The UNHCR Demographic Projection Tool: estimating the future size and composition of forcibly displaced populations

Contributed by: Diego Alburez-Gutierrez[1] and Carlota Segura García[2]

The United Nations High Commissioner for Refugees (UNHCR) has the mandate to protect and assist forcibly displaced populations around the world. The agency constantly collects and analyses demographic data on these populations to understand their composition, geographic distribution, and overall needs. The main purpose of this demographic data is to provide the evidence base for programmatic response planning and advocacy in support of durable solutions for persons of concern (PoCs). The aggregated data published by UNHCR is also an important resource for academic and policy-oriented researchers. It is also used for evidence-based advocacy regarding the state of the world's forcibly displaced population.

Understanding the demographic composition of forcibly displaced populations helps UNHCR workers to allocate resources, plan, and respond in support of meeting urgent needs and implementing durable solutions. Data about the demographic breakdown of a population are particularly important to inform the design of interventions because needs differ by age and sex. For example, demographic data can be used to guide interventions directed at women of childbearing age or school-age children. This clearly benefits both the population in need of intervention and UNHCR, which can more effectively allocate resources. For the same reason, it is helpful to be able to project what a population will look like in the future.

UNHCR operations routinely make estimates about the future size of PoC groups. However, there is presently no systematic approach or standard guidance to help UNHCR staff arrive at these figures. Outside of conflict-induced displacement contexts, other organs of the United Nations have a long history of projecting the age

[1] UNHCR Field Information and Coordination Support Section (FICSS).

[2] Independent researcher.

and sex composition of human populations (UN DESA 1955), and demographers have developed different methodologies to project the future size and composition of populations (Siegel & Swanson 2004).

In this paper we summarise UNHCR's experience of applying demographic methods for projecting forcibly displaced populations in conflict situations. The objective of this paper is to examine the use of cohort-component methods for projecting populations of PoC. To do this we focus on the development of the Demographic Projection Tool (DPTool)[3], a project led by FICSS (Field Information and Coordination Support Section), a section within the Division of Programme Support and Management (DPSM). The paper provides a brief introduction to UNHCR's data on PoCs before outlining the principles of population projections in general and the design of the DPTool in particular. Relying on context-specific data is indispensable for projecting populations of concern, since UNHCR data show that the particular age-sex distribution of a displaced population is heavily dependent on the characteristics of the displacement event. Furthermore, the structure of these populations often differs considerably from that of non-displaced populations, even within the same country of origin. In this paper we provide a brief introduction to the registry data collected by UNHCR and discuss some of their limitations. In the last section we discuss lessons learned from implementing the projection tool in Kenya for the 2017-19 period.

Background

UNHCR Persons of Concern

The population of concern to UNHCR includes individuals who have been forcibly displaced (refugees, asylum-seekers, and internally displaced persons—IDPs), those who have been able to return to their countries or areas of origin after displacement (returnees and returned IDPs), stateless persons, and other groups of people who do not fall under UNHCR's mandate but to whom it has extended its protection. The size of the forcibly displaced population around the world reached 65.6 million at the end of 2016, the highest level since the end of World War Two. Syria, Colombia, and Afghanistan are the three most common countries of origin for refugees (UNHCR

[3] The DPTool is available online: http://DemographicProjection.unhcr.org/

2007a). The empirical section of this paper focuses on the refugee population, which by June 2017 amounted to about 24.5 million worldwide, including the approximate 5.3 million Palestinian refugees under UNRWA's mandate but excluding the more than 650,000 refugees who have fled from Myanmar to Bangladesh since August 2017 (UNHCR 2017b, UNRWA 2017). Most of the refugee populations of concern to UNHCR are located in Africa and the Middle East, but Turkey continues to host the largest number of refugees— mostly Syrian. Nine out of the ten major refugee-hosting countries (with the exception of Germany) are low- and middle-income countries.

Data on Persons of Concern to UNHCR

In this paper we consider UNHCR's PoC as a demographic population: a collection of individuals that meet certain criteria at a given point in time. This population can only change through a) natural increase[4]; b) forced displacement, migration, or resettlement[5]; and c) changes in administrative status. Natural increase includes births and deaths in the population. To understand the second component of population change, it is necessary to distinguish between 'stock' and 'flow' populations. The former refers to the existing population at a specific time. The stock of PoC in Kenya at the start of 2017, for example, are all the PoC in the country at that time. The flow, by contrast, is the sum of those who join or leave the population over a specific period. Lastly, changes in administrative status also affect the population of PoC—refugees that become citizens, for example, no longer fall under UNHCR's mandate. The demographic balancing equation shown below summarises this equivalence. Note that *arrivals* and *departures* in equation (1) include changes in status (e.g. individuals who successfully applied for asylum)[6]. Figure 1 below summarises this graphically.

$$P_t = P_{t-1} + Births - Deaths + Arrivals - Departures \quad (1)$$

[4] The difference between the number of births and deaths over a given period.

[5] Refugees and IPDs should not be confused with economic migrants or those who migrate due to climate change or natural disasters.

[6] We use 'arrivals' and 'departures' in this way throughout the document in order to be consistent with the demographic literature.

UNHCR's work in 130 countries is managed by a number of operations organised in different types of hierarchies. Some operations cover several countries (e.g. the Panama Regional Operation covers all of Central America). The Profile Global Registration System (proGres) is UNHCR's enterprise system for registration and case management of refugees and asylum-seekers. It was introduced in 2004 to implement the registration standards defined in ExCom Conclusion No. 91 on the Registration of Refugees and Asylum-Seekers (i.e. to substitute a set of unstandardized registration systems).[7] The most recent version of the system, proGres v.4, consists of over fifty relational tables containing micro-data collected during the registration process and in subsequent verification and follow-up activities. ProGres v.4, which is expected to be deployed in all relevant operations by the end of 2018, offers a centralised and secure repository for registration data, as well as a harmonized platform for data input and sharing. Individual-level proGres data are not publicly available, but the aggregate data are regularly updated and made available online[8].

