Statistics South Africa

Preferred supplier of quality statistics

Methodological note

P0302

A methodology for estimating national and provincial populations:

The approach of Statistics South Africa

email: info@statssa.gov.za
www.statssa.gov.za

170 Andries Street • Private Bag X44, 0001 Pretoria, South Africa
Tel: +27(12) 310 8911, Fax: +27(12) 321 7381

Published by:
Statistics South Africa
Private Bag X44
Pretoria
South Africa
0001

Electronic services
For more information about our on-line electronic services, contact user information services at: www.statssa.gov.za

Contact details

Telephone:	(012) 310 8600/ 8390/ 8351/ 4892/ 8496/ 8095 (user information services)
	(012) 310 8490 (library)
Fax:	(012) 310 8500/ 8495 (user information services)
Email:	info@statssa.gov.za (user information services) distribution@statssa.gov.za (orders)

1. Overview

In order to meet the need for population estimates, Statistics South Africa publishes national and provincial population estimates annually. This paper provides users of the mid-year population estimates with a detailed description of the methodology employed by Statistics South Africa. The actual estimates may be downloaded from the Statistics South Africa website.

We distinguish between two phases in our projections. These are:

(a) Projection of the national population by using the SPECTRUM system;

(b) Provincial projection by applying a UN method (United Nations, 1992) in which internal migration plays an important role.

2. National projections

2.1 *Models for projection*

Demographic-based programmes that incorporate an epidemiological component were developed by different organisations. Four of the most important models from a South African viewpoint is (a) the Spectrum Modeling System developed by the Futures Group; (b) the RupHivAids model developed by the US Census Bureau; (c) the abcDIM model which is developed and used by the United Nations Population Division (UNPD) and (d) the ASSA model developed by the Actuarial Society of Southern Africa.

When Statistics South Africa started its new round of population projections in 2003, the documentation of the RupHivAids model was not completed and the abcDIM software was at that stage an in-house programme that was used to do population projection of all the countries in the world. The choice was therefore between ASSA and Spectrum and Stats SA decided to follow the Spectrum route. Better documentation of RupHivAids recently became available and it might be feasible to investigate and compare the assumptions of this model and Spectrum for a later round of South African projections.

The Spectrum Policy Modeling System (Futures Group) consists of 7 components, but Statistics South Africa used only two of them in this projection, namely: (a) **Demproj** for population projections and (b) **AIM** in which the consequences of the AIDS epidemic were projected.

The population projections, which apply the cohort-component method, must first be prepared before AIM can be used. In this projection with base-year 1985, fertility, mortality and internal migration for the projection period are required. Life expectancies and Total fertility rates without taking AIDS into account were used as input in the model.

The base year populations and projected populations on the 1996 and 2001 census dates are given in Table 1 below. It can be seen from this table that according to the Stats SA projections the estimated 1996 population figures were about 1,2 million higher than the published 1996 Census, while a very small difference can be found between the published 2001 Census and this projection on that date. Greater differences were however found regarding the sex-ratios and age-distributions (not shown here).

Table 1: Base and projected South African populations

	Base (1985) population	Reported 1996 Census	Stats SA projection on 1996 Census date	Reported 2001 Census	Stats SA projection on 2001 Census date
African	24 452 430	31 127 631	32 224 274	35 416 166	35 260 840
Coloured	3 001 856	3 600 446	3 674 984	3 994 505	3 958 337
Asian	881 367	1 045 596	1 058 435	1 115 467	1 118 109
White	4 847 619	4 434 697	4 792 336	4 293 640	4 501 676
Total	**33 183 272**	**40 583 573**	**41 750 029**	**44 819 778**	**44 838 962**

2.2 *Adjustment of Antenatal Clinics (ANC) data and HIV Prevalence estimates*

AIM requires an assumption about the future course of adult HIV prevalence. Our knowledge of the HIV epidemic in South Africa is based primarily on the prevalence data collected annually since 1990 from pregnant women attending public antenatal clinics (ANC). These data have been used to obtain national estimates of HIV prevalence among the adult population and to determine epidemic trends over time.

ANC data are biased estimators of the general population HIV prevalence because only a select group of people (pregnant women attending public health services) are included in the sample. HIV prevalence among women who attend public health services is generally estimated to be higher than prevalence among those who attend private health services. Furthermore, the distribution of HIV prevalence among pregnant women is likely to be different than the distribution among the general adult population. In order to correct for these biases we need to adjust the ANC prevalence estimates, firstly (a) by adjusting for

relative attendance rates at antenatal clinics and (b) secondly by adjusting for the difference in prevalence between pregnant women and the general adult population.

Ideally, national HIV prevalence surveys would be an unbiased sample of the general adult population but with the 2005 HSRC population-based survey (Shisana *et al*., 2005) the overall response rate for adults 15–49 years was only 55%. If non-responders are more likely to be HIV positive than responders, then the HSRC survey is likely to underestimate the true population prevalence. The approach used here is therefore to start from the ANC estimates of prevalence, and to use the HSRC data to calculate various correction factors.

This adjustment requires that the ANC data be standardised for race because the race distribution of ANC attendees is different from the race distribution in the general population. Ideally this standardisation should be based on the ANC prevalence of HIV by race and on the race distribution of all pregnant women. However, because the prevalence by race among ANC attendees is not available, we have used the prevalence by race from the HSRC survey, as shown in Table 2. The race distribution for the South African adult population aged 15–49 is based on the closest available population estimates from Statistics South Africa (Statistics South Africa, 2005). The correction factor is then estimated as the ratio of the unadjusted prevalence to the race-standardised prevalence. From Table 2 this ratio is calculated as 16,2 / 18,5 = 0,88.

