UN variants we consider the modified logistic model [18] for the world population P (measured in 10^9) in time t,

$$\frac{dP}{dt} = \frac{1}{\tau}(P-L)\Big(1 - \frac{P-L}{K-L}\Big). \tag{2.1}$$

Here, L refers to an asymptotic population level found at large negative t, K refers to the carrying capacity (the maximum population), and τ is a characteristic time scale. The solution of the model (2.1) reads

$$P = L + (K-L)\Big[1 - \Big(1 - \frac{K-L}{P_0 - L}\Big)\exp\Big\{-\frac{t-t_0}{\tau}\Big\}\Big]^{-1}, \tag{2.2}$$

where t_0 and $P_0 \ge L$ represent initial values of t and P, respectively.

The application of the model (2.2) to the known population data ($1950 \le t \le 2010$) and the median variant ($2010 \le t \le 2100$) of the 2010 UN projections reveals the suitability of this model. As shown in Figure 1 (a), the known evolution of the world population ($1950 \le t \le 2010$) and the median variant ($2010 \le t \le 2100$) can be described very well by equation (2.2) combined with $t_0 = 1950$, $P_0 = 2.53$ and the parameters values $K = 10.35$, $L = 1.06$, and $\tau = 27$. These parameter values were obtained here by minimizing the relative model error, which is found to be smaller than 0.8%.

The model (2.2) cannot only describe the median variant, but it can also describe the high and low UN variants ($2010 \le t \le 2100$) if K is provided as an appropriate function of time t. Knowledge about which

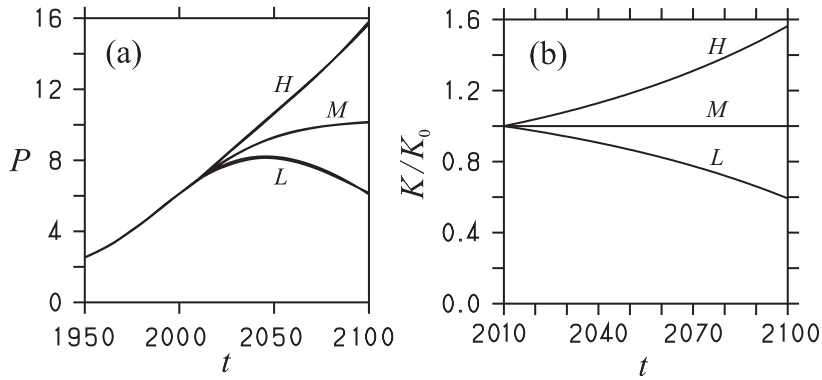

Figure 1. The solid lines in (a) show the high (H), median (M), and low (L) 2010 UN population projection data (measured in 10^9). The corresponding dashed lines represent the model (2.2) combined with the K model (2.4), which is illustrated in (b).

functions $K(t)$ are needed to cover the high and low UN variants can be obtained by using equation (2.2) for writing K as a function of P,

$$K = P_0 + (P - P_0)\left[1 - \frac{P - L}{P_0 - L}\exp\left\{-\frac{t - t_0}{\tau}\right\}\right]^{-1}. \tag{2.3}$$

The use of the known UN data $P(t)$ in equation (2.3) then enables the calculation of $K(t)$. The result of this calculation is given by

$$K = K_0[1 + (K_1 c + K_2 c^2)g(t)]. \tag{2.4}$$

Here, $K_0 = 10.35$ is the value involved in the model for the median UN projection, we used the abbreviations $K_1 = 0.565$ and $K_2 = 0.182$, and the function

$$g(t) = \exp\left\{\frac{t - 2010}{90}\right\} - 1 \tag{2.5}$$

is introduced here to describe the K variations in time. The model parameter c in equation (2.4) reflects the assumptions of UN variants: $c = 0$ corresponds to the UN median variant, $c = 0.5$ corresponds to the UN high variant (the total fertility rate is half a child higher), and $c = -0.5$ corresponds to the UN low variant (the total fertility rate is half a child lower). The use of equation (2.4) in equation (2.2) combined with $c = (0.5, 0, -0.5)$ covers the high, median, and low UN projections for $t > 2010$, respectively. The variation of K according to equation (2.4) is shown in Figure 1 (b). The corresponding curves for the high and low UN projections are shown in Figure 1 (a). It may be seen that these models perform very well. The relative error of the high and low variant models is smaller than 1% and 1.5%, respectively.

3 The stochastic world population model

Equation (2.2) combined with equation (2.4) provides for $c = (0.5, 0, -0.5)$ accurate analytical models for the high, median, and low UN projections for $t > 2010$, respectively. The population trajectories for $c = (0.5, 0, -0.5)$ represent realizations for the case that the total fertility rate is half a child higher than assumed in the median variant, equal to the median variant assumption, and half a child lower than assumed in the median variant, respectively. These available realizations can be extended to a stochastic model by assuming that the model parameter c is a random variable. In particular, we will assume that c is normally distributed with zero mean and standard deviation σ. The consideration of a normal random variable c is the most reasonable assumption [26]. The assumption of a zero mean of c reflects the relevance of c to account for deviations of the total fertility rate from the median variant assumption. Insight into appropriate values for the standard deviation σ of c can be obtained by means of Figure 2, which shows the PDF of c for three values of σ. In the

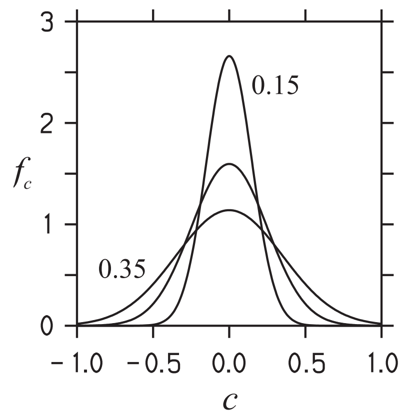

Figure 2. The solid lines shows the PDF f_c of c for $\sigma = 0.25$. The dashed lines show f_c for $\sigma = 0.15$ and $\sigma = 0.35$ (the value of σ is given at the curve).

UN scenarios, values $c = \pm 0.5$ are considered to be high deviations from zero. It may be seen that the PDF provides for $\sigma = 0.15$ random variations of c which are, basically, bounded by $-0.5 < c < 0.5$. On the other hand, the PDF for $\sigma = 0.35$ provides random variations of c which are, basically, bounded by $-1 < c < 1$. For this PDF, values $c = \pm 0.5$ can hardly be considered to represent high deviations from zero anymore, which means $\sigma = 0.35$ can be seen as an upper limit for the standard deviation of the c PDF. Therefore, a reasonable range of σ variations is given by $\sigma = 0.25 \pm 0.1$. It is worth noting that the mean value $\sigma = 0.25$ assumed here agrees well with the assumption $\sigma = 0.234$ used by Lutz, Sanderson and Scherbov [26]. A consequence of assuming the model parameter c to be normally distributed is that the mean $\langle K \rangle = K_0(1 + \Delta)$, i.e., $\langle K \rangle > K_0$. Here and below, the bracket symbol refers to an ensemble average, and $\Delta = K_2 \sigma^2 g(t) > 0$. For $\sigma = 0.25$ we find $\Delta = (0.21, 0.64, 1.20, 1.95)\%$ for $t = (2025, 2050, 2075, 2100)$, respectively. Thus, the deviation of $\langle K \rangle$ from K_0, which characterizes the median variant, is negligibly small with a maximum of 2% at $t = 2100$. The significant advantage of accepting this very minor deviation of $\langle K \rangle$ from K_0 is the simplicity of the model formulation.