UNHCR proGres data are not available to the agency in contexts were registration is carried out by the government or its partners. In some cases, individual-level data on PoC are available in a different format or not available at all. The lack of registration data creates problems of compatibility and harmonisation and in some cases makes it difficult to obtain basic demographic information disaggregated by age and sex.

That being said, all operations where UNHCR is in charge of registration do use an instance of proGres that includes individual-level data on country of origin, country of asylum, date of birth, gender, date of arrival, and date of registration. A separate set of codes record the status of a given record (e.g. 'active' or 'inactive'). Records are 'closed' after a PoC leaves the population definitively—for example, when a 'solution' is achieved (i.e., voluntary repatriation, resettlement, or local integration).

The registration of large numbers of refugees and asylum-seekers presents unique challenges that affect the accuracy of the data. For example, preliminary analysis of

[7] Note that the data on refugees and asylum-seekers presented in UNHCR's Global Trends report is based on official data provided by national statistical offices, not on proGres data.

[8] http://popstats.unhcr.org

proGres data showed that dates and ages are affected by response heaping (e.g. rounded to the nearest multiple of 5) and birth registration is likely to be incomplete (i.e. children of PoCs born in displacement counted as new arrivals, not as new births).[9] A separate report in this series will focus on the quality of proGres data and suggest methods for improving it. Nonetheless, UNHCR registration data continues to be the most comprehensive repository of microdata on refugees and asylum-seekers around the world.

Methods

The cohort-component method for projecting populations

All population projections rely on historical data and a set of assumptions about demographic behaviour to estimate the size and composition of a population in the future. The cohort-component method is one of the most common projection methodologies. The UN population forecasts were produced using this technique up until 2014, when Bayesian methods were incorporated to quantify the uncertainty around projection estimates (Raftery et al., 2014; UN DESA, 2017). Other projection methods, including trend extrapolation and structural models, are also available (Siegel and Swanson, 2004).

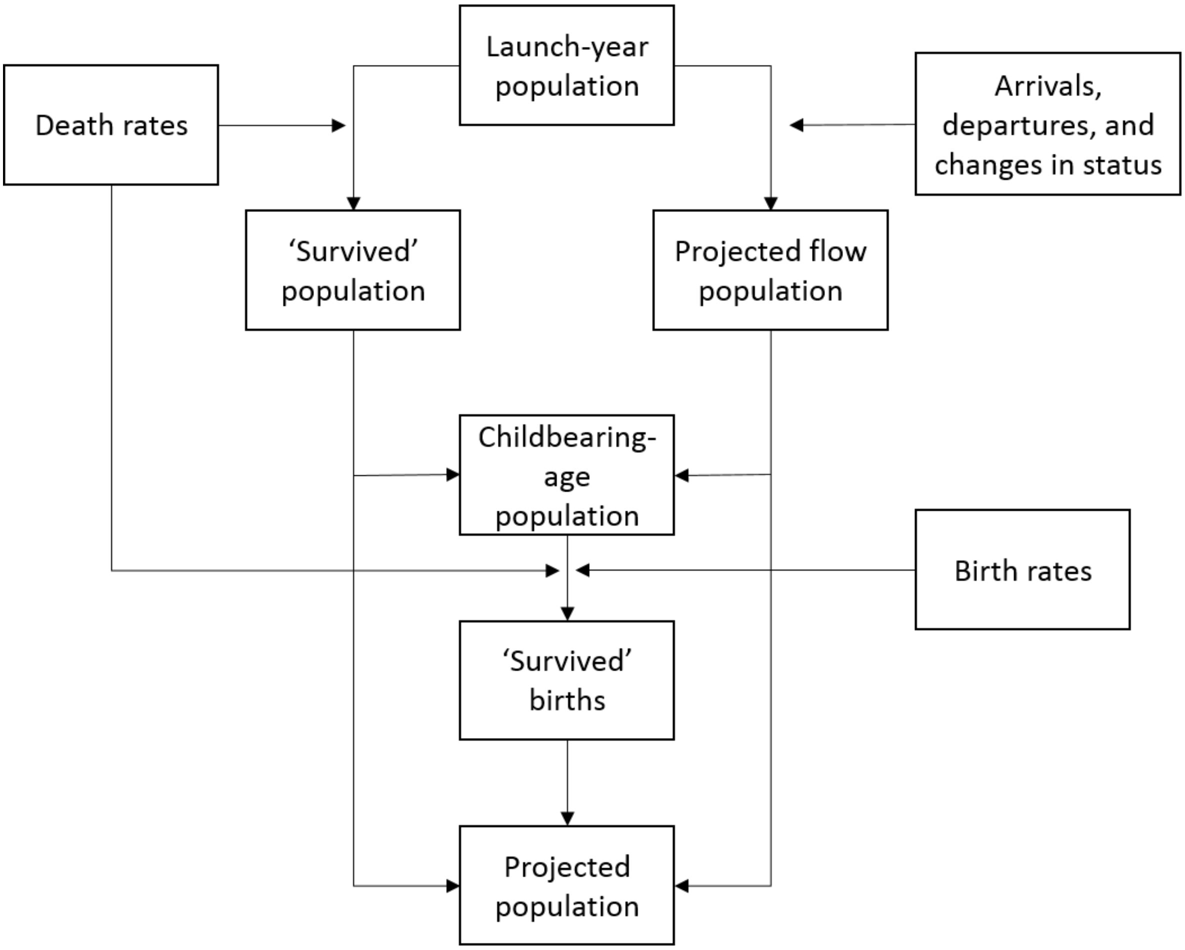
Cohort-component methods project births, deaths, and migration separately for each age-sex group of an initial or *launch population*, usually using single-year or 5-year birth cohorts. The cohort-component method can be applied to national and subnational populations. The simplest variants of the method assume that the age-specific rates remain constant over the projected time, or *projection horizon*. More complex models allow the rates to vary over time (Siegel and Swanson, 2004). The base demographic rates may be obtained from administrative records, censuses, surveys, or derived from model life tables[10]. In the absence of substantial migratory flows, fertility account for the largest share of the variation in long-term projections.

---

[9] An analysis of the 2017 Kenya proGres v.3. data showed an unusually large number of age 0 arrivals not recorded as births. Many of these were likely unregistered births.