Table 2: Standardising ANC data for race using adult (15–49 year) prevalence data from the HSRC survey

	Distribution Stats SA midyear 2005	Adult Prevalence (15-49 years)*	Proportional ANC attendance	Standardised estimate of HIV prevalence
African	79,4	19,9	0,92	18,31
Coloured	8,8	3,2	0,05	0,16
White	9,3	0,5	0,01	0,00
Indian	2,5	1,0	0,03	0,03
Total	**100,0**	**16,2**		**18,5**

* Shisana *et al*., 2005

Antenatal clinic surveillance is generally not representative of the entire population (e.g., deep rural areas are often excluded from antenatal clinic surveillance). A further adjustment is therefore necessary and is done by again using data from the HSRC study (2005). In this survey it is shown that prevalence among pregnant women was 23,3%, while in the general

population it was 20,2% among women aged 15–49 years and 16,2% among both men and women aged 15–49 years. The higher prevalence among pregnant women as compared to women in the general population of the same age could be attributed to higher levels of contraceptive use in the general population.

To adjust for the difference between HIV prevalence in pregnant women and prevalence among adults in the general population, a correction factor based on the ratio of prevalence in these two population groups can be calculated using data from the HSRC, i.e. 16,2/23,2= 0,7. The combined correction factor is then estimated as the product of the two correction factors described above, i.e. 0,88×0,7 = 0,61. Using this approach the ANC prevalence estimate therefore has to be adjusted by a factor 0,61.

The HIV prevalence data from national ANC surveillance in South Africa from 1990 to 2005 were entered into the Estimation and Projection Package (EPP) to obtain a national epidemic curve. Prevalence data were adjusted using the correction factor as described above. The adjusted national prevalence curve was then entered into the Spectrum software package to estimate the demographic impact of HIV.

3. Provincial projections

3.1 *Overview*

When projections for all the regions of a country are desired and the appropriate data are available, a multi-regional approach should be considered, as this is the only way to guarantee that the total migration flows between regions will sum to zero, or to the assumed level of international migration (United Nations, 1992). Developed by Willekens and Rogers (1978), these methods have not been widely used in developing countries, largely due to the lack of adequate migration data and the difficulty of applying these methods.

Multi-regional methods require the estimation of separate age-specific migration rates between every region of the country and every other region, and such detailed data are rarely available. Although it is possible to estimate some of the missing data (see Willekens *et al*., 1979), the task of preparing data can become overwhelming if there are many regions. If there are only a few streams, however, the multi-regional method is the best method to use. In South Africa, 2448 (9x8x17x2) migration streams are derived if the multi-regional model is applied in calculating migration streams by age group (17 in total) and sex for each

province. The number of streams increase dramatically if this approach is followed for the 52 district councils/metros, namely to about 90 000.

3.2 The age-structure of the provinces

The base from which a population projection is done is very important as it has a big effect on the outcome of a projection. It also forms the base from which the provincial fertility and mortality rates were adjusted in this study. For a number of reasons that we will discuss now, it was decided to make adjustments to the total 2001 Census provincial populations as well as to the age and sex structures.

We will discuss two aspects, namely the way that provincial adjustment factors were constructed for Census 2001 and the average provincial growth rate between Census 1996 and Census 2001. We will also point out some inconsistencies from these aspects, which have to be rectified before the data could be used for projections.

To understand why adjustments for the base provincial populations were made, one will have to consider how adjustment factors were derived for Census 2001:

- An assessment of the magnitude and direction of errors were made by conducting a Post-Enumeration Survey (PES) one month after the Census.
- By using the PES, the unadjusted 37,3 million population was adjusted to 44,82 million (17,6 % net undercount).

The net undercount of the provinces was not the same and this fact will be taken into account. Table 3 shows the net undercount, absolute errors, and their confidence intervals by province.

Table 3: Census 2001: Net undercount rate by province

	Net undercount	Absolute error	Confidence intervals	
			Lower	Upper
Eastern Cape	14,7	1,8	12,9	16,6
Free State	17,6	1,2	16,4	18,9
Gauteng	18,7	3,5	15,3	22,2
KwaZulu-Natal	22,5	5,6	16,9	28,1
Limpopo	14,4	0,4	14,0	14,7
Mpumalanga	16,1	1,0	15,1	17,0
Northern Cape	14,1	0,8	13,2	14,9
North West	16,0	1,2	14,8	17,3
Western Cape	16,3	1,5	14,8	17,7

If we further examine the provincial growth rates between Census 1996 and Census 2001 (see Table 4) the following is revealed:

- The very low growth rates of the Eastern Cape and the Free State
- The negative growth of the Northern Cape
- The high growth rates of Gauteng, KwaZulu-Natal and the Western Cape

Table 4: Comparing growth rates between Census 1996 and Census 2001

	Census results (October)			**Stats SA mid-year adjustments**	
	1996	**2001**	**Average annual growth rate**	**2001**	**Average annual growth rate**
EC	6 302 525	6 436 763	0,42	6 929 869	2,01
FS	2 633 504	2 706 775	0,55	2 893 541	1,99
GT	7 348 423	8 837 178	3,69	8 254 103	2,46
KZN	8 417 021	9 426 017	2,55	9 263 134	2,03
LIM	4 929 368	5 273 642	1,35	5 474 683	2,22
MP	2 800 711	3 122 990	2,18	3 103 451	2,18
NC	840 321	822 727	-0,42	870 657	0,75
NW	3 354 825	3 669 349	1,79	3 686 162	2,00
WC	3 956 875	4 524 335	2,68	4 207 044	1,30