The high and low UN variants represent the sum of the high or low variants of all countries without considering any country correlations. Such correlations may imply that the population development uncertainties of several countries balance each other so that the total population development uncertainty of these countries is smaller than the sum of all country uncertainties. In terms of the model obtained above for the high and low UN variants, the consequence of accounting for country correlations would be a reduction of $g(t)$. A reasonable assumption would be the replacement of $g(t)$ by $rg(t)$. Here, the parameter r, which may be expected to be bounded by $0.75 \le r \le 1$, accounts for the reduction of $g(t)$ due to country correlations. Such a factor r will be not considered here because of three reasons. First, such correlations seem to be relatively small: a mean correlation of 0.15 is quoted in reference [11]. Second, the analysis of such correlations represents a complex issue that requires empirical estimates [2]. Third, the influence of r values can be covered by modifications of the standard deviation σ. The latter can be seen by considering the standard deviation $\sigma_K = K_1 \sigma g K_0 [1 + 2(K_2/K_1)^2 \sigma^2]^{1/2}$ of K, which follows from equation (2.4) in conjunction with the consideration of a normally distributed c. For $\sigma = 0.25$ we find $2(K_2/K_1)^2 \sigma^2 = 0.013$. Hence, $\sigma_K = K_1 \sigma K_0 g$ represents an accurate approximation for σ_K. Then, the effect of r can be absorbed in σ. For example, the use of $r = 0.75$ corresponds to $\sigma = 0.19$, which is inside the range of σ variations considered above, $\sigma = 0.25 \pm 0.1$.

The use of equation (2.4) in equation (2.2) results in the following stochastic population model,

$$P = m_3 - \frac{m_1}{\epsilon + m_2}. \tag{3.1}$$

Here, the abbreviations m_1, m_2, and m_3 are given by the expressions

$$m_1 = h(t)\frac{P_0 + h(t) - L}{K_0 g(t)}, \quad m_2 = \frac{h(t) + K_0 - L}{K_0 g(t)}, \quad m_3 = h(t) + P_0, \tag{3.2}$$

and we introduced the function

$$h(t) = (P_0 - L)\left[\exp\left\{\frac{t - t_0}{\tau}\right\} - 1\right]. \tag{3.3}$$

The functions m_1, m_2, and m_3 are positive. In particular, m_1 and m_2 have minimum values at $t = 2024.1$ and $t = 2047.8$, whereas m_3 is always increasing with t. The abbreviation ϵ in equation (3.1) refers to the expression

$$\epsilon = K_1 c + K_2 c^2. \tag{3.4}$$

Here, c is normally distributed with zero mean and variance σ^2, this means the PDF f_c of c is given by the expression

$$f_c = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{c^2}{2\sigma^2}\right\}. \tag{3.5}$$

The range of σ variations is given by $\sigma = 0.25 \pm 0.1$, see the discussion in the first paragraph of Section 3.

The cumulative distribution function (CDF) $F(p)$ and the corresponding world population PDF

$$f(p) = \frac{dF}{dp}$$

implied by the stochastic population model equation (3.1) can be calculated in the following way. The CDF $F(p)$ is defined by the probability Pr to find population values P smaller or equal to p,

$$F(p) = \Pr(P \le p) = \Pr\left(m_3 - \frac{m_1}{\epsilon + m_2} \le p\right) = \Pr(c \le H(p)), \tag{3.6}$$

where we introduced the abbreviation

$$H(p) = \frac{K_1}{2K_2}\left\{\sqrt{1 - 4\frac{K_2}{K_1^2}\left(m_2 + \frac{m_1}{p - m_3}\right)} - 1\right\}. \tag{3.7}$$

Equation (3.1) was used for P in equation (3.6), and the last inequality in equation (3.6) represents the rewritten previous inequality. Due to the properties of c we find, therefore,

$$F(p) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{H(p)} \exp\left\{-\frac{x^2}{2\sigma^2}\right\} dx. \tag{3.8}$$

The application of equation (3.8) can be simplified by rewriting $F(p)$ in terms of the error function $\mathrm{erf}(x)$,

$$F(p) = \frac{1}{2}\left[1 + \mathrm{erf}\left\{\frac{H(p)}{\sqrt{2}\sigma}\right\}\right]. \tag{3.9}$$

Here,

$$\mathrm{erf}(x) = \pm\left[1 - \exp\left\{\frac{-\left(\frac{4x^2}{\pi} + 0.14x^4\right)}{1 + 0.14x^2}\right\}\right]^{1/2}$$

can be used to analytically approximate $\mathrm{erf}(x)$ (see [18]). The positive (negative) sign applies to positive (negative) x values. The relative error of the latter approximation for the error function $\mathrm{erf}(x)$ is smaller than 0.04% (see [18]). The implied world population PDF $f(p) = dF/dp$ is found to be given by

$$f(p) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{H^2(p)}{2\sigma^2}\right\} \frac{dH}{dp} = \frac{1}{\sqrt{2\pi\sigma^2}} \frac{m_1}{(p - m_3)^2} \frac{\exp\left\{-\frac{H^2(p)}{2\sigma^2}\right\}}{K_1 + 2K_2 H(p)}. \tag{3.10}$$

The last expression specifies the derivative dH/dp.

The validity of the analytical PDF model (3.10) can be confirmed by Monte Carlo simulation. This simulation was performed by using equation (3.1) to generate 10^7 sample values of P at a time considered, which were used to calculate the PDF. The latter PDF will be referred to as Monte Carlo PDF. The numerical method to calculate such Monte Carlo PDFs is described elsewhere [18]. A PDF filter interval $\sigma_P/10$ was used, where σ_P refers to the PDF standard deviation. The PDFs were calculated at positions separated by $\sigma_P/20$. The Monte Carlo PDF calculation required less than two minutes on a Pentium(R) 4 CPU 3.2 GHz personal computer with 1 GB memory. The comparison between such Monte Carlo PDFs and analytical PDFs obtained by equation (3.10) is shown in Figure 3 for $t = (2025, 2050, 2075, 2100)$. It may be seen that there is no observable difference between the Monte Carlo and analytical PDFs. All the PDFs shown in Figure 3 are characterized by a significant skewness.

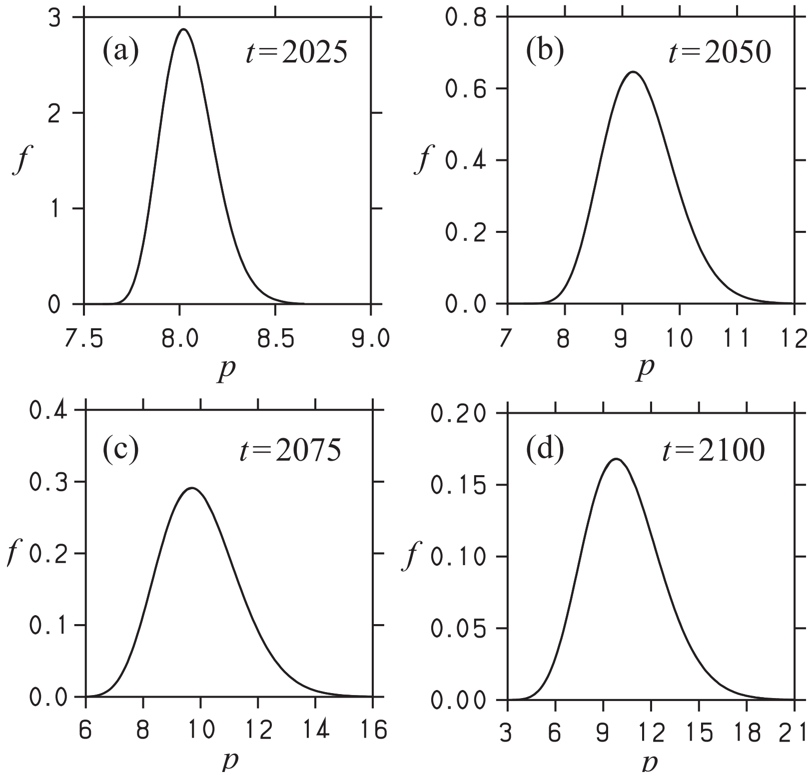

Figure 3. The solid lines show the world population PDF $f(p)$ obtained from the stochastic model (3.1) by Monte Carlo simulation at the given time $t = 2025, 2050, 2075$, and 2100, respectively. The dashed lines show the corresponding analytical PDF (3.10).