[10] Life tables are demographic instruments that summarise the mortality experience of a given populations. Model life tables have been constructed from reference populations (see UN 1982).

Figure 1. Cohort-component projection for a given projection interval.



Adapted from Smith *et al.* (2001, p.41)

Migration rates, however, are likely to be especially relevant in contexts of forced migration. Unfortunately, migration is the most volatile component of population change and is generally more susceptible to abrupt changes (Smith & Swanson, 2002). This is especially true for contexts of forced displacement, where it is nearly impossible to make robust assumptions about long-term migratory flows. Knowledge of the situation in the field and awareness of the geopolitical setting may be more relevant for making valid short-term assumptions about the magnitude of forced migration flows. Figure 1 summarises the process of projecting a population using a cohort component-method. Applications of the technique vary in practice, but most implementations start by (1) applying age-specific mortality rates to a launch population, (2) adding or subtracting the number of migrants by sex and age, (3) applying age-specific fertility rates to women in childbearing age and infant mortality rates to the new births.

The UNHCR Demographic Projection Tool (DPTool)

All UNHCR operations conduct yearly planning exercises to request budgets and help determine the optimal distribution of resources. Planning figures, including the estimated future size and composition of PoC groups, are a key component of these exercises. In the absence of systematic methodologies, planning figures are estimated using a plethora of different methods that consider trends from previous years and assumptions about natural growth and future mortality rates. Generally, the sophistication of the methodology used by a given operation is determined by the availability and quality of existing data and the time and resources available.

In this context, UNHCR FICSS developed the DPTool in 2017-2018 to apply demographic projection methodologies for estimating population planning figures. The tool takes advantage of the new data collection and management practices introduced by the proGres v.4 system. The main objective of the projection tool is to promote a consistent and harmonized approach to generating planning figures across operations. The DPTool is intended to project population of PoC disaggregated by age, sex, origin, and population type within a 3 years projection horizon.

The tool was developed by the creators of the Spectrum projection package and, like the DemProj model in Spectrum (Stover 2007), it combines data on a launch-year population with a series of assumptions about demographic behaviours to create cohort-component projections within a given projection horizon. The model requires inputs about (a) base population by age and sex, (b) fertility, (c) mortality, and (d) forced migration. It is theoretically possible to obtain (a) to (c) from the proGres data but these estimates could be prone to error. Underreporting of births and deaths and unregistered changes in the status of PoCs could affect the estimation of fertility and mortality rates. Therefore, subject to future research, demographic rates are obtained from the World Population Prospects, WPP (UN DESA 2017), data relying on the assumptions summarised in Table 1 (e.g. life expectancy at birth is assumed to be that of the country of asylum). Fertility rates, life expectancy, and the sex ratio at birth all change throughout the projection horizon, based on WPP trends.

Table 1. Data sources and assumptions for the DPTool.

Input	Demographic component	Source	
		proGres	WPP (if proGres unavailable)
Launch-year population	Total population by sex (stock)	Yes	No
	Age-sex distribution (stock)	Yes	Yes (origin country)
Mortality	Life expectancy at birth by sex	No	Yes (asylum country)
	Age-specific mortality rates	No[a]	Yes (asylum country)
Fertility	Total Fertility Rate	No[a]	Yes (origin country)
	Age-specific fertility	No[a]	Yes (origin country)
Forced migration	Total new arrivals (predicted)	No	No
	Total departures (predicted)	No	No
	Age distribution of arrivals and departures	Yes[b]	Yes (origin country)

a. Currently not advisable, given under-registration of birth and death registration.
b. Using data on the previous year's flow population. See details in text.

Assumptions about the population flow deserve special attention since displacement is a major driver of demographic change in refugee populations. The DPTool accounts for net migration, meaning that separate data on in-migration (arrivals) and out-migration (departures) are required. At present, there is no systematic way of estimating the future size of PoC groups because the causes of forced displacement are difficult to predict.[11] This means that users are prompted to estimate the size of the future PoC arrivals and departures (i.e. through voluntary repatriation, resettlement, or local integration). The DPTool provides no guidance on how to reach this figure, but operations routinely rely on a combination of expert knowledge and first-hand experience. In this respect, it is good practice to produce multiple

[11] As of the time of writing (May 2018), the UNHCR Innovation Service was working on predicting refugee flows using online social media data. For more details, see: http://www.unhcr.org/innovation/tag/jetson/.

projections to represent alternative scenarios—e.g. a new massive refugee influx, or a growing number of repatriations due to conflict resolution.

As mentioned before, the DPTool applies a cohort component method to project a given type of population of PoC by age and sex into the future. Initially, the base population is adjusted for mortality and forced migration, with arrivals and departures being subject to a half year of mortality. This includes making a number of assumptions about the age-sex distribution of the migrant flow population. If proGres data are available, it is assumed that the structure of future refugee flows will resemble that of past flows. In practice, this means that the age-sex composition of the incoming population is estimated by analysing the population that arrived during *t-1*, the year before the launch-year. The composition of the outgoing population, on the other hand, can be obtained by considering the records that were resettled or repatriated during *t-1*.

The model requires the population to be disaggregated by single (i.e. one-year) age groups. Aggregate age groups (e.g. 0-4, 5-11, 12-17, 18-59, 60+) are hence split into single ages according to a model pattern estimated by calculating the median percentage of the broad age group in each single age within the group for all the countries in the world in 2015 (UN DESA 2017). An alternative approach is needed for cases where proGres data are unavailable for *t-1*, either because the displacement event is recent or because UNHCR was not in charge of registration at the time. In these situations, the DPTool assumes that the population flow has an equivalent age-sex structure to that of the non-displaced population in the country of origin. This assumption is less than ideal since the demographic composition of refugee populations differs considerably from the distribution of non-displaced population in their countries of origin[12].

$$P_{a,s,t} = P_{a-1,s,t-1} \times sv_{a,s,t} + (A_{a,s,t} - D_{a,s,t}) \times (1 - \frac{(1 - sv_{a,s,t})}{2}) \quad (2)$$

Equation (2) shows how the model accounts for mortality and forced migration for individuals in age groups 1 to 79. In it, $P_{a,s,t}$ is the population of each age *a*, sex *s*, and

[12] As shown by internal analysis of UNHCR registration data. The forthcoming paper on proGres data referred to above will explore this further.

time *t*; *A* and *D* stand for the total number of arrivals and departures; and *sv* represents the survival rate.