If we take into account the high level of uncertainty of the census adjustments (the wide confidence intervals of especially Guateng and KwaZulu-Natal in Table 3) and growth rates that seemed to be questionable, Stats SA decided to make use of information of the Bureau of Market Research in which the total provincial populations for 2001 were adjusted (Van Aardt, 2006). The final age/sex structures of the provinces were determined through an iteration process and using the following datasets:

- The projected RSA population by sex and five-year age groups;
- The provincial Census 2001 populations by age and sex;
- Dataset of the BMR in which the Census 2001 provincial and magisterial totals were adjusted.

3.3 *Calculation of migration rates from the 2001 Census*

For subnational areas, migration is often the major determinant of population growth and can also be seen as the most difficult component of growth to forecast accurately as they are subject to much greater volatility than either fertility or mortality rates.

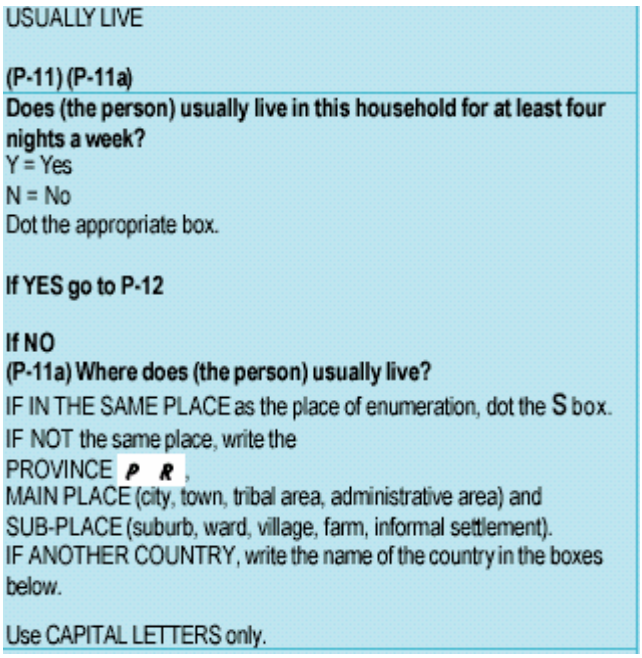
To determine flow patterns a census or large survey is usually the only source of data and two procedures most frequently used for measuring internal migration are:

(a) Procedures based on questions in censuses or surveys intended to detect migration.

(b) Procedures based on other population characteristics such as age-structure and place of birth.

As special questions in Census 2001 were available, it was decided to follow the approach in (a) above in the Stats SA study. The following questions were used to determine usual residence, residence 5 years ago, and previous residence.

Questions to determine usual residence


USUALLY LIVE

(P-11) (P-11a)

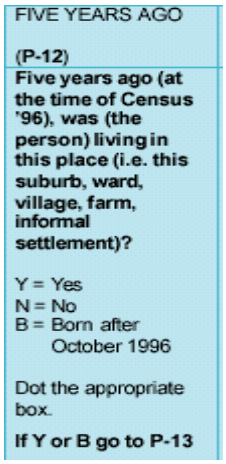
Does (the person) usually live in this household for at least four nights a week?
Y = Yes
N = No
Dot the appropriate box.

If YES go to P-12

If NO
(P-11a) Where does (the person) usually live?
IF IN THE SAME PLACE as the place of enumeration, dot the S box.
IF NOT the same place, write the
PROVINCE P R ,
MAIN PLACE (city, town, tribal area, administrative area) and
SUB-PLACE (suburb, ward, village, farm, informal settlement).
IF ANOTHER COUNTRY, write the name of the country in the boxes below.

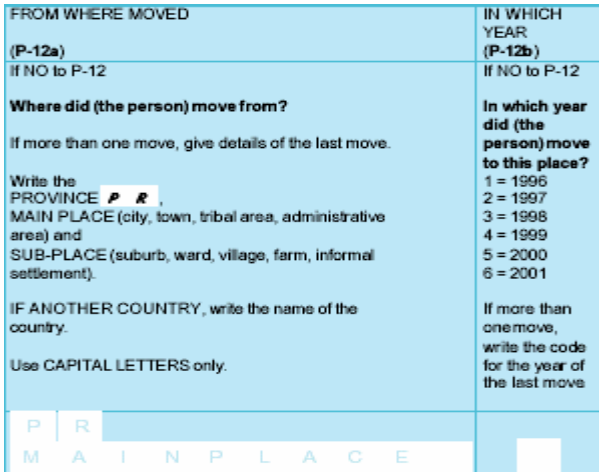
Use CAPITAL LETTERS only.