4 Theoretical PDF model validation

Next, let us consider whether the stochastic population model (3.1) satisfies the desired properties (P1)–(P5) of a reference world population model described in the introduction. The way of developing the model (3.1) reveals that this model satisfies property (P1) to represent a global model. The model does also satisfy property (P2): the only parameter that enters the model is the standard deviation σ, which determines the total fertility rate distribution. As shown in Section 3, the model does also satisfy property (P3) to represent an analytical PDF model.

Let us consider whether the stochastic population model (3.1) satisfies property (P4), this means whether this model represents a bridging model that relates the world population PDF to deterministic scenarios. This question will be addressed by calculating the probability covered by the UN high and low variants. This probability $p_{LH}(t)$ is given by

$$p_{LH}(t) = \Pr(P_L(t) \le P \le P_H(t)).$$

Here, $P_L(t)$ and $P_H(t)$ refer to the population values predicted by the UN low and high variants, respectively. The functions $P_L(t)$ and $P_H(t)$ are determined by equation (2.2) combined with $K = K_0(1 + [-K_1/2 + K_2/4]g(t))$ and $K = K_0(1+[K_1/2+K_2/4]g(t))$, respectively, which follow from equation (2.4) by setting $c = -0.5$ and $c = 0.5$. According to equation (3.1), $P_L(t)$ and $P_H(t)$ are given by

$$P_L(t) = m_3 - \frac{m_1}{-\frac{K_1}{2} + \frac{K_2}{4} + m_2}, \quad P_H(t) = m_3 - \frac{m_1}{\frac{K_1}{2} + \frac{K_2}{4} + m_2}. \tag{4.1}$$

In terms of $\Pr(P \le p) = F(p)$, the probability $p_{LH}(t) = \Pr(P_L(t) \le P \le P_H(t))$ can be written

$$p_{LH}(t) = F(P_H(t)) - F(P_L(t)). \tag{4.2}$$

By using relation (3.9) we obtain

$$p_{\mathrm{LH}}(t) = \frac{1}{2}\left[\operatorname{erf}\left(\frac{H[P_{\mathrm{H}}(t)]}{\sqrt{2}\sigma}\right) - \operatorname{erf}\left(\frac{H[P_{\mathrm{L}}(t)]}{\sqrt{2}\sigma}\right)\right]. \tag{4.3}$$

According to equation (3.7) for H we find $H[P_{\mathrm{H}}(t)] = 1/2$ and $H[P_{\mathrm{L}}(t)] = -1/2$. Hence, equation (4.3) reduces to

$$p_{\mathrm{LH}}(t) = \frac{1}{2}\left[\operatorname{erf}\left(\frac{1}{\sqrt{2}2\sigma}\right) - \operatorname{erf}\left(-\frac{1}{\sqrt{2}2\sigma}\right)\right] = \operatorname{erf}\left(\frac{1}{\sqrt{2}2\sigma}\right). \tag{4.4}$$

We observe that the probability $p_{\mathrm{LH}}(t)$ is independent of time and fully determined by σ. Relation (4.4) represents a relevant result because it explains how the standard deviation σ of the PDF of c, which specifies the variability of deviations from the median fertility assumption, is related to the probability covered by the UN high and low variants. By applying the values $\sigma = (0.15, 0.25, 0.35)$ considered above we find the reasonable values $p_{\mathrm{LH}} = (1.00, 0.95, 0.85)$, respectively.

The remaining question is whether the PDF model also has the property (P5), this means whether the model is supported by constraints for the random evolution of complex systems. A convenient basis for the discussion of constraints for the random evolution of complex systems is the consideration of standardized random variables, which is the standardized population density $P_* = (P - \langle P\rangle)/\sigma_P$ for the case considered here ($\langle P\rangle$ refers to the mean population density). There are two possibilities in which P_* can evolve in time. The first possibility is that P_* does not change in time, i.e., $P_* = P_*(0)$, where $P_*(0)$ refers to P_* at the initial time. Such a process is characterized by a self-similar PDF, which has the same shape over all time. The second possibility is that P_* changes in time, this means $P_* = P_*(t)$. This case is given if new randomness enters the model over all time. The PDF of the standardized variable will then change in time, i.e., the PDF is nonself-similar.

Should the population PDF be a self-similar PDF? The complete specification of the population PDF requires two ingredients: the temporal evolution of the mean and standard deviation, and the temporal evolution of all standardized higher order statistics (the skewness, flatness, etc.), which determine the evolution of the PDF of the standardized variable P_*. The latter requirement can be split into two problems:

(a) the need to provide characteristic estimates for the skewness, flatness, etc. at the time of forecast considered ($t = 2010$),

(b) the need to provide the temporal evolution of deviations of these characteristic estimates.

Reasonable solutions for problem (a) can be obtained, but there is no reliable information available to solve problem (b). Therefore, the skewness, flatness, etc. should hardly vary in time to enable reliable predictions. Consequently, the population PDF should be well approximated by a self-similar PDF.

In the Appendix it is shown that the nonself-similar standardized population PDF (A.3), which is equal to the standardized PDF (3.10), can be approximated by the self-similar PDF (A.6). According to equation (A.7), the relative error of this approximation is found to be very small. Another way to assess the accuracy of this approximation is to compare the self-similar PDF (A.6) with the nonself-similar standardized population PDF (A.3). The use of the latter PDF requires the calculation of the mean $\langle P\rangle$ and standard deviation σ_P of the population PDF (3.10). These values were calculated here on the basis of equation (3.1) by Monte Carlo simulations using 10^7 sample values. The difference between the self-similar PDF (A.6) and the nonself-similar standardized PDF (A.3) is maximal at $t = 2047.8$ (because W has a maximum and m_2 has a minimum at $t = 2047.8$: see the Appendix). This t value was considered for the comparison in Figure 4 (a) to illustrate for $\sigma = 0.25$ the maximum difference between the self-similar PDF (A.6) and the nonself-similar standardized PDF (A.3). It may be seen that there is hardly any difference between the two PDFs considered. Thus, the nonself-similar standardized population PDF (A.3) is extremely well represented by the self-similar PDF (A.6), i.e., property (P5) is also satisfied. The self-similar PDF (A.6) is only affected by variations of σ. The effect of σ is illustrated in Figure 4 (b) for the range of σ variations considered. We see that the effect of these σ variations is rather limited, i.e., the self-similar PDF (A.6) is little affected by the σ variations considered.

