Projection method for South Australian Statistical Local Areas (SLAs) and Census Collection Districts (CDs) - more details

The three-tiered approach outlined in Appendix I has been further disaggregated in this accompanying paper. Apart from the births formulae all equations apply to both sexes, so sex has not been denoted.

Step 1

This involved projecting the South Australian population by age and sex, 2001 - 2008.
The cohort component method used can be summarised in the formulae below:

x	-> age
max	-> highest age projected (100+ for State, 85+ for SLAs)
t	-> base year
P	-> population
F	-> fertility rate
f	-> females
B	-> births
Q	-> death probability
OM	-> net overseas migration
IM	-> net interstate migration
NM	-> net migration (SLA projections only)

In Step 1 the following refer to interstate migration; in Step 2 they refer to overseas + inter-SLA migration.

DEP	-> departures
ARR	-> arrivals
DEPRATE	-> per capita departure rate (donor state/SLA)
ARRRATE	-> per capita arrival rate (receiving states)

For ages 0 to maximum age - 1:

(i) $$P_{x+1}(t+1) = P_x(t) * [1-Q_x(t)] + (0.5 * OM_x(t)) * (1-(0.5 * Q_x(t))) + (0.5 * OM_{x+1}(t)) * (1-(0.5 * Q_{x+1}(t)))$$

(ii) $$P_{max}(t+1) = P_{max}(t) * [1-Q_{max}(t)] + P_{max-1}(t) * [1-Q_{max-1}(t)] + OM_{max}(t) * (1-(0.5 * Q_{max}(t))) + (0.5 * OM_{max-1}(t)) * (1-(0.5 * Q_{max-1}(t)))$$

Births were then calculated:

(iii) $$B(t) = 0.5*[\sum_{x=15}^{49} (F_x(t)*P_{f,x}(t)) + \sum_{x=15}^{49} (F_x(t+1)*P_{f,x}(t+1))]$$

After constraining to projected State-level births, these were then used to calculate age 0 in the projected year:

(iv) $P_0(t+1) = B(t)*(1-Q_b(t)) + (0.5*OM_0(t)) * (1-(0.5*Q_0(t)))$

Interstate migration was calculated by applying departure rates to the South Australian population and arrival rates to the population of the remaining States and Territories (to obtain numbers departing other States to reside in South Australia). These rates were derived from 1996 Census data and were held constant for the duration of the projection.

(v) $DEP_x(t+1) = P_x(t+1) * DEPRATE_x$

(vi) $ARR_x(t+1) = P_x(t+1)^{\text{Non-SA}} * ARRRATE_x$

The resulting total arrivals and departures were then scaled to a predetermined total net interstate migration assumption. Finally, the arrivals and departures by age and sex were scaled to the new arrival and departure totals, then combined to give net age/sex interstate migration.

(vii) $IM_x(t+1) = ARR_x(t+1) - DEP_x(t+1)$

Then add the interstate migration:

(viii) $P_x(t+1) = P_x(t+1) + IM_x(t+1)$

After constraining of State age/sex population sum to the Australian-level, year t+1 then became the base for projecting the next year and the cycle was repeated until the final projection year was reached.

Step 2

This used the cohort component method to project all South Australian SLAs. The formulae in Step 1 also apply to the SLA projections, except that Net Migration (overseas + inter-SLA) was used instead of overseas and inter-SLA separately.

This slightly simpler approach was warranted as the overseas component is negligible in most SLAs in comparison with inter-SLA migration. Furthermore an annual historical time-series only exists at the SLA level for net migration, any overseas/inter-SLA split must approximated using past Census data.

