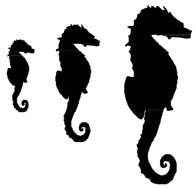


Age and Cohort Models - when CJS doesn t apply

Up until now we have made the tacit assumption that our basic "underlying" model has been the CJS time-dependent model. This has been our usual starting point.

- However, it is obvious to most biologists that there are many instances when the CJS "assumptions" are not met.
- By "assumptions" we are referring to the assumptions concerning **i**ndependence of fate and **i**dentity of rates among **i**ndividuals (the **iii** assumptions - see Lebreton *et al*. 1992). More specifically, the CJS model generally assumes that "all individuals, whatever their age or capture history, should have the same probabilities of capture and survival" (Cormack 1979 - cited in Lebreton *et al.* 1992).
- What exactly is this referring to? Recall the basic structure of the CJS time-dependent model:

survival

cohort									
		ϕ_1		ϕ_2		ϕ_3		ϕ_4	
1	**1**	→	**2**	→	**3**	→	**4**	→	**5**
				ϕ_2		ϕ_3		ϕ_4	
2			**2**	→	**3**	→	**4**	→	**5**
						ϕ_3		ϕ_4	
3					**3**	→	**4**	→	**5**
								ϕ_4	
4							**4**	→	**5**

recapture

cohort	**1**	→	**2**	→	**3**	→	**4**	→	**5**
1			p_2		p_3		p_4		p_5
			2	→	**3**	→	**4**	→	**5**
2					p_3		p_4		p_5
					3	→	**4**	→	**5**
3							p_4		p_5
							4	→	**5**
4									p_5

Age models

- It is perhaps exceedingly obvious to most biologists that individuals of different age classes (or developmental stages) differ in the probability of surviving to the next age or stage. In fact, life history theory is to a large degree focussed on analysis of such differences.
- The reasons for this are well-established. Organisms of a given age (for simplicity, we will refer only to age transitions - the logic however applies reasonably well to simple stage-structured systems as well) may be more or less vulnerable to sources of mortality than are individuals of different age(s). The reasons for these differences might reflect differences in size, behaviour, or physiological maturation. It is probably safe to say that there have been as many papers in the ecological and evolutionary literature devoted to "age-dependence" of one trait or another as any other subject.
- So, we need to be able to address the question: are there age-specific differences in survival?
- First - some definitions. Clearly, the aging process begins when individuals are born. All individuals born in a given breeding season

can be said to belong to the same birt cohort - a **cohort** is simply some criterion by which individuals are group together (birth year, in this case). Within a birth cohort, age and time are (logicall synonymous). In fact, age, time (e.g., year) and cohort are related by the following expression:

$$\text{age} = \text{current year} - \text{birth cohort}$$

- Consider the first row of a simple CJS time-dependence survival model.

cohort		ϕ_1		ϕ_2		ϕ_3		ϕ_4	
1	**1**	$\rightarrow$	**2**	$\rightarrow$	**3**	$\rightarrow$	**4**	$\rightarrow$	**5**

- The first row also corresponds to the first cohort. Let's assume that this is a "birth" cohort - all individuals that are newly marked and released at occasion 1 were marked as newborns. Let's also assume (for simplicity) that the occasions mark years.
- Thus, individuals marked at occasion 1 (0 years of age) are, if they survive, 1 year of age at occasion 2, 2 years of age at occasion 3, and so forth.
- Now, the parameters shown in this table ($\phi_1 \rightarrow \phi_4$) were originally written to show simple time-dependence. But, since individuals also age through time, we cannot differentiate between age-specific differences in survival, and simple time-specific differences. They are completely analogous.
- How do we then separate age and time effects? Clearly, this cannot be accomplished using a single cohort alone.
- However, what happens if we use multiple birth-cohorts? To examine this situation, let's make the following assumptions about some arbitrary population. First, let's assume that all individuals in each cohort are marked as newborns. Let's also assume that survival between age 0 and age 1 year is different from survival after age 1 year. For simplicity, let's call the survival from $0 \rightarrow 1$ "juvenile" survival, and survival from any age x (where x≥1 year) to x+1 "adult" survival.
- If you think about it, this is not at all an uncommon situation. However, let's make it even simpler. Let's assume for the time being that juvenile survival is constant among cohorts, and that adult survival is constant both within cohorts and among cohorts. What would the parameter structure of this model look like? Let's use the "j" subscript for juvenile survival, and the "a" subscript for adult survival.