What is the advantage of proving that the population PDF can be approximated by a self-similar PDF? First, a theoretical advantage is given by the demonstration that the population PDF is supported by theory, this means the PDF is in consistency with the theoretical constraint that the PDF should be approximately

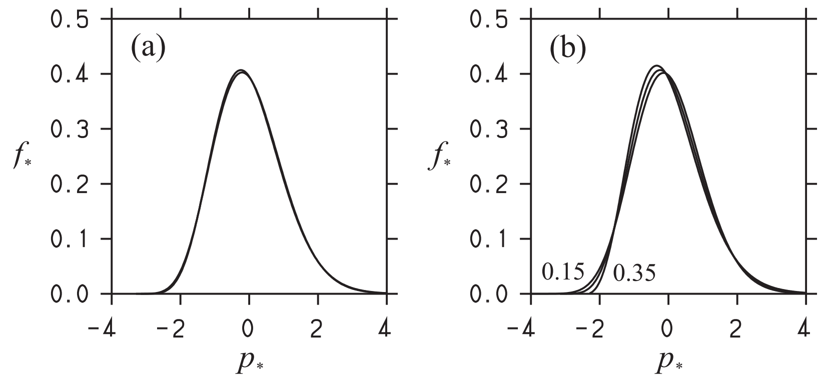

Figure 4. The standardized population PDF. The solid line in (a) shows the self-similar PDF (A.6) at $t = 2047.8$ for $\sigma = 0.25$. The corresponding nonself-similar standardized PDF (A.3) is given by a dashed line. The effect of variations on the self-similar PDF (A.6) is shown in (b). The solid line shows f_* for $\sigma = 0.25$. The dashed lines show f_* for $\sigma = (0.15, 0.35)$: the value of σ is given at the curve.

self-similar. This requirement was discussed in the fourth paragraph of this section: it was argued that the standardized PDF should be hardly affected by assumptions about the temporal evolution of standardized model statistics (i.e., the skewness, flatness, etc.) to increase the reliability of forecasts and to decrease the sensitivity to parameter variations. Second, a practical advantage is the identification of accurate and simple expressions for the mean and variance of the population PDF (in contrast to the mean and variance of the nonself-similar PDF (A.3) which can only be calculated numerically): see expressions (A.9) in the Appendix. The latter advantage will be explained in more detail in Section 6.

5 PDF model validation by observations

How it is possible to validate the suitability of the world population PDF model (3.10)? A closer look shows that there are two problems related to

(i) the concrete model compared to available observations,
(ii) the method applied to compare a probabilistic prediction with deterministic available data.

The problem (i) is that there is no way to compare the PDF model (3.10), which provides a probabilistic forecast for $2010 \leq t \leq 2100$, directly with observations (also, the latter would have little value because there is no need for a forecast if the population development is known). Observations are available, e.g., for $1980 \leq t \leq 2010$ (see [36]). There are two ways to compare these observations for $1980 \leq t \leq 2010$ with forecast methods. One forecast method, which will be referred to below as old forecast, is to consider the median, high, and low UN forecasts made in 1980 and to redo the analysis performed here on the basis of this forecast. A second forecast method, which will be referred to below as new forecast, is to apply the model derived here at the time of forecast considered, this means by using $t = 1980$ instead of $t = 2010$ in $g(t)$: see equation (2.5). For the case that only the old (or new) forecast provides a good agreement with observations, it would be questionable whether this result provides indeed support for the validity of the forecast for $2010 \leq t \leq 2100$ because the PDF model structure will differ from the PDF model (3.10) (or the applicability of the PDF model structure (3.10) at $t = 1980$ is questionable). Therefore, a good agreement between the old and new forecasts is a requirement to support the validity of the PDF model (3.10) in this way. In particular, it needs a good agreement between the old and new median, high, and low forecasts because this information is used to develop the old and new forecasts.

To address the question described in the preceding paragraph, let us compare the old and new forecasts. Figure 5 shows the old UN median, high and low variants of the world population development as predicted in 1980 [36]. Figure 5 also shows the new analytical models implied by equations (2.2), (2.3), and (2.4) for the median, high and low variants of the world population as predicted in 2010 (the difference to the

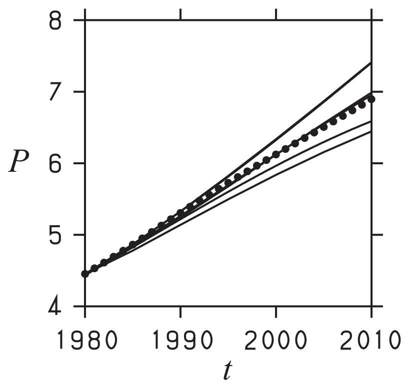

Figure 5. The solid lines show the high, median, and low 1980 UN population projection data [36] in comparison with the population development seen in reality (dots). The dashed lines show the corresponding new forecasts.

corresponding curves shown in Figure 1 is that the initial time of the high and low variants in equation (2.5) is set to be 1980 instead of 2010). It may be seen that there is an excellent agreement between these curves regarding the mean and high variants. The new low UN forecast deviates from the mean variant slightly less than the old UN forecast made in 1980, but there is also a relatively good agreement between the old and new forecasts. Hence, the old and new UN projections provide very similar results. Thus, the comparison of these forecasts with the known $1980 \le t \le 2010$ population data represents an appropriate approach to validate the suitability of the world population PDF model (3.10).

The problem (ii) regarding the validation of the world population PDF model (3.10) is the question of how the old and new forecasts will be compared to the available observations for $1980 \le t \le 2010$. A deterministic approach to assess the accuracy of these forecasts is to compare the relative errors of the mean old and new forecast with the available observations for $1980 \le t \le 2010$. However, this approach does not reflect the probabilistic nature of the new PDF model applied at $t = 1980$. One possibility of reflecting the probabilistic structure is to calculate the probability of population values in some neighborhood of the observed population development for $1980 \le t \le 2010$. This leads to questions about the definition of the neighborhood. To simplify the approach significantly, let us assume that the known world population development represents one trajectory predicted by the probabilistic forecast (which is well justified regarding the excellent agreement between the observed population development and the development predicted by the new forecast). Then, the question is how likely it is to find this trajectory considered. Obviously, the calculation of a very high probability of finding the trajectory observed in reality reflects a very good performance of the probabilistic forecast. We have to introduce a trajectory probability measure to quantify this idea. To do so, the probability $p_c(t)$ to find a trajectory defined by the setting of c is given by

$$p_c(t) = \Pr(P_{c-\Delta c}(t) \le P \le P_{c+\Delta c}(t)).$$

Here, $P_{c-\Delta c}(t)$ and $P_{c+\Delta c}(t)$ refer to P given by equation (3.1) where c is replaced by $c - \Delta c$ and $c + \Delta c$, respectively, and Δc is an infinitesimal interval. It turns out that $p_c = 2\Delta c f_c$ is twice the probability to find a value of c between c and $c + \Delta c$, where f_c refers to the normal PDF of c. Relation $p_c = 2\Delta c f_c$ shows that the probability p_c is independent of time t, this means p_c represents a global measure. The variable p_c provides the probability for a trajectory specified via the setting of c, but it does not represent a probability measure ranging between zero and one. Such a trajectory probability measure, which will be denoted by P_T, is given by normalizing p_c in terms of its maximum. According to $p_c = 2\Delta c f_c$, the maximum of p_c is given at $c = 0$ by $p_0 = 2\Delta c/(2\pi\sigma^2)^{1/2}$. Hence, the trajectory probability measure $P_T = p_c/p_0$, which has the property $0 \le P_T \le 1$ by definition, reads $P_T = \exp[-c^2/(2\sigma^2)]$.