Question to determine if person lived in this place five years ago



FIVE YEARS AGO (P-12)
Five years ago (at the time of Census '96), was (the person) living in this place (i.e. this suburb, ward, village, farm, informal settlement)? Y = Yes N = No B = Born after October 1996 Dot the appropriate box. **If Y or B go to P-13**

Diagram 3: Question to determine previous residence



FROM WHERE MOVED (P-12a)	IN WHICH YEAR (P-12b)
If NO to P-12 **Where did (the person) move from?** If more than one move, give details of the last move. Write the PROVINCE P R, MAIN PLACE (city, town, tribal area, administrative area) and SUB-PLACE (suburb, ward, village, farm, informal settlement). IF ANOTHER COUNTRY, write the name of the country. Use CAPITAL LETTERS only. P R M A I N P L A C E	If NO to P-12 **In which year did (the person) move to this place?** 1 = 1996 2 = 1997 3 = 1998 4 = 1999 5 = 2000 6 = 2001 If more than one move, write the code for the year of the last move

To convert the response to these migration questions to migration rates per thousand of the population in specific age groups, the following steps were followed:

(a) Determine the usual place of residence in 2001 of each person.

(b) Determine the previous place of residence (in this analysis the residence in 1996).

(c) If the usual place of residence in 2001 and the previous residence in 1996 were the same, then the person is classified as a non-migrant.

(d) If the usual place and previous place are different the direction of the migration stream is determined.

(e) Cross-tabulate usual and previous province (see Table 7 below).

(f) Cross-tabulate age (five-year groups) with usual province for every previous province (see Table 8).

By using the unit record count of Census 2001, we were quite successful in determining the previous and usual residence of every person in this census. In Tables 5 and 6 below the result of the migration calculations are given.

Table 5: Result of migration calculations regarding usual province

	Numbers	Percentage
Province determined	44 752 441	99,85
No response	53 705	0,12
Place outside RSA	10 910	0,02
Invalid main place	2 722	0,01
Total	**44 819 778**	**100,0**

Table 6: Result of migration calculations regarding previous province

	Numbers	Percentage
Province determined	42 392 170	94,58
No response	143 595	0,32
Place outside RSA	181 408	0,40
Invalid main place	24 430	0,05
Children born in 5 years before the census and who are not children of the head of the household	2 078 175	4,64
Total	**44 819 778**	**100,0**

We are now ready to convert the migration information in Census 2001 to migration rates per thousand of the population. This is done by using two sets of information, namely the cross-tabulations of previous and usual provinces for the total population and by age groups.

Table 7 is used to determine the level of migration movements between the provinces. To illustrate the procedure only the female movements are given in this table. Table 7 shows that 94,12 % of the population of the Eeastern Cape reported that they had been living in the same province in 1996 and 2001. This, however, does not mean that they did not move as the census question did not focus on the last move. Of those that moved, 2,4 % moved to Western Cape, 1,5 % to Gauteng and smaller proportions to other provinces.

Table 7: Total female movement between provinces: 1996-2001 (percentages)

Prov	Province in 2001								
In 1996	EC	FS	GT	KZN	LIM	MP	NC	NW	WC
EC	94,12	0,2748335*	1,50	1,00	0,11	0,16	0,06	0,40	2,38
FS	0,32	95,04	2,53	0,33	0,18	0,25	0,25	0,77	0,50
GT	0,34	0,30	96,57	0,54	0,49	0,41	0,08	0,55	0,72
KZN	0,22	0,10	1,47	97,55	0,08	0,21	0,02	0,09	0,27
LIM	0,06	0,08	3,48	0,10	95,21	0,62	0,03	0,31	0,10
MP	0,12	0,19	2,88	0,37	0,63	95,32	0,05	0,26	0,20
NC	0,33	0,87	1,39	0,22	0,20	0,15	93,36	1,01	2,46
NW	0,18	0,36	3,53	0,16	0,41	0,19	0,44	94,49	0,26
WC	0,69	0,13	0,84	0,24	0,07	0,08	0,27	0,09	97,58

* This FS cell has more decimals in order to illustrate the calculations in Table 9.

The next step is now to determine the age distribution of the out-migrants from each province to every other province. The results of these analyses are given in Table 8 to illustrate the out-migration of the female Eastern Cape population to the other eight provinces. To complete the analyses eight similar tables (not shown here) have been constructed.

Table 8 does not show the volume of migration, but only indicates the age distribution of the migrants. It is clear from this table that in the age group 15–24 years the highest percentage of migrants was found. For example in Gauteng about 45 % of the out-migrants were from the age group 15–24 years.

It is therefore clear that the information in both Tables 7 and 8 must be used to calculate out-migration rates per thousand of the population. An example of such a calculation is given in Table 9. In this calculation it is assumed that the rates for the period 1996–2001 can be used in the period 2001–2006.

Table 8: Female migrants from EC to other provinces (percentage distribution)

Age group	Province in 2001							
	FS*	GT	KZN	LIM	MP	NC	NW	WC
Births after 2001	0,060519991	0,0557	0,0626	0,0747	0,0820	0,0706	0,0810	0,0567
0-4	0,078544898	0,0524	0,0635	0,0836	0,0676	0,0879	0,0650	0,0638
5-9	0,075486126	0,0483	0,0628	0,0494	0,0646	0,0855	0,0406	0,0652
10-14	0,120166048	0,1073	0,1291	0,0922	0,1145	0,1257	0,0814	0,1325
15-19	0,187786760	0,2446	0,2402	0,1598	0,1768	0,1299	0,1979	0,2448
20-24	0,151846187	0,2126	0,1682	0,1638	0,1745	0,1327	0,2059	0,1823
25-29	0,107712475	0,0982	0,0881	0,1355	0,1168	0,0850	0,1320	0,0867
30-34	0,071444177	0,0638	0,0612	0,0857	0,0768	0,0729	0,0882	0,0543
35-39	0,049705047	0,0436	0,0428	0,0593	0,0462	0,0579	0,0531	0,0381
40-44	0,029495303	0,0266	0,0261	0,0379	0,0233	0,0519	0,0252	0,0241
45-49	0,019008084	0,0151	0,0176	0,0198	0,0157	0,0276	0,0102	0,0154
50-54	0,011361154	0,0095	0,0112	0,0097	0,0094	0,0257	0,0066	0,0110
55-59	0,011033428	0,0075	0,0095	0,0097	0,0105	0,0140	0,0047	0,0091
60-64	0,008302381	0,0051	0,0057	0,0044	0,0073	0,0084	0,0025	0,0060
65-69	0,006663754	0,0033	0,0055	0,0052	0,0050	0,0098	0,0022	0,0043
70-74	0,005571335	0,0031	0,0025	0,0044	0,0046	0,0065	0,0018	0,0025
75-79	0,003604981	0,0019	0,0020	0,0026	0,0029	0,0042	0,0008	0,0018
80+	0,001747870	0,0013	0,0013	0,0024	0,0017	0,0037	0,0010	0,0013
Total	1	1	1	1	1	1	1	1