cohort		ϕ_j		ϕ_a		ϕ_a		ϕ_a	
1	**1**	$\rightarrow$	**2**	$\rightarrow$	**3**	$\rightarrow$	**4**	$\rightarrow$	**5**
				ϕ_j		ϕ_a		ϕ_a	
2			**2**	$\rightarrow$	**3**	$\rightarrow$	**4**	$\rightarrow$	**5**
						ϕ_j		ϕ_a	
3					**3**	$\rightarrow$	**4**	$\rightarrow$	**5**
								ϕ_j	
4							**4**	$\rightarrow$	**5**

- Now, let's look at this table and see how it makes sense. The best starting point is to look at each cohort separately. Within a cohort, we see that between the first occasion (in the cohort), and the second occasion after marking, individuals survive with probability ϕ_j. However, from the second occasion after marking onwards (within a cohort), they survive with probability ϕ_a.
- Now, as we discussed previously, within a cohort we cannot differentiate between "time" and "age". However, note that we can test whether a model with this structure differs from one where (say) survival is constant (no age or time effect). But what we're really after is "age" as separate from "time".
- Here is where multiple cohorts come in. By contrasting parameter estimates among cohorts, but within a time interval, we can differentiate age and time effects. For example, concentrate on the interval between occasion 2 and occasion 3 (shaded - below).

cohort	
1	$1 \xrightarrow{\phi_j} 2 \xrightarrow{\phi_a} 3 \xrightarrow{\phi_a} 4 \xrightarrow{\phi_a} 5$
2	$2 \xrightarrow{\phi_j} 3 \xrightarrow{\phi_a} 4 \xrightarrow{\phi_a} 5$
3	$3 \xrightarrow{\phi_j} 4 \xrightarrow{\phi_a} 5$
4	$4 \xrightarrow{\phi_j} 5$

- In terms of time, both cohorts 1 and 2 are experiencing the same "temporal effect". In other words, all individuals, whether they were newly marked at occasion 1 or occasion 2, are experiencing the aspects of the "environment" causing mortality during this interval. However, in cohort 1, the individuals at occasion 2 are 1 year of age, while those from cohort 2 are newborn. Thus, for cohort 1 individuals, they will survive from 2 → 3 with rate ϕ_a.
- In contrast, for cohort 2 individuals, they are surviving from 2 → 3 at rate ϕ_j.
- If there were no age-specific differences in survival, then the ratio of survival of cohort 1 individuals over this interval would be the same as the survival of cohort 2 individuals over this interval. In other words, the ratio of ϕ_a/ϕ_j would equal 1.
- Again, it is the contrast among cohorts (i.e., rows), but within columns (i.e., intervals between occasions) that allows us to test for age differences in survival. This is a very important concept to grasp, so it is critical that you spend the time now to make sure you do.
- Let's expand our model somewhat, adding some more "flexibility". Suppose that juvenile survival varies over time, but that adult survival is constant through time. What is the time axis of our model? The time axis which we need to follow in an age-structured model is along the diagonal. For example, for our model with constant juvenile and constant adult survival, we essentially have 2 parameters. The different shaded areas of our model (below) show the juvenile and adult age classes, respectively:

cohort	
1	$1 \xrightarrow{\phi_j} 2 \xrightarrow{\phi_a} 3 \xrightarrow{\phi_a} 4 \xrightarrow{\phi_a} 5$
2	$2 \xrightarrow{\phi_j} 3 \xrightarrow{\phi_a} 4 \xrightarrow{\phi_a} 5$
3	$3 \xrightarrow{\phi_j} 4 \xrightarrow{\phi_a} 5$
4	$4 \xrightarrow{\phi_j} 5$