The use of the validation concepts described in the preceding paragraph provides the following results. By following the deterministic validation approach, we assess the accuracy of old and new population forecast in terms of the relative errors of corresponding mean predictions defined by $e_{M1} = 100(P^{M1} - P^R)/P^R$ and $e_{M2} = 100(P^{M2} - P^R)/P^R$. Here, the observed development is denoted by P^R, and the subscripts and superscripts 1 and 2 refer to the old and new projections, respectively. It is found that the absolute value of the error of the old forecast $|e_{M1}| < 1.3\%$. Thus, we may conclude that the old UN projection of the world population represents an accurate population projection. On the other hand, the absolute value of the error of the new forecast $|e_{M2}| < 0.8\%$. Hence, the new UN forecast is even more accurate than the old UN forecast. The probabilistic validation approach described in the preceding paragraph requires the determination of c related to the observed population data to calculate the corresponding trajectory probability measure P_T. The new UN mean variant, which corresponds to the setting $c = 0$ in equation (3.1), was shown to represent an excellent approximation to the population data observed in reality (the absolute value of the relative error is smaller than 0.8%). For $c = 0$, we find for the trajectory probability measure $P_T = 1$. Thus, we can conclude that the new forecast has a very high accuracy because there is a maximum probability for finding the observed population data. These findings provide relevant support for the suitability of the world population PDF model (3.10).

6 PDF model benefits

Finally, the benefits of the PDF model (3.10) presented here will be illustrated in comparison to the deterministic median, high, and low UN projections, another deterministic forecast, and another probabilistic forecast.

Compared to the deterministic median, high, and low UN forecasts, the PDF model (3.10) offers several advantages. First, the PDF model does not only provide the median, high, and low UN variants: the PDF model does also provide an infinite number of other realizations, and it specifies the probability P_T for finding such realizations. Second, the PDF model (3.10) can be used to calculate the probability of any events, like the probability covered by the UN high and low variants, or the probabilities $\Pr(P \leq 4.44)$ and $\Pr(P \leq 8)$ calculated below in this section. Third, the PDF model (3.10) can be used to quantify the uncertainty range of the median UN projection. The latter advantage will be explained in more detail in the next paragraph. Fourth, the probabilistic extension of the deterministic UN projections presented here offers the possibility to use the PDF model as a reference model for comparisons with other deterministic and stochastic methods: the PDF model (3.10) can be used to calculate the probability for conclusions of deterministic forecast methods, and the PDF model can be used for proving the validity of calculations obtained by means of probabilistic methods. Two examples for the use of the PDF model (3.10) as a reference model will be provided in the last two paragraphs of this section.

The uncertainty range of the median UN projection can be quantified in the following way. According to relations (A.9) below, the world population mean and standard deviation are given by $\langle P \rangle = m_3 - m_1/m_2$ and $\sigma_P = \sigma G(t)$, where the abbreviation $G(t) = K_1 m_1/m_2^2$ is used. It turns out that the mean $\langle P \rangle$ agrees with equation (3.1) combined with equation (3.4) and $c = 0$. As it has to be expected, the standard deviation σ_P is proportional to σ because G is independent of σ. The development of G in time is shown in Figure 6 (a). This figure does also show that G can be characterized very well by a quadratic function of time, $G = s(2.2 + 7.5s)$. Here, s is defined by $s = (t - 2010)/90$, which varies between 0 and 1 for $2010 \leq t \leq 2100$ considered. The equation $G = s(2.2+7.5s)$ was found by modeling G/s as a linear function of s. The expressions $\langle P \rangle = m_3 - m_1/m_2$ and $\sigma_P = \sigma G(t)$ for the world population mean and standard deviation can be used to quantify the uncertainty of forecasts. Corresponding to the definition of the intensity of turbulence [17], we define a population forecast uncertainty measure in terms of the coefficient of variation $i = \sigma_P/\langle P \rangle$. According to $\langle P \rangle = m_3 - m_1/m_2$ and $\sigma_P = \sigma G(t)$, the population forecast uncertainty measure is given by $i = \sigma H(t)$, where $H(t)$, which is independent of σ, is given by

$$H(t) = \frac{K_1 m_1}{m_2(m_3 m_2 - m_1)}. \tag{6.1}$$

The development of H in time is shown in Figure 6 (b). This figure does also demonstrate that H can be characterized very well by the quadratic function, $H = s(2.2 + 4.75s)/7.3$. The writing $H = s(2.2 + 4.75s)/7.3$ takes reference to $G = s(2.2+7.5s)$ and shows that H is not given by dividing G by a constant mean population value. For population predictions over thirty, sixty, and ninety years, this means at $t = (2040, 2070, 2100)$,

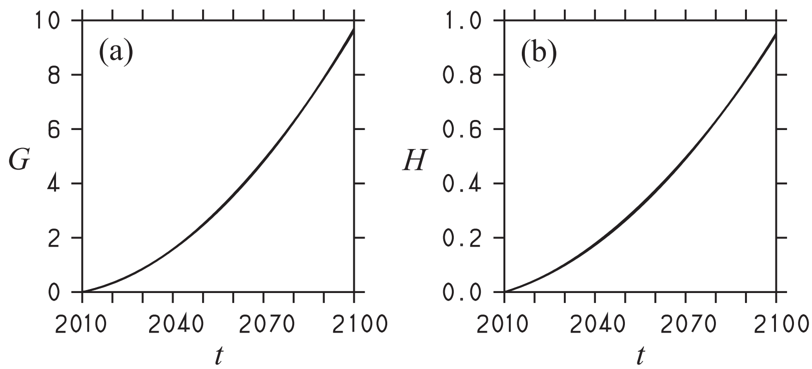

Figure 6. The solid lines in (a) and (b) show the functions $G = K_1 m_1/m_2^2$ and H given by equation (6.1). The dashed lines show the corresponding approximations $G = s(2.2 + 7.5s)$ and $H = s(2.2 + 4.75s)/7.3$ for G and H.

we find the population forecast uncertainty measure i to be given by $i = (4.3 \pm 1.7, 12.3 \pm 4.9, 23.8 \pm 9.5)\%$, respectively. The variation of i values in this relation, which arises from the range $\sigma = 0.25 \pm 0.1$ of σ values considered, characterizes the accuracy of the uncertainty measure prediction. Hence, we find an acceptable population forecast variability for the first thirty years. The i values for predictions over sixty and ninety years reflect the significant nonlinear increase of the forecast uncertainty. However, even the value $i = 23.8 \pm 9.5$ can be considered to be reasonable for a forecast over ninety years.

The value of the world population PDF model (3.10) as a reference model will be illustrated by a comparison with conclusions obtained by a deterministic forecast method: the World3 model. The latter model predicts in its standard scenario that the world population in 2100 will be equal to the world population in 1980 [18, 30, 32, 35]. According to equation (2.2), the population in 1980 was $P = 4.44$. The probability p_{W3} for finding the world population in 2100 to be equal or below the world population in 1980 is then given by

$$p_{w3} = \Pr(P \leq 4.44) = \frac{1}{2} + \frac{1}{2}\,\mathrm{erf}\Big(\frac{H(4.44)}{\sqrt{2}\sigma}\Big) = \frac{1}{2} + \frac{1}{2}\,\mathrm{erf}\Big(-\frac{0.5577}{\sigma}\Big), \tag{6.2}$$

where $\Pr(P \leq p) = F(p)$ and equation (3.9) were used. For $\sigma = 0.25$ we obtain $p_{W3} = 0.08\%$. This means that the conclusions of the World3 model are not very likely.