* The FS column has more decimals in order to illustrate the calculations in Table 9.

Table 9: Calculation of 'scaled' migration rates (out-migration of EC Females to the FS in the period 2001 to 2006)

Age in 2001	EC Female Population in 2001	Out migration rates from the EC to the FS (Table 8)	Estimated migrants	Scaled migration rates
Births after 2001		0,060519991		0,001936
0-4	387 164	0,078544898	30 410	0,002513
5-9	446 561	0,075486126	33 709	0,002415
10-14	450 993	0,120166048	54 194	0,003844
15-19	403 541	0,187786760	75 780	0,006007
20-24	305 412	0,151846187	46 376	0,004858
25-29	253 482	0,107712475	27 303	0,003446
30-34	199 671	0,071444177	14 265	0,002285
35-39	182 789	0,049705047	9 086	0,001590
40-44	182 212	0,029495303	5 374	0,000944
45-49	153 021	0,019008084	2 909	0,000608
50-54	127 907	0,011361154	1 453	0,000363
55-59	108 174	0,011033428	1 194	0,000353
60-64	119 473	0,008302381	992	0,000266
65-69	91 015	0,006663754	607	0,000213
70-74	61 401	0,005571335	342	0,000178
75-79	35 740	0,003604981	129	0,000115
80+	31 958	0,001747870	56	0,000056
Total	3 540 514	1	304 177	
Calculated total out-migration rate			304 177 / 3 540 514 = 0,0859132	
Desired total out-migration rate (from Table 7)			0,002748335	
Scale factor			0,002748335 / 0,0859132 = 0,03198967	

The scaled migration rates in the last column of Table 9, was calculated by applying the scale factor to the census migration rates (third column). These out-migration rates for the Eastern Cape females to the Free State will be used in the cohort-component projections. For each of the nine provinces, sixteen tables of the format of Table 9 will be created.

3.4 *Provincial Fertility Rates*

The following steps were used to obtain a set of age-specific fertility rates for each province to be used in the provincial cohort-component projections:

(a) For each province a preliminary set of age-specific fertility rates were obtain. In this analysis the rates were taken from a report by Moultrie and Dorington (2004).

(b) The preliminary age-specific fertility rates were then applied to the female population 15 to 49 years to calculate the number of births. This was done for each province.

(c) The total number of births generated from the provinces were then compared with the total number of births in the RSA projection (phase one). Proportional adjustments were made if necessary and adjusted TFRs calculated.

(d) Using these adjusted TFRs and age specific fertility rates as well as survival ratios, the number of birth and the 0–4 projected population were obtained. The projected 0–4 year and 5–9 year populations were checked for consistency. Provision was made to adjust the TFR manually if inconsistencies were found.

(e) The process ((a) to (d) above) was repeated if inconsistencies were found in (d).

3.5 *Provincial survival ratios*

The following steps were used to obtain a set of survival ratios for each province to be used in the provincial cohort-component projections:

(a) For each province a preliminary set of life expectancies at birth for males and females were obtained. In this analysis these rates were taken from ASSA 2000 and 2003.

(b) Life Tables for males and females separately and for each province were then constructed from these preliminary set of life expectancies at birth (Far-East pattern of the UN used).

(c) By using the m(x)-values (age-specific mortality rates) in these life tables, the number of deaths by age could then be calculated for both males and females in each province.

(d) The numbers of male and female deaths calculated for each province were then compared with the total number of male and female deaths in the RSA projection respectively. Proportional adjustments were made if necessary.

(e) The adjusted m(x) values were then used to construct life tables. Life expectancies at birth as well as survival ratios by age can be read from the obtained life tables.

An example of the calculations for the female population follows the steps as described above and is given in Table 10. The first step ((a) above) is to make assumptions about a preliminary set of life expectancies. In Table 10 below these values are indicated as shaded areas in column 4. The second step [(b) above] is to use these life expectancies to create life tables. This was done by using the MATCH application in Mortpak4 and the m(x)-values from the life expectancies in column 4 of Table 10. Applying the derived m(x)-values to the

population figures will result in the number of deaths per age group. In Table 10 these calculations were done for every province and according to step ((d)) above the total number of deaths for all the provinces are compared with RSA number of deaths. In this example the total number of RSA deaths was 294 666 and the deaths by province total 333 314. The adjusted deaths (for each age group in a province) are calculated by applying an adjustment factor of 0,884049274 (294 666 divided by 333 314). The adjusted deaths are given in column 6 in the table below. The last step is to use the adjusted m(x)-values (shaded values in column 7) to construct life tables. This is done by applying the LTMXQXAD spreadsheet in PAS. The survival ratios read from the constructed life tables and which will be used in the projection, are given in the last column of Table 10.