- The juvenile age class in this example spans one time interval (i.e., one year). The adult age class (above the diagonal) spans n-1 years, where n is the number of occasions, and "1" is the duration of the juvenile age class. If we re-write the matrix using numbers for subscripts instead of letters (let "1" = "j", and "2" = "a"), then

cohort	
1	$1 \xrightarrow{\phi_1} 2 \xrightarrow{\phi_2} 3 \xrightarrow{\phi_2} 4 \xrightarrow{\phi_2} 5$
2	$2 \xrightarrow{\phi_1} 3 \xrightarrow{\phi_2} 4 \xrightarrow{\phi_2} 5$
3	$3 \xrightarrow{\phi_1} 4 \xrightarrow{\phi_2} 5$
4	$4 \xrightarrow{\phi_1} 5$

- Now, if we simplify this and re-write the parameter structure the way that SURGE interprets it (using only the subscripts), this 2 age-class model (juvenile, adult) is written as:

1	2	2	2
	1	2	2
		1	2
			1

- So, if we add time-dependence to the juvenile survival rate, but leave adult survival constant, the model structure would now look like:

cohort	
1	$\mathbf{1} \xrightarrow{\phi_1} \mathbf{2} \xrightarrow{\phi_5} \mathbf{3} \xrightarrow{\phi_5} \mathbf{4} \xrightarrow{\phi_5} \mathbf{5}$
2	$\mathbf{2} \xrightarrow{\phi_2} \mathbf{3} \xrightarrow{\phi_5} \mathbf{4} \xrightarrow{\phi_5} \mathbf{5}$
3	$\mathbf{3} \xrightarrow{\phi_3} \mathbf{4} \xrightarrow{\phi_5} \mathbf{5}$
4	$\mathbf{4} \xrightarrow{\phi_4} \mathbf{5}$

- In the SURGE parameter format, this would reduce to:

1	5	5	5
	2	5	5
		3	5
			4

- Let's extend it one more step: let's add time dependence to the adult survival as well.
- This particular model is important since (as we'll see shortly) it is one of the "built-in" models SURGE provides in the model choice menus. This reflects the fact that 2 age-class models, with a juvenile age-class spanning one year, and a single adult age class, with time-dependence in both, are very commonly seen in analysis of mark-recapture data from wild populations.
- Here is what the 2 age-class model with time-dependence in both age classes would look like:

cohort	
1	$\mathbf{1} \xrightarrow{\phi_1} \mathbf{2} \xrightarrow{\phi_5} \mathbf{3} \xrightarrow{\phi_6} \mathbf{4} \xrightarrow{\phi_7} \mathbf{5}$
2	$\mathbf{2} \xrightarrow{\phi_2} \mathbf{3} \xrightarrow{\phi_6} \mathbf{4} \xrightarrow{\phi_7} \mathbf{5}$
3	$\mathbf{3} \xrightarrow{\phi_3} \mathbf{4} \xrightarrow{\phi_7} \mathbf{5}$
4	$\mathbf{4} \xrightarrow{\phi_4} \mathbf{5}$

- This matrix reduces to:

1	5	6	7
	2	6	7
		3	7
			4

- As a final test, to see if you really understand the structure of these models, consider the following situation. We have an 8 occasion study of a long-lived organism with indeterminate growth, and we believe that survival may be age-dependent. We decide to model survival in the following way. 3 age classes, the first age class spanning 1 year, the second age class spanning 1 year, and the final age class spanning all remaining years. In other words, a single-year duration "juvenile" phase, a 1-year duration "sub-adult" phase, and final the "adult" phase. Juvenile and sub-adult survival are time-dependent, but adult survival is constant over time. Here is the structure for this model:

cohort	1 → 2	2 → 3	3 → 4	4 → 5	5 → 6	6 → 7	7 → 8
1	ϕ_1	ϕ_8	ϕ_{14}	ϕ_{14}	ϕ_{14}	ϕ_{14}	ϕ_{14}
2		ϕ_2	ϕ_9	ϕ_{14}	ϕ_{14}	ϕ_{14}	ϕ_{14}
3			ϕ_3	ϕ_{10}	ϕ_{14}	ϕ_{14}	ϕ_{14}
4				ϕ_4	ϕ_{11}	ϕ_{14}	ϕ_{14}
5					ϕ_5	ϕ_{12}	ϕ_{14}
6						ϕ_6	ϕ_{13}
7							ϕ_7

- With a bit of thought, you should be able to see how this model was constructed (look carefully at the ordering of the parameter subscripts). If not, go back through the preceding few pages, and try again. Age models are very important.
- If you do "get the basic idea", then let's proceed to analysis of a simple 2 age-class model. We will examine how to use the built-in menu options in SURGE to test 2 different types of age models -as we will see in the next chapter, by using a user-defined model, we can construct an age (or cohort) model of an arbitrary design. For the moment, we'll focus on the models you can construct from the SURGE menus.
- In this example data set (AGE.SUR), we suspect that there are 2 age-classes. We want to "confirm" our suspicion by using SURGE to test the fit of a 2 age-class model versus the standard time-dependent CJS model with no age effects. There is only one group.
- Start SURGE, and send the output to AGE.LST. Let's start with the standard CJS model (title "Phi(t),p(t)"). By now, you should be able to run the CJS model without any prompting, so go ahead and run the model, with full time-dependence in both survival and recapture. Then, choose option "2" from the final menu to allow you to run another model.
- This time, we'll run one of the age models. Which one? Well, remember the structure of the model specification menu (Fig. 7.1).

```
M* MODEL CHOICE : SURVIVAL  PROBABILITIES

M*           AGE  DEPENDENCE                = 1
M*           TIME DEPENDENCE                = 2
M*           CONSTANCY  OVER AGE AND TIME   = 3
M*           USER DEFINED MODEL             = 4
M*   TIME DEPENDENCE,  DEP. ON EXT.VAR.     = 5
M*   2 AGE CLASSES 1=TIME DEP. 2=TIME DEP.  = 6
M*                 1=TIME DEP. 2=EXT.VAR.   = 7
M*                 1=EXT.VAR.  2=TIME DEP.  = 8
M*                 1=EXT.VAR.  2=EXT.VAR.   = 9

M* OPTION FOR DATA SET males.sur ?
```

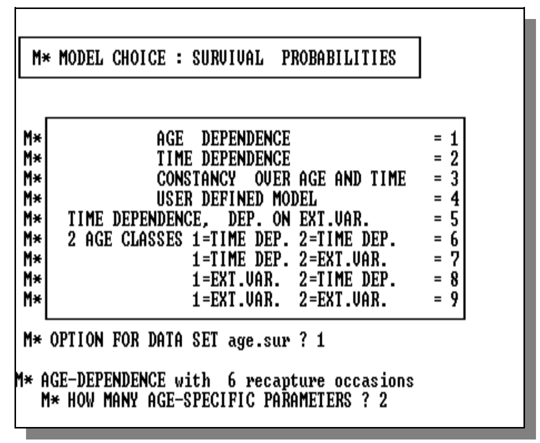
Fig. 7.1

- Notice that there are 2 basic groupings of age models on this menu. First, choice "1" - AGE DEPENDENCY. And then, down below, a series of 2 age class models - specifically, choice "6" with TIME DEPENDENCE for age class 1 and age class 2. This reflects the 2 age-class structure of these models - the model for the youngest age class (1) and the oldest (2). For the moment, we will consider only choice “6” from this group of 2 age-class models.