The value of the world population PDF model (3.10) as a reference model will be also illustrated by a comparison with conclusions obtained by another probabilistic forecast method [34]. In particular, Scherbov, Lutz and Sanderson [34] calculated the temporal development of the probability for a world population of more than 8 billion by using their probabilistic population projection, which is based on the empirical expert-based determination of model parameters. The probability for finding population values higher than $P = 8$ is defined by $p_8(t) = 1 - \Pr(P \leq 8)$. By using $\Pr(P \leq p) = F(p)$, the corresponding result implied by the population CDF model (3.9) is given by

$$p_8(t) = 1 - F(8) = \frac{1}{2}\Big[1 - \mathrm{erf}\Big(\frac{H(8)}{\sqrt{2}\sigma}\Big)\Big]. \tag{6.3}$$

For the range of σ values considered here, the curve resulting from equation (6.3) is compared with the result of Scherbov, Lutz and Sanderson [34] and the predictions of the UN high, median, and low variants in Figure 7. It may be seen that there is a significant discrepancy between the predictions of Scherbov, Lutz and Sanderson and the results implied by the population model (3.9). The reason for the discrepancy between both models is that the model (3.9) developed here is based on the recent UN projections, whereas the model of Scherbov, Lutz and Sanderson is based on a combination of their previous probabilistic assumptions [26–28] with the 2010 UN data. In difference to the recent median UN prediction (which assumes that the population levels off asymptotically: see Figure 1), the previous probabilistic assumptions of Scherbov, Lutz and Sanderson assume in their median variant that the population will decrease asymptotically: see Figure 2 in reference [26]. As a consequence, the probability $p_8(t)$ is much lower in the model of Scherbov, Lutz and Sanderson than in the model presented here. The trend of a decreasing world population

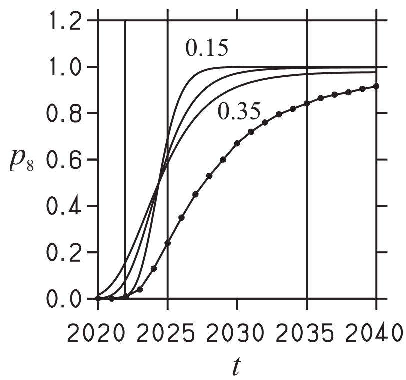

Figure 7. The probability p_8 for finding population values higher than $P = 8$ is shown as a function of time. The solid and dashed lines show the results according to equation (6.3), which follow from the population CDF model (3.9): the solid line shows p_8 for $\sigma = 0.25$ and the dashed lines show p_8 for $\sigma = 0.15$ and $\sigma = 0.35$ (the value of σ is given at the curve). The solid line with dots shows the corresponding result obtained by Scherbov, Lutz and Sanderson [34]. The vertical dashed lines show (from left to right) the consequences of the recent UN high, median, and low projections that the world population reaches 8 billion by 2022, 2025, and 2035, respectively.

after 2070 (which is reflected in the median variant of the model of Scherbov, Lutz and Sanderson) has to be seen as an empirical assumption for which evidence is unavailable. Scherbov, Lutz and Sanderson do not compare their results with independent predictions. In contrast to that, the validation of the stochastic population model presented in Section 5 provided support for the suitability of the world population PDF model (3.10).

7 Summary

Existing methods for quantifying the uncertainty of the world population development face a variety of questions: see the discussion in the introduction. An essential problem of all such methods is the lack of direct evidence (e.g., by means of comparisons with independent observations like measurements). A way to support the validity of such forecast methods is to validate such models with reference models, which play the role of independent observations. Desired properties of such a reference model were formulated here and a reference world population model was formulated by a probabilistic extension of recent deterministic UN projections. The validation of this model in terms of theory, observations, and the benefits of using the model presented will be summarized in the following three paragraphs, respectively.

A theoretical validation of the suitability of the population PDF model (3.10) formulated here was performed by asking whether this model has the desired properties (P1)-(P5) of a reference model. The PDF model (3.10) was formulated as a model for the world population development, which avoids problems of other approaches based on country projections. The PDF model does only depend on the standard deviation σ of the fertility PDF, which specifies the variability of deviations from the median fertility assumption (the relatively minor model sensitivity to σ variations was discussed in terms of Figure 4 (b)). The PDF model represents an analytical and bridging model that relates the world population PDF to deterministic scenarios. The simple and relevant relation $p_{LH}(t) = \mathrm{erf}[1/(2^{3/2}\sigma)]$ was obtained, which explains how the probability $p_{LH}(t)$ covered by the UN high and low variants is related to the standard deviation σ of the fertility PDF. The fact that the PDF model is extremely well approximated by a self-similar PDF ensures a correct model formulation. Thus, the PDF model (3.10) has all the desired properties (P1)-(P5) of a reference model.

The suitability of the PDF model (3.10) was also validated by means of comparisons with existing observations. Such a validation faces non-trivial questions related to the concrete model compared to observations and the comparison of a probabilistic prediction with deterministic observations: see the discussions in Section 5. The latter questions were addressed by showing that the world population PDF model (3.10), which provides a probabilistic forecast for $2010 \le t \le 2100$, can be validated by the known $1980 \le t \le 2010$ population data. Both deterministic and probabilistic methods were used to quantify the PDF model performance. These comparisons showed that the PDF population model (3.10) formulated here performs very well (for a period of at least thirty years) in comparison with available population data.

The benefits of the population model (3.10) developed here were discussed in Section 6. Compared to the deterministic UN projections, one relevant advantage of the stochastic world population model presented here is the quantification of the uncertainty range of the median UN projection, including a determination of the accuracy of the uncertainty calculation. An acceptable population forecast uncertainty of $(4.3 \pm 1.7)\%$ was found for the first thirty years. Another relevant advantage of the PDF model (3.10) is that this model can be used as a reference model for comparisons with other deterministic and stochastic methods: it can be used to calculate the probability for conclusions of deterministic forecast methods and to prove the validity of conclusions obtained by probabilistic methods. Two examples for the application of the PDF model as reference model were shown. A comparison with consequences of the deterministic World3 model showed that the conclusions obtained by the World3 model are not very likely. The stochastic world population model was also used for a comparison with the probabilistic model of Scherbov, Lutz and Sanderson [34]. It was shown that the probabilistic model of Scherbov, Lutz and Sanderson and the PDF model presented here, which was (in contrast to the model of Scherbov, Lutz and Sanderson) validated by theoretical constraints and observations, do not lead to the same conclusions. The reasons for these differences were discussed.

It is worth emphasizing that the stochastic world population model presented is relatively simple, this means the model does not take into account all influence factors that can be expected. The only parameter involved is the constant standard deviation σ, which determines the total fertility rate distribution. Country correlations can be accounted for (see the discussion in Section 3), but they are not explicitly considered. Nevertheless, the model presented here is fully consistent with theoretical constraints for the development of such probabilistic models and recent and former UN forecasts. Hence, the model presented can be expected to represent a reasonable approximation to the complex development seen in reality. Thus, the model can be used as a reference tool for the evaluation of the suitability of more complex models. As demonstrated in terms of the comparison with the probabilistic model of Scherbov, Lutz and Sanderson, such inter-model comparisons can raise relevant questions about the modeling assumptions applied, and they may contribute to a harmonization of the assumptions used in various probabilistic projections.