Estimates of the population in age group 80+ are a sum of the population aged 79 at *t-1* that survives to age 80, and the population aged 80 or older at *t-1* that survives to year *t*, accounting for arrivals and departures.

$$P_{80+,s,t} = P_{80,s,t-1} \times sv_{80+,s,t} + P_{80+,s,t-1} \times sv_{80+,s,t} + (A_{a,s,t} - D_{a,s,t}) \times \left(1 - \frac{1-sv_{a,s,t}}{2}\right) \quad (3)$$

Finally, the population at age 0 is calculated by multiplying the number of births during the year, *B,* by the survival rate at birth, sv_{birth}, to obtain the number of new-borns that survive to the end of the year.

$$P_{0,s,t} = B_{s,t} \times sv_{birh,t} \quad (4)$$

The number of births, required by equation (4), is estimated from the total fertility rate, *TFR*, the age-specific fertility rates, *ASFR,* and the average number of women in reproductive age (i.e. 15-49) over the past year summed over the reproductive ages, *W.*

$$B_{s,t} = \sum_a TFR_t \times ASFR_{a,t} \times SP_{s,t} \times \frac{(W_{a,t-1} + W_{a,t})}{2} \quad (5)$$

The final variable needed to complete the projection model is *SP*—the proportion of births that are of sex *s.* This can be derived from the sex ratio at birth for men and women respectively. The default behaviour of the tool is to obtain the fertility and mortality data from the WPP, but the values can also be entered by the user.

$$SP_{male,t} = \frac{SRB_t}{1+SRB_t} \quad (6)$$

$$SP_{female,t} = \frac{1}{1+SRB_t} \quad (7)$$

Case study: South Sudanese refugees in Kenya

Kenya is the tenth largest refugee hosting country in the world and the fourth in Africa. The country has hosted a large number of PoC ever since Somali refugees first arrived in 1991 fleeing from armed conflict. They were followed by Sudanese and Ethiopian refugees later that decade. There are currently an estimated 324,400 Somali refugees in Kenya. This is the largest refugee population in the country, followed by refugee populations from South Sudan (87,100), Ethiopia (19,100), and the Democratic Republic of the Congo (13,300)[13]. Relative improvements in the political situation of Somalia and a deterioration in conditions in South Sudan have meant that the proportion of refugees from South Sudan has steadily increased in Kenya over the last few years. Figure 2 shows that Somalia was by far the most common country of origin for new refugees arriving in Kenya up until 2013, when it was overtaken by South Sudan. Nonetheless, the refugee population in Kenya declined by almost 20% from 553,900 to 451,100 in the course of 2016.

The Dadaab Refugee Complex is the larger of the two major refugee complexes in Kenya[14]. At the end of January 2018, it hosted 235,299 refugees and asylum seekers in four refugee camps. The first camps were established in 1991 to host Somali refugees and have now developed into towns with developed commercial hubs. New camps were established consequently to host a growing influx of forcibly displaced individuals. The Kakuma Refugee Camp and the Kalobeyei Integrated Settlement are composed of four and three refugee camps respectively. As of January 2018, these camps hosted 185,449 refugees. Finally, many refugees in Kenya live in urban areas. UNHCR's Urban refugee program is currently in charge of 65,175 asylum-seekers and refugees (as of end January 2018), who mainly reside in Nairobi and other urban locations.

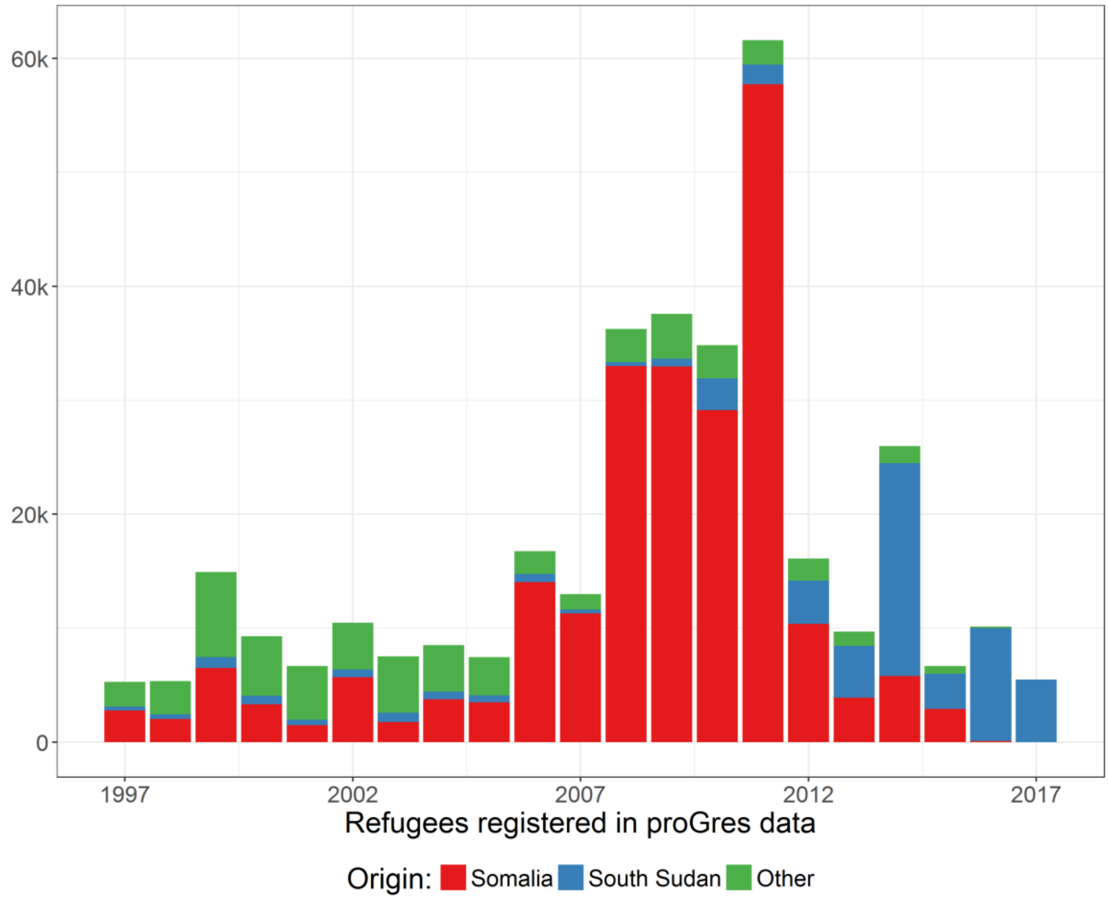
The UNHCR Kenya operation routinely collects data on new arrivals and departures using the latest version of the proGres system, making this an appropriate case study