Table 10: Calculation of female survival ratios for provincial projections

	Age category	Population in 2001	Derived m(x) values	Deaths	Adjusted		Survival ratios	
					Deaths	m(x)	Categories	Values
EC	0	69 939	0,084125316	5884	5 202	0,0744	Birth	0,9177
	1-4	317 225	0,009896939	3140	2 776	0,0088	0- 4	0,9745
	.	.	.	.	.	.	.	.
	80-84	21 446	0,154634765	3316	2 932	0,1367	75-79	0,5717
	85+	10 512	0,236319752	2484	2 196	0,2089	80+	0,3913
	EC Total		LE=53,6	52111	46 069	LE=56,3		
FS	0	31 344	0,094427415	2960	2 617	0,0835	Birth	0,9068
	1-4	122 286	0,012152275	1486	1 314	0,0107	0- 4	0,9688
	.	.	.	.	.	.	.	.
	80-84	7 209	0,161594571	1165	1 030	0,1429	75-79	0,5549
	85+	4 782	0,242804251	1161	1 026	0,2146	80+	0,3794
	FS Total		LE=50,7	23263	20 565	LE=53,6		
GT	0	85 909	0,076565555	6578	5 815	0,0677	Birth	0,9256
	1-4	318 539	0,008381398	2670	2 360	0,0074	0- 4	0,9783
	.	.	.	.	.	.	.	.
	80-84	15 073	0,149157647	2248	1 987	0,1319	75-79	0,5851
	85+	9 722	0,231211611	2248	1 987	0,2044	80+	0,4008
	GT Total		LE=55,8	51045	45 127	LE=58,4		
KZN	0	110 491	0,114533793	12655	11 188	0,1013	Birth	0,8851
	1-4	464 769	0,017124547	7959	7 036	0,0151	0- 4	0,9563
	.	.	.	.	.	.	.	.
	80-84	20 673	0,173734742	3592	3 176	0,1536	75-79	0,5260
	85+	11 800	0,254100355	2998	2 650	0,2246	80+	0,3591
	KZN Total		LE=45,4	93169	82366	LE=48,4		

Table 10: Calculation of female survival ratios for provincial projections (continued)

	Age category	Population in 2001	Derived m(x) values	Deaths	Adjusted		Survival ratios	
					Deaths	m(x)	Categories	Values
LIM	0	61 848	0,085874091	5311	4 695	0,0759	Birth	0,9159
	1-4	273 599	0,010264272	2808	2 482	0,0091	0- 4	0,9735
	.	.	.	.	.	.	.	.
	80-84	17 848	0,155854418	2782	2 459	0,1378	75-79	0,5688
	85+	11 932	0,237456566	2833	2 505	0,2099	80+	0,3892
	LIM Total		LE=53,1	40481	35 787	LE=55,9		
MP	0	43 463	0,099171047	4310	3 810	0,0877	Birth	0,9017
	1-4	171 468	0,013264737	2274	2 010	0,0117	0- 4	0,9660
	.	.	.	.	.	.	.	.
	80-84	7 503	0,164633858	1235	1 092	0,1455	75-79	0,5476
	85+	4 384	0,245634068	1077	952	0,2172	80+	0,3742
	MP Total		LE=49,4	26865	23 750	LE=52,3		
NC	0	12 576	0,066936442	842	744	0,0592	Birth	0,9355
	1-4	48 485	0,006619639	321	284	0,0059	0- 4	0,9828
	.	.	.	.	.	.	.	.
	80-84	2 427	0,141625996	344	304	0,1253	75-79	0,6038
	85+	1 820	0,224177177	408	361	0,1982	80+	0,4140
	NC Total		LE=58,7	6201	5 482	LE=61,2		
NW	0	36 840	0,087281737	3215	2 842	0,0772	Birth	0,9144
	1-4	140 356	0,010564783	1483	1 311	0,0093	0- 4	0,9728
	.	.	.	.	.	.	.	.
	80-84	7 284	0,156823615	1142	1 010	0,1386	75-79	0,5664
	85+	5 319	0,238359776	1268	1 121	0,2107	80+	0,3875
	NW Total		LE=52,7	22997	20331	LE=55,5		
WC	0	43 634	0,046253351	2018	1 784	0,0409	Birth	0,9564
	1-4	168 864	0,003477976	587	519	0,0031	0- 4	0,9909
	.	.	.	.	.	.	.	.
	80-84	8 708	0,122492877	1067	943	0,1083	75-79	0,6526
	85+	6 439	0,206225574	1328	1 174	0,1823	80+	0,4489
	WC Total		LE=65,4	17183	15190	LE=67,6		
Total number of deaths (all provinces) = 333314								
RSA	0	496 043	0,077923398	38 653				
	1-4	2 025 591	0,008644808	17 511				
	.	.	.	.				
	80-84	108 170	0,150167930	16 244				
	85+	66 710	0,232154254	15 487				
			RSA LE=55,4	RSA Total deaths = 294666				

3.6 Calculating provincial population estimates for South Africa

The format that explains the cohort-component method used to project the provincial populations is shown in Table 11 below. The projection is for the Eastern Cape female population. Please note that the same population was used in the discussion of mortality (Table 10) and migration (Table 9).