A Appendix: Self-Similar PDF

To prepare the derivation of the self-similar population PDF, let us consider first the stochastic model and PDF of the standardized population density $P_* = (P - \langle P\rangle)/\sigma_P$. By replacing ϵ by the standardized variable $\epsilon_* = (\epsilon - \langle\epsilon\rangle)/\sigma_\epsilon$, where the mean and standard deviation of ϵ are given by $\langle\epsilon\rangle = K_2\sigma^2$ and $\sigma_\epsilon = [K_1^2\sigma^2 + 2\langle\epsilon\rangle^2]^{1/2}$, respectively, we can write equation (3.1) as

$$P = m_3 - \frac{m_1}{\sigma_\epsilon \epsilon_* + m_2 + \langle\epsilon\rangle} = m_3 - m_1 \frac{W/\sigma_\epsilon}{1 + W\epsilon_*}, \tag{A.1}$$

where the relative weight W of randomness $W = \sigma_\epsilon/(m_2 + \langle\epsilon\rangle)$ is introduced. The standardized population density $P_* = (P - \langle P\rangle)/\sigma_P$ then reads

$$P_* = \frac{1}{\sigma_P}\left[m_3 - \frac{m_1 W}{\sigma_\epsilon} Z - \langle P\rangle\right], \tag{A.2}$$

where we introduced the abbreviation $Z = 1/(1 + W\epsilon_*)$. By definition, there are two requirements for P_* given by $\langle P_*\rangle = 0$ and $\langle P_*^2\rangle = 1$. The first requirement $\langle P_*\rangle = 0$ implies the condition

$$\langle P\rangle = m_3 - \frac{m_1 W \langle Z\rangle}{\sigma_\epsilon}.$$

Consequently, we can write

$$P_* = -\frac{m_1 W \widetilde{Z}}{\sigma_P \sigma_\epsilon},$$

where $\widetilde{Z} = Z - \langle Z \rangle$. The second requirement $\langle P_*^2 \rangle = 1$ then implies

$$\sigma_P = \frac{m_1 W \langle \widetilde{Z}^2 \rangle^{1/2}}{\sigma_\epsilon}.$$

In terms of the latter relation we can write $P_* = -\widetilde{Z}/\langle \widetilde{Z}^2 \rangle^{1/2}$. Thus, the standardized stochastic population density is given by the negative standardized Z. The PDF of P_*, this means the nonself-similar standardized population PDF (3.10), is given by

$$f_*(p_*) = \frac{\sigma_P}{\sqrt{2\pi\sigma^2}} \frac{m_1}{(p - m_3)^2} \frac{\exp\{-\frac{H^2(p)}{2\sigma^2}\}}{K_1 + 2K_2 H(p)}. \tag{A.3}$$

This PDF is a function of p_* via $p = \sigma_P p_* + \langle P \rangle$. A closed form expression for this PDF requires expressions for the population mean $\langle P \rangle$ and standard deviation σ_P. According to

$$\langle P \rangle = m_3 - \frac{m_1 W \langle Z \rangle}{\sigma_\epsilon} \quad \text{and} \quad \sigma_P = \frac{m_1 W \langle \widetilde{Z}^2 \rangle^{1/2}}{\sigma_\epsilon},$$

the calculation of $\langle P \rangle$ and σ_P requires explicit expressions for $\langle Z \rangle$ and $\langle \widetilde{Z}^2 \rangle$. The latter two variables are determined by the PDF of Z, which can be explicitly calculated by following again the approach described in Section 3. However, it is impossible to use these integral representations for the analytical calculation of $\langle Z \rangle$ and $\langle \widetilde{Z}^2 \rangle$. Thus, the only way to provide $\langle Z \rangle$ and $\langle \widetilde{Z}^2 \rangle$ is their numerical calculation either by Monte Carlo simulation or the numerical solution of integrals.

Let us consider now whether the nonself-similar standardized population PDF (A.3) is well approximated by a self-similar PDF. According to $P_* = -\widetilde{Z}/\langle \widetilde{Z}^2 \rangle^{1/2}$, $Z = 1/(1 + W\epsilon_*)$ determines the standardized stochastic population density P_*. The PDF of P_* will change its shape in time due to the variation of $W = \sigma_\epsilon/(m_2 + \langle \epsilon \rangle)$ via changes of $m_2(t)$. The maximum of $W(t)$ is given for a minimal $m_2(t)$. An analysis shows that $m_2(t)$ has a minimum for $t = 2047.8$. For $t = 2047.8$, it turns out that W is bounded from above by $W = (0.0172, 0.0122, 0.0073)$ for the considered $\sigma = (0.35, 0.25, 0.15)$, respectively. Hence, W is very small, i.e., the fraction involving $W\epsilon_*$ in $Z = 1/(1 + W\epsilon_*)$ can be expanded in terms of its Taylor series (ϵ_* is of the order of one). The first-order Taylor series expansion of Z reads $Z = 1 - W\epsilon_*$. Thus, we find $\langle Z \rangle = 1$, $\widetilde{Z} = Z - \langle Z \rangle = -W\epsilon_*$, and $\langle \widetilde{Z}^2 \rangle = W^2$, which implies $P_* = \epsilon_*$ according to $P_* = -\widetilde{Z}/\langle \widetilde{Z}^2 \rangle^{1/2}$. The PDF of P_* is self-similar because the statistics of ϵ_* do not change in time. The PDF of $P_* = \epsilon_*$ can be obtained in the following way. The CDF $F_*(p_*) = \Pr(\epsilon_* \le p_*)$ is given by

$$F_*(p_*) = \Pr\left(\frac{K_1 c + K_2(c^2 - \sigma^2)}{\sigma_\epsilon} \le p_*\right) = \Pr(c \le H_*(p_*)), \tag{A.4}$$

where we introduced the abbreviation

$$H_*(p_*) = \frac{K_1}{2K_2}\left\{\sqrt{1 - 4\frac{K_2}{K_1^2}(\sigma_\epsilon p_* + \langle \epsilon \rangle)} - 1\right\}. \tag{A.5}$$

The implied standardized world population PDF $f_*(p_*) = dF_*/dp_*$ is found to be given by

$$f_*(p_*) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{H_*^2(p_*)}{2\sigma^2}\right\} \frac{dH_*}{dp_*} = \frac{\sigma_\epsilon}{\sqrt{2\pi\sigma^2}} \frac{\exp\{-\frac{H_*^2(p_*)}{2\sigma^2}\}}{K_1 + 2K_2 H_*(p_*)}. \tag{A.6}$$

This expression represents the self-similar PDF approximation of the nonself-similar standardized population PDF (A.3). The PDF (A.6) can be used for the representation of the non-standardized population PDF by using the relations $f_* = f\sigma_P$ and $p_* = (p - \langle P \rangle)/\sigma_P$ between both PDFs. The accuracy of the approximation $Z = 1 - W\epsilon_*$, which implies the self-similar PDF (A.6), can be evaluated by calculating the relative error E of this approximation. This leads to the expression

$$E = \frac{\frac{1}{1+W\epsilon_*} - (1 - W\epsilon_*)}{\frac{1}{1+W\epsilon_*}} = W^2 \epsilon_*^2. \tag{A.7}$$

Hence, the mean relative error is given by $E = W^2$. For the range $0.15 \le \sigma \le 0.35$ considered we have $W \le 0.0172$. Hence, the mean error is extremely small, $E \le 0.03\%$.

Regarding the discussion in Section 6 it is worth noting that the approximation $Z = 1 - W\epsilon_*$ implies explicit expressions for the world population mean and standard deviation. According to

$$\langle P\rangle = m_3 - \frac{m_1 W\langle Z\rangle}{\sigma_\epsilon} \quad \text{and} \quad \sigma_P = \frac{m_1 W\langle \widetilde{Z}^2\rangle^{1/2}}{\sigma_\epsilon}$$

combined with $\langle Z\rangle = 1$ and $\langle \widetilde{Z}^2\rangle = W^2$ implied by $Z = 1 - W\epsilon_*$ we find

$$\langle P\rangle = m_3 - \frac{m_1 W}{\sigma_\epsilon} \quad \text{and} \quad \sigma_P = \frac{m_1 W^2}{\sigma_\epsilon} \tag{A.8}$$

for the world population mean and standard deviation. For the t and σ values considered, this means for $2010 \le t \le 2100$ and $0.15 \le \sigma \le 0.35$, it turns out that $m_2 W/\sigma_\epsilon$ is bounded by $0.998 \le m_2 W/\sigma_\epsilon \le 1$. Thus, $W/\sigma_\epsilon = 1/m_2$ represents an excellent approximation. In addition, we find $\sigma_\epsilon/(K_1\sigma) = (1.0023, 1.0065, 1.0126)$ for the $\sigma = (0.15, 0.25, 0.35)$ values considered. Consequently, $\sigma_\epsilon = K_1\sigma$ represents a very accurate approximation. The use of the approximations $W/\sigma_\epsilon = 1/m_2$ and $\sigma_\epsilon = K_1\sigma$ in combination with equation (A.8) enables us to represent the population mean and variance by the very accurate approximations

$$\langle P\rangle = m_3 - \frac{m_1}{m_2} \quad \text{and} \quad \sigma_P = K_1\sigma\frac{m_1}{m_2^2}. \tag{A.9}$$

Acknowledgement: I would like to thank Dr. G. Heilig (United Nations, Chief, Population Estimates and Projections Section, DESA – Population Division) for several interesting discussions and valuable comments.