[13] As of the end of 2016 (UNHCR 2017a).

[14] UNHCR Kenya: http://www.unhcr.org/ke/

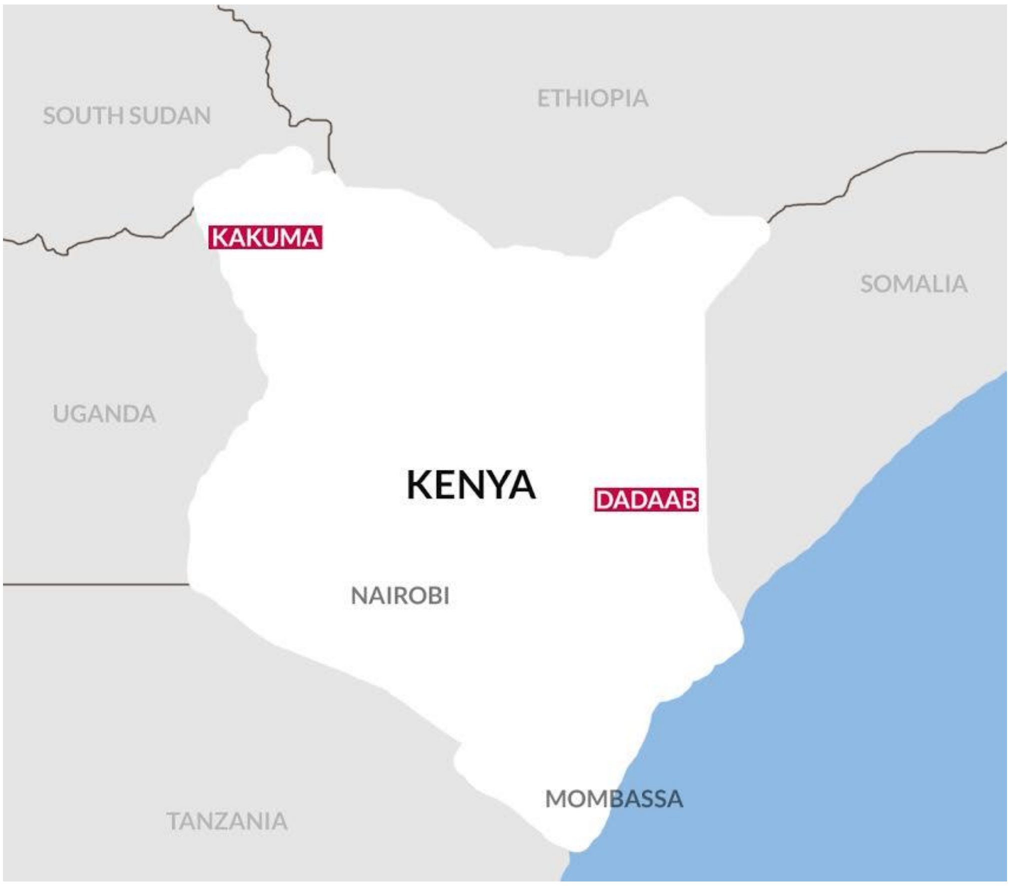
for implementing the DPTool to project populations of PoC in the country. In this section we present an overview of the refugee population in Kenya and we describe how the population of South Sudanese refugees in the country was projected for the 2017-19 period.

Figure 2. New refugee arriving to Kenya by country of origin (flow)



South Sudan gained its independence in 2011 after a prolonged armed conflict. In 2013, a new conflict broke out in Juba, the capital city, and spread out to the Greater Upper Nile. The fighting intensified dramatically during 2016, when violence reached most parts of the country. The civil war has had a devastating effect on South Sudan and thousands have been displaced both internally and to neighbouring countries. During 2016, 752,300 South Sudanese returned to their country. However, by the end of that year the conflict had produced at least 865,000 new displacements. The number of South Sudanese IDPs, refugees, and asylum seekers has increased considerably since. The displaced population has taken refuge in all of the countries bordering South Sudan—Uganda hosts the largest number of South Sudanese refugees (639,000), followed by Ethiopia (338,000), and Sudan (297,200). Kenya hosts the fourth largest number of refugees from South Sudan (87,100).

Figure 3. Major Refugee Complexes in Kenya.