Table 11: Projection of the Eastern Cape female population (part 1)

Age	Population 2001	Survival ratio (Table 10)	Age specific fertility	Migration rates (per thousands of population) to:							
				FS (Table 9)	GT*	KZN	LIM	MP	NC	NW	WC
Births after 2001	426 668	0,9177		0,00194	0,00952	0,00567	0,00106	0,00154	0,00055	0,00393	0,01649
0-4	387 164	0,9745		0,00251	0,00895	0,00575	0,00118	0,00127	0,00069	0,00316	0,01854
5-9	446 561	0,9908		0,00241	0,00826	0,00570	0,00070	0,00121	0,00067	0,00197	0,01895
10-14	450 993	0,9878		0,00384	0,01833	0,01170	0,00131	0,00215	0,00099	0,00396	0,03849
15-19	403 541	0,9802	0,0609	0,00601	0,04179	0,02177	0,00226	0,00332	0,00102	0,00962	0,07114
20-24	305 412	0,9742	0,1380	0,00486	0,03634	0,01525	0,00232	0,00328	0,00104	0,01000	0,05298
25-29	253 482	0,9696	0,1644	0,00345	0,01678	0,00799	0,00192	0,00219	0,00067	0,00641	0,02520
30-34	199 671	0,9645	0,1431	0,00229	0,01090	0,00555	0,00121	0,00144	0,00057	0,00428	0,01579
35-39	182 789	0,9579	0,0994	0,00159	0,00745	0,00388	0,00084	0,00087	0,00045	0,00258	0,01107
40-44	182 212	0,9476	0,0436	0,00094	0,00455	0,00237	0,00054	0,00044	0,00041	0,00123	0,00701
45-49	153 021	0,9305	0,0162	0,00061	0,00258	0,00160	0,00028	0,00029	0,00022	0,00049	0,00449
50-54	127 907	0,9049		0,00036	0,00162	0,00101	0,00014	0,00018	0,00020	0,00032	0,00318
55-59	108 174	0,8693		0,00035	0,00128	0,00087	0,00014	0,00020	0,00011	0,00023	0,00264
60-64	119 473	0,8217		0,00027	0,00087	0,00052	0,00006	0,00014	0,00007	0,00012	0,00175
65-69	91 015	0,7580		0,00021	0,00057	0,00050	0,00007	0,00009	0,00008	0,00011	0,00125
70-74	61 401	0,6754		0,00018	0,00054	0,00023	0,00006	0,00009	0,00005	0,00009	0,00073
75-79	35 740	0,5717		0,00012	0,00033	0,00018	0,00004	0,00005	0,00003	0,00004	0,00052
80+	31 958	0,3913		0,00006	0,00022	0,00011	0,00003	0,00003	0,00003	0,00005	0,00036

* The calculations of the migration rates to the other provinces as well as the age specific fertility rates are not given in this document

The main steps in deriving provincial mid-year population estimates for South Africa are as follows.

3.6.1 *Calculate the number of out-migrants (5 years and older)*

Whereas a projection for a single region involves multiplying the population at the first time-point in each five-year age group by a survival rate to obtain the survivors to the next five-year age group at the second time point, a multi-regional projection involves a compound

survival rate which specifies the probability of surviving and being in a particular region at the second time-point. A compound survival rate is the product of the survival rate and the out-migration rate(s) to each of the other provinces. The number of out-migrants from province A to each of the other provinces (B to I) is then defined as:

$$OUT^{AB}_{t+5,x+5} = P^{A}_{t,x} * S^{A}_{t,x} * MR^{AB}_{t,x}$$
$$OUT^{AC}_{t+5,x+5} = P^{A}_{t,x} * S^{A}_{t,x} * MR^{AC}_{t,x}$$
$$\vdots$$
$$OUT^{AI}_{t+5,x+5} = P^{A}_{t,x} * S^{A}_{t,x} * MR^{AI}_{t,x}$$

Where:

$S^{A}_{t,x}$ is the survival ratio of province A, age group x, first projection period; $MR^{AB}_{t,x}$ is the migration rate of province A to province B, age group x, first projection period; $MR^{AC}_{t,x}$ is the migration rate of province A to province C, age group x, first projection period; and $MR^{AI}_{t,x}$ is the migration rate of province A to province I, age group x, first projection period. The migration rate is defined as the number of migrants per thousand of the population in a specific age group.

3.6.2 *Calculate the number of survivors by province (5 years and older)*

For survival in the same province, the compound rate is the survival rate times one minus the sum of the out-migration to the other provinces. That is, the survivors (those that have not died or migrated) for people in age group x+5 and period t+5 of province A are obtained by the following formula:

$$SUR^{A}_{t+5,x+5} = P^{A}_{t,x} * S^{A}_{t,x} * (1 - MR^{AB}_{t,x} - MR^{AC}_{t,x} - MR^{AD}_{t,x} - .....MR^{AI}_{t,x})$$

Where:

$P^{A}_{t,x}$ is the population of province A, age group x, first time period; and the other symbols are defined as before. The number of survivors in each of the other provinces is calculated in the same way.

Applying the formulas in sections 3.6.1 and 3.6.2 and using the data in Table 11 will result in the number of out-migrants as set out in Table 12. The calculations in Tables 11 and 12 will

have to be repeated for all the other female population in the other provinces (not shown). The same format, except for the fertility assumptions, is used for the male populations.