- What is the difference between choice “1” and choice “6”? In simplest terms:

Menu Choice	Model
choice 1	• arbitrary number of age classes - up to n, where n = number of cohorts. Age classes of arbitrary range (i.e., number of years) • survival constant over time within each age class
choice 6	• only 2 age classes • first age class has one time interval duration (i.e., “juvenile” and “adult”) • full time-dependence in both age classes

- So, these 2 menu choices represent extreme models of a sort: choice 1 lets you have many age classes, but the parameter estimates are constant within age class, while choice 6 gives you full time-dependence, but only 2 age classes, with the first age class spanning only one interval.
- While these may seem limited, in fact they are adequate for a large number of studies. Although we will see in the next chapter how to use a user-defined model to greatly extend the range of base models SURGE can use, for the moment, let’s look at these 2 choices, and apply them to our sample data set.
- Let’s start with choice “6” - full time-dependence - the closest (in some respects) to our CJS starting model, with the addition of age-structure.
- For a title, “Phi(a2*t),p(t)” - in other words - 2 age classes for survival, both time dependent, and no age classes for recapture, but full time-dependence”.
- Now, when you get to the model choice menu for survival, select choice “6” - 2 age classes with time dependence. SURGE assumes that for this model you want full time-dependence for both age classes, so it doesn’t ask you any further questions about the survival model. It proceeds immediately to the recapture model.
- Here, full time-dependence (choice “2”).
- Run the analysis, and when finished, go back to start another analysis (choice “2” from the final menu”).
- This time, we’re going to use choice “1” - which allows for more age classes, but no time-dependence. As a starting point, let’s use 2 age-classes. For a title, “phi(a2),p(t)” - 2 age classes in survival, but constant (no time), and time-dependence in recapture, but no age structure.
- From the model specification menu, select choice “1”. Now, however, since this choice gives you the option of using many age classes, SURGE asks you how many age classes you want to use. For this analysis, we’re going to use 2, so we answer “2” (Fig 7.2).



```
M* MODEL CHOICE : SURVIVAL  PROBABILITIES

M*            AGE  DEPENDENCE                  = 1
M*            TIME DEPENDENCE                  = 2
M*            CONSTANCY  OVER AGE AND TIME     = 3
M*            USER DEFINED MODEL               = 4
M*  TIME DEPENDENCE,  DEP. ON EXT.VAR.         = 5
M*  2 AGE CLASSES 1=TIME DEP. 2=TIME DEP.      = 6
M*                1=TIME DEP. 2=EXT.VAR.       = 7
M*                1=EXT.VAR.  2=TIME DEP.      = 8
M*                1=EXT.VAR.  2=EXT.VAR.       = 9

M* OPTION FOR DATA SET age.sur ? 1

M* AGE-DEPENDENCE with  6 recapture occasions
  M* HOW MANY AGE-SPECIFIC PARAMETERS ? 2
```

Fig. 7.2

- Now that we’ve told SURGE how many age classes we want to use, we next have to tell it how “wide” to make each of the two age classes

- in other words, how many intervals in duration are each of the age classes. In our 2 age-class, time-dependent model, we had an interval of 1 for the first age class - thus with only 2 age classes, the interval for second age class spans the remainder of the study.

- If we use this as the basic interval structure for the constant model, we need to tell SURGE we want the first age class to span one interval, and the second age class to span the remaining intervals. In this case, the first age class spans 1 year, while the second age class spans the remaining years - 5 years (since there are 6 total recapture occasions). So, we enter "1 5" (Fig. 7.3). Obviously the actual "width" of the final (or second) age class depends on the cohort (5 years for the first cohort, 4 years for the second cohort, and so forth) - the menu is asking for the *maximum* size of a particular age class (5 in this example).