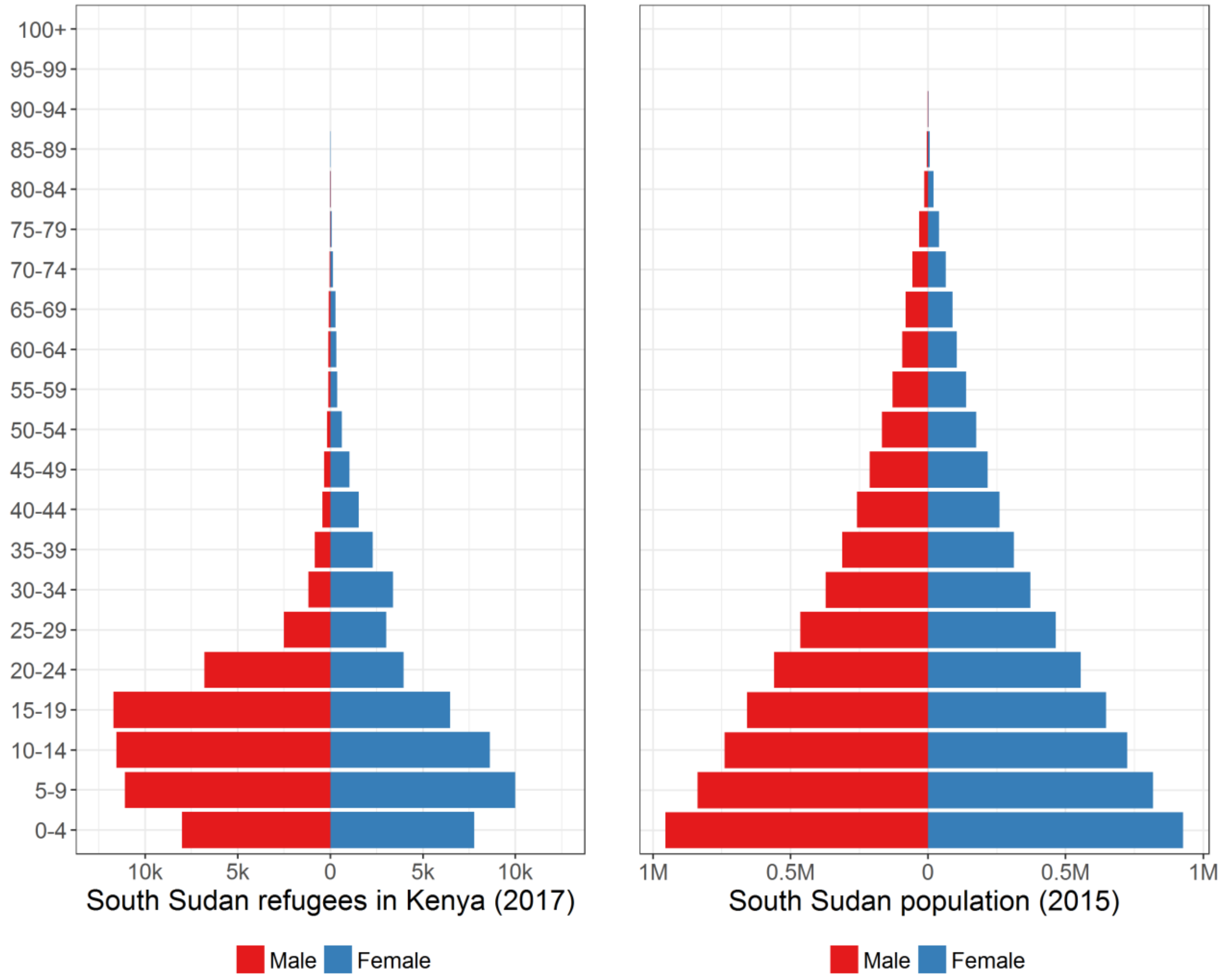
Source: https://tinyurl.com/y79t5a34

Internal analyses of UNHCR registration data have shown that the demographic composition of refugee populations are systematically different from that of non-displaced populations in their countries of origin. Origin-destination comparisons are useful because they highlight that segments of a population are affected differently by conflict. Selection effects are a particular concern. The oldest members of a population may not be physically able to flee during an armed conflict, for example, or young men may be displaced in larger numbers as they try to avoid forced recruitment. These comparisons provide an insight into the selection effects that take place during forced displacement.

Figure 4 compares the population structure of South Sudanese refugees in Kenya (i.e. the population stock) with the structure of the non-displaced population in South

Sudan. A visual analysis of the pyramids shows that the displaced population has a higher share of young males (of which many are unaccompanied minors) and a lower share of children under 5. The latter are possibly affected by problems in registration practices, particularly by the under-registration of births.

Figure 4. Age-sex composition of South Sudanese refugees in Kenya and non-displaced population in country of origin



Source: Author with data from Kenya proGres registration system and WPP statistics (UN DESA 2017).

Focusing on population stocks exclusively hides the fact that populations are highly dynamic. The largest share of the demographic change in refugee stock populations is explained by the international mobility of individuals (in Kenya, the arrivals greatly outnumber the departures, but this depends on the context of the crisis). Since natural increase—the difference between births and deaths—is likely to play a minor role in explaining demographic change in refugee populations, it is necessary to consider the magnitude and composition of the displacement flows. These will ultimately determine the size and shape of the populations of interest.

We now present a working example of how the DPTool can be used to project refugee populations in Kenya using a combination of UNHCR registration data, model life tables, and reference data from WPP. The total size and demographic composition of the 2017 population of South Sudanese refugees in Kenya—the launch population—was derived from the 2017 proGres data provided by the UNHCR Kenya operation. In this exercise, estimates of the size of the future refugee flow over the 2017-19 period were loosely based on the operation's 2017 planning figures. As we mentioned earlier, there is currently no standardised method for estimating future refugee arrivals and departures—assumptions about the population flow are usually derived from contextual information, data on previous arrivals, and professional experience. In this section we project three alternative scenarios predicting a low, medium, and high influx of refugees respectively, with all other parameters unchanged (Table 2). [15] Projecting multiple scenarios is helpful for considering the outcome of different contextual variables, even if only one set of estimates can be included in the final planning figures.

Table 2. Assumptions about the estimated flow of South Sudanese refugees in Kenya

End-year estimate	Low		Medium		High	
	Arrivals	Departures	Arrivals	Departures	Arrivals	Departures
2017	12,000	400	21,000	400	40,000	400
2018	12,000	450	21,000	450	40,000	450
2019	8,000	500	10,000	500	35,000	500

The estimated flow figures were initially broken down by age and sex by applying the age-sex distribution of the 2016 population of arrivals and departures (i.e. the year before the launch year). The tool sources the fertility rates of the country of origin

15 The 'Medium' scenario is closest to the 2017 Kenya planning figures, but the estimates were modified for presentation purposes.

(South Sudan in this case), and the mortality rates of the country of destination (Kenya) from the WPP[16].