For ages 0 to maximum age - 1:

(ix) $P_{x+1}(t+1) = P_x(t) * [1-Q_x(t)]$

(x) $P_{max}(t+1) = P_{max}(t) * [1-Q_{max}(t)] + P_{max-1}(t) * [1-Q_{max-1}(t)]$

Births were then calculated:

(xi) $B(t) = 0.5*(\sum_{x=15-19}^{45-49}[F_x(t)*P_{f,x}(t)] + \sum_{x=15-19}^{45-49}[F_x(t+1) * P_{f,x}(t+1)])$

After constraining to projected State-level births, these were then used to calculate age 0 in the projected year:

(xii) $P_0(t+1) = B(t) * (1-Q_b(t))$

SLA migration departures were calculated by applying mean 1996 & 2001 Census-derived departure rates to the population:

(xiii) $DEP_x(t+1) = P_x(t+1) * DEPRATE_x$

Total SLA arrivals were then derived using the pre-set net migration assumptions:

(xiv) $$ARR(t+1) = NM(t+1) - \sum_{x=0}^{x=max} DEP_x(t+1)$$

(xv) The assumed age-specific arrival levels were derived from 1996 and 2001 Census data. Together with departures from (xiii) these were simultaneously constrained (via IPF - see xvii - xix) to:

(a) SLA arrival and departure totals
(b) State age-specific net migration

Then the arrivals and departures were applied to the population:

(xvi) $P_x(t+1) = P_x(t+1) + ARR_x(t+1) - DEP_x(t+1)$

Year t+1 then became the base for projecting the next year and the cycle was repeated until the final projection year was reached. However, before $P_x(t+1)$ became the new base, the projected SLAs were constrained to sum to the State-level projection. This involved a final iterative proportional fitting process, the year is t+1:

SLA	->	Statistical Local Area
SA	->	South Australia
a	->	first SLA
z	->	last SLA
r	->	SLA number

Scale the SLA totals to the State total.

(xvii) $$P^{SLA} = P^{SLA} * (P^{SA} / \sum_{r=a}^{r=z} P_r^{SLA})$$

For each SLA, scale all ages to sum to the new SLA total.

(xviii) $$P_x^{SLA} = P_x^{SLA} * (P^{SLA} / \sum_{x=0}^{x=max} P_{xr}^{SLA})$$

For each age, scale all SLAs to sum to the State total.

(xix) $$P_x^{SLA} = P_x^{SLA} * (P_x^{SA} / \sum_{r=a}^{r=z} P_{xr}^{SLA})$$

Stages (xviii) and (xix) were then iterated several times before the resulting matrix was rounded.

Step 3

This involved splitting the completed SLA projections into Census Collection Districts.

From 2001 Census data an SLA to CD concordance was constructed giving the proportion of the SLA population aged 18+ in each CD. This concordance was then updated for each year up to 2008 to allow for differing rates of CD population growth.

CD building approval data was used for this updating, with 2001/02 approvals as an indicator of population growth. The building approvals were converted to extra population by applying dwelling-type-specific occupancy ratios from the 2001 Census for SA. The conservative assumption taken was that such differentials in CD migration would decline linearly to zero by 2008. This approach was to allow CDs to grow/decline at different rates within SLAs while reducing the risk that if 2001-02 was an unusual year for building in a CD the effect would not compound too drastically by 2008.

BAP	-> building approvals
t	-> year
CD	-> Census Collection District
F	-> CD split factor for ages 18+ (as a proportion of its SLA)
P_{2001}	-> Population in 2001
OR	-> Occupancy Ratio
d	-> Dwelling type

$$(xx) \qquad F_t^{CD} = \left(P_{2001}^{CD} + \sum_{r=2001}^{r=t} [BAP_t^{d,CD} \times OR^d] \right) / \left(P_{2001}^{SLA} + \sum_{r=2001}^{r=t} [BAP_t^{d,SLA} \times OR^d] \right)$$

where SLA represents the SLA to which the CD belongs.

The CD split factors were then applied to the SLA projections to give projected CD populations:

$$(xxi) \qquad P_t^{CD} = P_t^{SLA} * F_t^{CD}$$

While this approach produces CD projections for the age group 18+, it only indirectly 'ages' individual cohorts through distributing the effect of the SLA cohort-component projections. This means that projected CD populations may reflect the 2001 Census profile more than would actually be the case in the future.