References

[1] M. Alders, N. Keilman and H. Cruijsen, Assumptions for long-term stochastic population forecasts in 18 European countries, *Eur. J. Population* **23** (2007), 33–69.

[2] J. Alho, Aggregation across countries in stochastic population forecasts, *Intern. J. Forecasting* **24** (2008), 343–353.

[3] J. M. Alho and B. D. Spencer, The practical specification of the expected error of population forecasts, *J. Official Statist.* **13** (1997), 203–225.

[4] J. M. Alho and B. D. Spencer, *Statistical Demography and Forecasting*, Springer-Verlag, New York, 2005.

[5] L. Alkema, A. E. Raftery, P. Gerland, S. J. Clark and F. Pelletier, Estimating the total fertility rate from multiple imperfect data sources and assessing its uncertainty, Working Paper 89, Center for Statistics and the Social Sciences, University of Washington, Seattle, 2008.

[6] L. Alkema, A. E. Raftery, P. Gerland, S. J. Clark, F. Pelletier and T. Buettner, Probabilistic projections of the total fertility rate for all countries, Working Paper 97, Center for Statistics and the Social Sciences, University of Washington, Seattle, 2010.

[7] L. Alkema, A. E. Raftery, P. Gerland, S. J. Clark, F. Pelletier, T. Buettner and G. K. Heilig, Probabilistic projections of the total fertility rate for all countries, *Demography* **48** (2011), 815–839.

[8] U. Bardi, *The Limits to Growth Revisited*, Springer-Verlag, New York, 2011.

[9] J. Bijak and A. Wiśniowski, Bayesian forecasting of immigration to selected european countries by using expert knowledge, *J. R. Stat. Soc. Ser. A* **173** (2010), 775–796.

[10] F. C. Billari, R. Graziani and E. Melilli, Stochastic population forecasts based on conditional expert opinions, *J. R. Stat. Soc. Ser. A* **175** (2012), 491–511.

[11] J. Bongaarts and R. A. Bulatao, Appendix F: Estimating expected errors from past errors, in: *Beyond Six Billion. Forecasting the World's Population*, National Academy Press, Washington DC (2000), 326–348.

[12] J. L. Chunn, A. E. Raftery and P. Gerland, Bayesian probabilistic projections of life expectancy for all countries, Working Paper 105, Center for Statistics and the Social Sciences, University of Washington, Seattle, 2010.

[13] R. Costanza, R. Leemans, R. Boumans and E. Gaddis, Integrated global models, in: *Sustainability or Collapse? An Integrated History and Future of People on Earth*, MIT Press, Cambridge (2007), 417–446.

[14] F. Girosi and G. King, *Demographic Forecasting*, Princeton University Press, Princeton, 2008.

[15] H. C. J. Godfray, J. R. Beddington, I. R. Crute, L. Haddad, D. Lawrence, J. F. Muir, J. Pretty, S. Robinson, S. M. Thomas and C. Toulmin, Food security: The challenge of feeding 9 billion people, *Science* **327** (2010), 812–818.

[16] D. G. Heilig, United Nations, Chief, Population Estimates and Projections Section, DESA – Population Division, personal communication.

[17] S. Heinz, *Statistical Mechanics of Turbulent Flows*, Springer-Verlag, Berlin, 2003.

[18] S. Heinz, *Mathematical Modeling*, Springer-Verlag, Heidelberg, 2011.

[19] R. J. Hyndman and H. Booth, Stochastic population forecasts using functional data models for mortality, fertility and migration, *Intern. J. Forecasting* **24** (2008), 323–342.
[20] N. Keilman, D. Q. Pham and A. Hetland, Why population forecasts should be probabilistic – illustrated by the case of Norway, *Demogr. Res.* **6** (2002), 409–454.
[21] R. Lee, The outlook for population growth, *Science* **333** (2011), 569–573.
[22] R. D. Lee and S. Tuljapurkar, Stochastic population forecasts for the united states: Beyond high, medium, and below, *J. Amer. Statist. Assoc.* **89** (1994), 1175–1189.
[23] W. Lutz and K. C. Samir, Dimensions of global population projections: What do we know about future population trends and structures?, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **365** (2010), 2779–2791.
[24] W. Lutz, W. C. Sanderson and S. Scherbov, Doubling of world population unlikely, *Nature* **387** (1997), 803–805.
[25] W. Lutz, W. C. Sanderson and S. Scherbov, Expert-based probabilistic population projections, *Popul. Dev. Review* **24** (1998), suppl. S, 139–155.
[26] W. Lutz, W. C. Sanderson and S. Scherbov, The end of world population growth, *Nature* **412** (2001), 543–545.
[27] W. Lutz, W. C. Sanderson and S. Scherbov, *The End of World Population Growth in the 21th Century. New Challenges for Human Capital Formation & Sustainable Development*, Earthscan, London, 2004.
[28] W. Lutz, W. C. Sanderson and S. Scherbov, The coming acceleration of global population ageing, *Nature* **451** (2008), 716–719.
[29] D. H. Meadows, D. L. Meadows, J. Randers and W. W. Behrens III, *The Limits to Growth: A Report for the Club of Rome's Project on the Predicament of Mankind*, Universe Books, New York, 1972.
[30] D. H. Meadows, D. L. Meadows, J. Randers and W. W. Behrens III, *The Limits to Growth: A Report for the Club of Rome's Project on the Predicament of Mankind*, 2nd ed., Universe Books, New York, 1974.
[31] D. H. Meadows, J. Randers and D. L. Meadows, *Limits to Growth: The 30-Year Update*, Chelsea Green Publishing, White River Junction, 2004.
[32] D. L. Meadows, W. W. Behrens III, D. H. Meadows, R. F. Naill, J. Randers and E. K. O. Zahn, *Dynamics of Growth in a Finite World*, Wright–Allen Press, Cambridge, MA, 1974.
[33] W. C. Sanderson, S. Scherbov, B. C. O'Neill and W. Lutz, Conditional probabilistic population forecasting, *Intern. Statist. Review* **72** (2004), 157–166.
[34] S. Scherbov, W. Lutz and W. C. Sanderson, The uncertain timing of reaching 8 billion, peak world population, and other demographic milestones, *Popul. Dev. Review* **37** (2011), 571–578.
[35] G. M. Turner, A comparison of the limits to growth with 30 years of reality, *Global Environ. Chang.* **18** (2008), 397–411.
[36] *World Population Prospects as Assessed in 1980*, ST/ESA/SER.A/78, Population Studies No. 78, Department of International Economic and Social Affairs, United Nations, New York, 1981.
[37] *World Population Prospects: The 2010 Revision. Excel Tables – Population Data*, United Nations, New York, 2011, http://esa.un.org/unpd/wpp/Excel-Data/population.htm

Received May 3, 2014; accepted October 20, 2014.