An appropriate life table is automatically selected by the tool to model the age-sex patterns of mortality, but these settings can be changed by the user. For this exercise, we selected the Coale and Demeney 'West' regional model life table (UN 1982). Given this input, the DPTool computed the estimates over the projection horizon and displays them graphically and in tabular form[17]. The output includes estimates of (a) total population, (b) population by age and sex, (c) births, and (d) deaths. Figure 5 shows the South Sudanese refugee population reported by the agency for the 2013-2017 period. The plot includes population projections from the three different scenarios outlined above, each with different assumptions about the number of future arrivals[18]. All the figures refer to end-of-year populations.

Figure 5 shows that the number of South Sudanese refugees in Kenya grew on average in the years leading to 2017. This growth, however was not steady—there was a sharp increase in the number of refugees during 2014 and a slight decrease during 2016. This irregular pattern exemplifies the difficulties that we face when forecasting forced migration. An unforeseeable intensification of the conflict may lead to sharp increases in the population, for example, but the official figures can also change because of administrative reasons or due to problems with data quality. There are reasons to believe, for example, that the latter partly explains the reduction of Somali refugees in Kenya during 2016 (UNHCR 2017a).

The DPTool provides data on the future size and composition of the projected population. The 'Medium' scenario in Figure 5, for example, predicted a 30% increase of the South Sudanese refugees (from 105,000 in 2017 to 137,000 in 2019). The projection anticipated a progressive decline in the relative number of children and

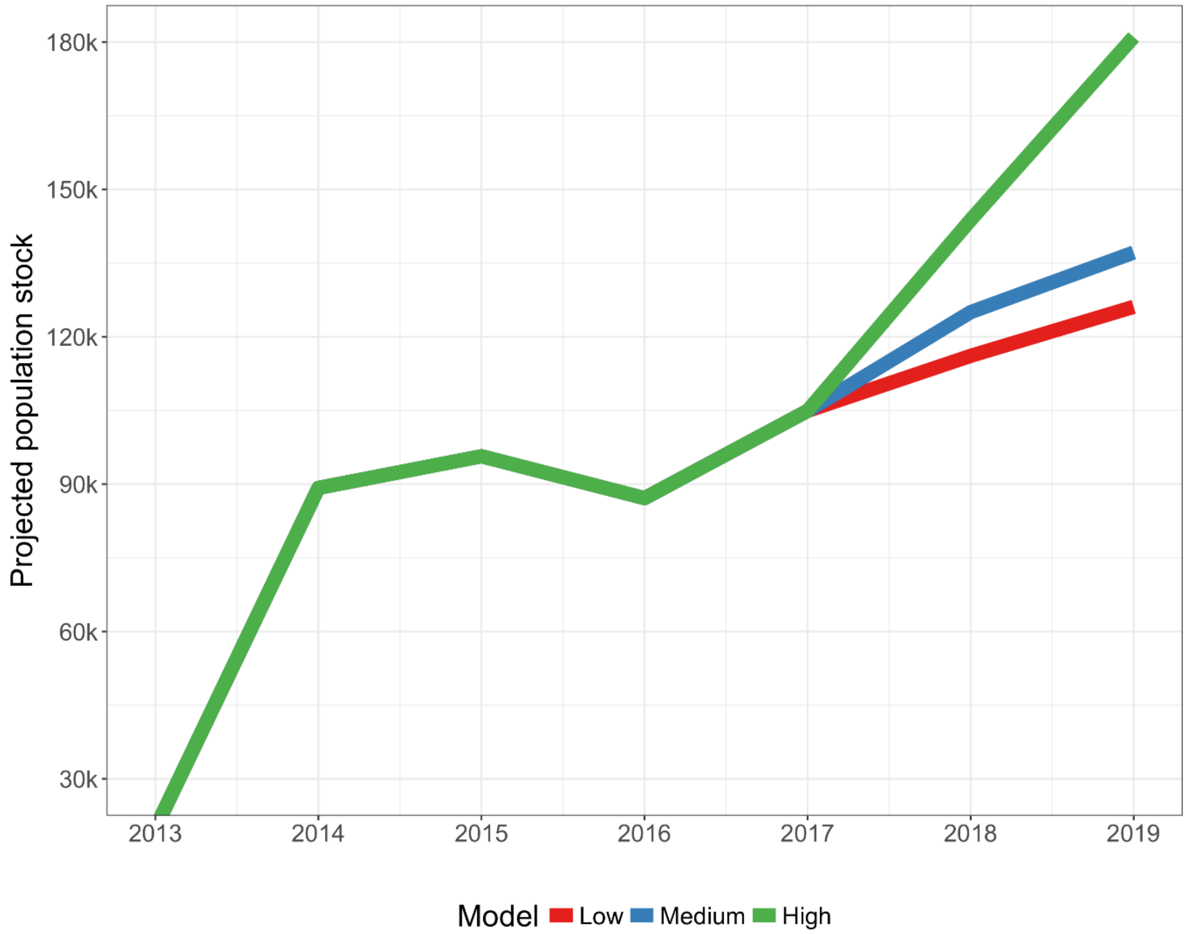
[16] The DPT assumes that change in fertility behaviour is slow, whereas short-term mortality outcomes depend on the availability of healthcare, access to food, sanitation, etc.

[17] At the moment, it is not possible to model different scenarios simultaneously. Because of this, the graphs presented on this paper were made separately by the author and have different formatting to the ones produced by the DPT directly.

[18] Note that different scenarios have to be modelled separately in the DPTool. The figure presented here summarises the output of three separate projections.

young adults, whilst first group adults were expected to constitute a larger share of the population (Table 3). This means that the predicted number of arriving children and young adults was smaller than the population that transitioned to the next age group, died, or otherwise left the population (e.g. due to durable solutions).[19] As mentioned above, these insights are helpful for conducting informed and efficient planning exercises.

Figure 5. Total size of refugee stock with different assumptions about forced migration.



Source: Author with data from three different models produced by the UNHCR Demographic Projection Tool.

It is worth emphasising that the initial assumptions about demographic change do affect the estimates of the DPTool. The size and composition of the flow population are of particular importance when projecting forced migration in the short-term (e.g.