Table 12: Projection of the Eastern Cape female population (part 2)

	Survivors in EC	Number of out migrants to:								In-migrants To EC*	Projected Population 2006
		FS	GT	KZN	LIM	MP	NC	NW	WC		
0-4	375 618	758	3 729	2 221	414	603	217	1 540	6 457	6 351	381 969
5-9	361 426	947	3 377	2 169	445	479	260	1 192	6 995	6 358	367 784
10-14	424 812	1 066	3 655	2 522	310	535	296	872	8 384	5 772	430 584
15-19	409 509	1 711	8 166	5 212	584	958	441	1 764	17 147	6 874	416 383
20-24	333 477	2 377	16 530	8 611	894	1 313	403	3 805	28 139	7 487	340 964
25-29	260 022	1 446	10 812	4 537	690	976	309	2 975	15 763	7 997	268 019
30-34	229 897	848	4 122	1 964	472	538	165	1 575	6 194	6 650	236 547
35-39	184 488	441	2 099	1 069	233	277	110	824	3 041	4 493	188 981
40-44	170 063	278	1 304	679	147	152	79	452	1 938	2 912	172 975
45-49	169 644	162	786	409	93	76	71	212	1 210	1 907	171 551
50-54	140 882	87	367	228	40	41	31	70	639	1 312	142 194
55-59	114 932	42	188	117	16	21	23	37	368	884	115 816
60-64	93 488	33	120	82	13	19	10	22	248	636	94 124
65-69	97 798	27	85	51	6	14	7	12	172	327	98 125
70-74	68 791	14	39	34	5	6	6	8	86	134	68 925
75-79	41 389	7	22	10	2	4	2	4	30	57	41 446
80+	32 900	3	10	5	1	1	1	2	16	27	32 927
Total	3 509 136	10 247	55 411	29 920	4 365	6 013	2 431	15 366	96 827		3 569 314

* To obtain the in-migrants to the EC, similar calculations were done for all the other provinces.

3.6.3 *Calculate the number of in-migrants (5 years and older)*

The number of in-migrants to province A (see second last column in Table 12) is obtained by adding the out-migrants from the other provinces (B to I) to province A, that is:

$$IN^{A}_{t+5,x+5} = OUT^{BA}_{t+5,x+5} + OUT^{CA}_{t+5,x+5} + OUT^{DA}_{t+5,x+5} + \ldots\ldots\ldots + OUT^{IA}_{t+5,x+5}$$

3.6.4 *Projected population (5 years and older)*

The projected provincial population of A in each age group aged 5 years and over (see last column in Table A12) is simply the sum of the survivors in province A and the number of in-migrants to province A, namely:

$$P^{A}_{t+5,x+5} = SUR^{A}_{t+5,x+5} + IN^{A}_{t+5,x+5}$$

3.6.5 *Calculate the number of births and survivors aged 0–4 years*

Annual births are estimated by applying the age-specific birth rates assumed for each province to the number of women in each of the reproductive age groups. This step is done separately for 1996 and 2001. The results are averaged and then multiplied by five to obtain the total number of births in the specific province for the first five-year projection interval. The total number of births is multiplied by the assumed sex ratio at birth to obtain the number of male births. The number of female births is calculated by subtracting the number of male births from the total number of births. The projected 0–4 population (see first entry in the last column of Table 12) is calculated by applying the formula in sections 3.6.1 to 3.6.4.

This projection process can be repeated for further time intervals and the assumed levels of mortality, fertility and migration can be altered for each projection period, if desired.

REFERENCES

Moultrie T & Dorrington R. 2004. *Estimation of fertility from the 2001 South African Census data*. Centre for Actuarial Research, University of Cape Town.

Shisana O, Rehle T, Simbayi LC, Parker W, Zuma K, Bhana A, Connolly C, Jooste S & Pillay V. 2005. *South African National HIV Prevalence, HIV Incidence, Behaviour and Communication Survey.* HSRC Press, Cape Town.

Statistics South Africa, 2005. *Mid-year population estimates, South Africa.* Statistical Release P0302. Statistics South Africa, Pretoria.

Stover J. 2003. *AIM version 4. A computer program for HIV/AIDS projections and examining the social and economic impacts of AIDS.* Spectrum system of Policy Models. The Futures Group International.

United Nations. 2002a. *HIV/AIDS and fertility in sub-Saharan Africa: A perspective of the research literature.* United Nations, New York.

United Nations. 2002b. *Fertility levels and trends in countries with intermediate levels of fertility: A background paper for the Expert Group Meeting on Completing the Fertility Transition.* 11-14 March 2002. United Nations, New York.

United Nations. 1992. *Preparing Migration Data for Subnational Population Projections.* Deparment of International and Economic and Social Affairs. United Nations, New York.

Van Aardt CJ. 2006. *Population estimates for South Africa by Magisterial district and Province, 2001 and 2006.* Bureau of Market Research, Pretoria.

Willekens F & Rogers A. 1978. *Spatial Population Analysis: Methods and Computer Programs.* International Institute for Applied System Analysis. Research Report, RR 78-18. Laxenberg, Austria.

Willekens F, Por A, & Raquillet R. 1978. *Entropy, multiproportional, and quadratic techniques for inferring detailed migration patterns from aggregate data.* International Institute for Applied System Analysis. Working Paper, WP-79-88. Laxenberg, Austria.

World Health Organisation. 2001. *Prevention of mother-to-child transmission of HIV: Selection and use of Nevirapine*. Technical notes. World Health Organisation, Geneva, Switzerland.