[19] Assuming that the demographic composition of the arrivals is similar to that of the 2016 flow population.

three years into the future). The projection tool assumes that the demographic composition of the flow population (arrivals and departures) is equivalent to that of the previous year's flow population. This assumption, which is reasonable if the nature of the crisis remains unchanged, can be relaxed by the user. Finally, it is important to be realistic about the quality of the proGres data that are used as a starting point for the projection. Are births under-reported? How accurate is death registration amongst PoC? Failing to consider these issues can affect the validity of the estimates. Local expertise continues to be essential for making sense of the projections.

Table 3. Age-sex distribution of projected population (medium scenario)

Projected end-year (%)	Children and young adults (0-19)	First group adults (18-60)	Second group adults (60+)	Total (all age groups)
2017	71.6	27.3	1.1	100
2018	69.1	29.8	1.1	100
2019	67.0	31.9	1.1	100

Conclusions

This paper has shown how a traditional demographic methodology can be applied to promote a systematic approach for modelling the future development of certain populations of persons of concern to UNHCR (PoC). The article provided a brief introduction to UNHCR registration data system (proGres) and outlined some of its limitations. We then discussed demographic projections in general and cohort-component projections in particular when discussing the Demographic Projection Tool (DPTool) developed by UNHCR FICSS. This tool takes as inputs data on population structure, mortality and fertility rates, and estimates of the future number of arrivals and departures into a given territory. It then applies a series of assumptions to produce reliable estimates of the future size and composition of PoC populations. The DPTool provides a flexible and user-friendly platform for projecting PoC groups. The tool automatically updates its assumptions about demographic dynamics depending on the given country of asylum and origin.

We have shown that the DPTool is sensitive to assumptions about the size of future arrivals and departures. Furthermore, the tool is intended for short-term projections only (planning figures usually cover a three-year projection horizon). Long-term predictions are subject to the uncertainty associated with predicting the development of ongoing crises. When will an armed conflict come to an end? How long will it take for the conditions necessary for voluntary repatriation to develop in its aftermath? Promising new research is currently being conducted on estimating future displacement flows, but for the time being projecting multiple scenarios with different assumptions about future forced displacement is a viable alternative (even if only one scenario is included in the official planning figures).

Another limitation of the DPTool is related to its use of proGres data. The tool functions optimally when proGres data are available, although it can also be used without registration data. Furthermore, the proGres registration system faces data quality issues, mostly related to response heaping and under-registration, that need to be addressed. Finally, the deployment of proGres v.4 since April 2015 and the migration of older versions to a data warehouse means that eligible staff outside operations will now have easier access to registration data via secured servers and tailored access rights. In this paper we have presented a tool for conducting evidence-based planning exercises. However, we firmly believe that the experience and expertise of staff and partners must continue to be an integral part of the decision–making process as UNHCR continues to operate in extremely challenging political and operational contexts.

Acknowledgements

Petra Nahmias, Rachelle Cloutier, and Aina Saetre provided helpful comments on earlier versions of this paper. Technical support from Htun Zaw Oo was essential. We would like to thank Yawa Awaho from the UNHCR Kenya operation for valuable support. Avenir Health developed the Demographic Projection Tool and John Stover provided a technical description of the projection model. Opinions in this report are those of the authors and do not reflect the views of the Office of the United Nations High Commissioner for Refugees (UNHCR). Figures and estimates presented in this report should not be taken as official.

Abbreviations

ASFR	Age-specific Fertility Rate
DHS	Demographic and Health Survey
DPSM	Division of Programme Support and Management
DPTool	UNHCR Demographic Projection Tool
FICSS	Field Information and Coordination Support Section (UNHCR)
IDP	Internally Displaced Person
TFR	Total Fertility Rate
PoC	Persons of Concern to UNHCR
proGres	Profile Global Registration System
SRB	Sex Ratio and Birth
WPP	World Population Prospects

References

Hill, K. (2004). *War, Humanitarian Crises, Population Displacement, and Fertility: A Review of Evidence*. Washington, DC: National Academies Press.

Preston, S. H., P. Heuveline, and M. Guillot. 2001. *Demography: Measuring and Modelling Population Processes*. Malden, MA: Blackwell Publishers.

Raftery, A. E., L. Alkema, and P. Gerland. 2014. Bayesian Population Projections for the United Nations, *Statistical Science* 29(1): 58–68.

Siegel, J. S., and D. A. Swanson (Eds.). 2004. *The Methods and Materials of Demography*. London: Elsevier Academic Press.

Smith, S. K. 1986. Accounting for Migration in Cohort-Component Projections of State and Local Populations, *Demography* 23(1): 127–135.

Smith, S. K. (2001). *State and Local Population Projections: Methodology and Analysis*. New York: Kluwer Academic.

Smith, S. K., Tayman, J., & Swanson, D. A. (2002). *State and Local Population Projections: Methodology and Analysis*. Springer Netherlands.

Stover, J. (2007). DemProj Manual: A Computer Program for Making Population Projections. Spectrum System of Policy Models.

UN. (1982). Model Life Tables for Developing Countries. (*United Nations Publication, Sales No. E.81.XIII.7*).

UN. (1955). Methods of Appraisal of Quality of Basic Data for Population Estimates (*United Nations Publications, Sales No. 56.XIII.2*).

UN DESA. 2017. World Population Prospects: The 2017 Revision, Methodology of the United Nations Population Estimates and Projections, *Working Paper No. ESA/P/WP.250*. New York: United Nations.

UNHCR. (2017a). *Global Trends: Forced Displacement in 2016*. Geneva: United Nations High Commissioner for Refugees.

UNHCR. (2017b). *Mid-Year Trends 2017*. Geneva: United Nations High Commissioner for Refugees.

UNRWA. (2017). *UNRWA in Figures*. Jerusalem: United Nations Relief and Works Agency for Palestine Refugees in the Near East.

CONTACT US

Petra Nahmias
Senior Statistician
Division of Programme Support and Management
Field Information and Coordination Support Section
Statistics Unit
+45 45 33 65 91
nahmias@unhcr.org

UNHCR
Marmorvej 51,
2100 København
www.unhcr.org