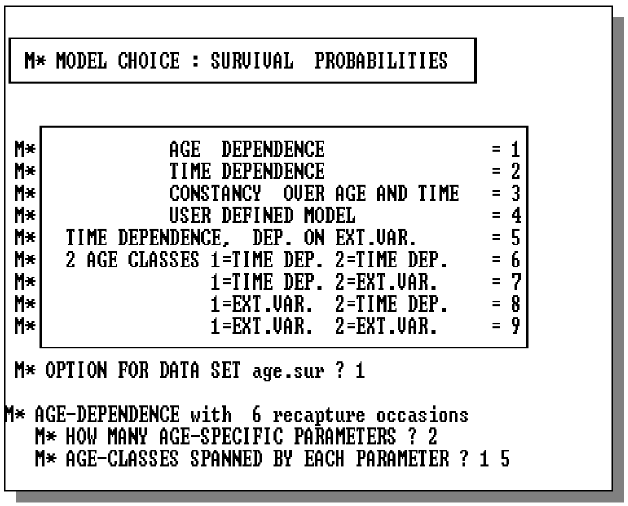


Fig. 7.3

- Next, for recapture, we again choose full time-dependence for recapture.
- Run the rest of the analysis, and then quit SURGE, and examine the output.
- First, the CJS estimates. The model deviance is 2837.939. With 11 parameters (AIC = 2859.939). The model structure for this model (which you should be able to do in your sleep by now!) is:

1	2	3	4	5	6
	2	3	4	5	6
		3	4	5	6
			4	5	6
				5	6
survival					6

7	8	9	10	11	12
	8	9	10	11	12
		9	10	11	12
			10	11	12
				11	12
recapture					12

- For the 2 age-class model with time dependence (the second part of the output file), the model deviance is 2772.750, which is substantially smaller than the deviance for the CJS model (65.190 smaller).
- Of course, the question is does it fit *significantly* better? In order to answer this question, we must know how many parameters there are in this model. Here is the model structure for both survival and recapture:

1	7	8	9	10	11
	2	8	9	10	11
		3	9	10	11
			4	10	11
				5	11
survival					6

12	13	14	15	16	17
	13	14	15	16	17
		14	15	16	17
			15	16	17
				16	17
recapture					17

- Note that these values are incremented started with the final value of the survival triangle.

- Now, how many individually identifiable parameters are in this model? Here are the saturated capture histories, and their associated probability functions - make sure you see how these functions were derived from the survival and recapture matrices:

capture history	**probability**
1111111	$\phi_1 p_2 \phi_7 p_3 \phi_8 p_4 \phi_9 p_5 \phi_{10} p_6 \phi_{11} p_7$
0111111	$\phi_2 p_3 \phi_8 p_4 \phi_9 p_5 \phi_{10} p_6 \phi_{11} p_7$
0011111	$\phi_3 p_5 \phi_9 p_5 \phi_{10} p_6 \phi_{11} p_7$
0001111	$\phi_4 p_5 \phi_{10} p_6 \phi_{11} p_7$
0000111	$\phi_5 p_6 \phi_{11} p_7$
0000011	$\phi_6 p_7$

- Now, of course, the next step is to determine which of the parameters in this table are individually identifiable. By now you may have realized that the "critical" part of this process typically involves looking at the terminal products. In this case, we have 2 different products: $\phi_{11}p_7$ and $\phi_6 p_7$. Are these β terms? If you think back to the earlier cohort example in this chapter, you will realize that the answer is yes - since p_7 is not identifiable. Thus, 16 identifiable paramters.
- This model corresponds to Table 7D in Lebreton et al. (1992). They note that for this model, the number of identifiable parameters is given as (3*k*-5), where k = the number of occasions. Since *k*=7 in this example, we see that (3*k*-5) = 21-5 = 16 parameters.
- Thus, with a model deviance of 2772.750, the AIC for this model is 2804.750. Even though this model has more parameters than the CJS model we first tried (16 versus 11), the AIC value for this model is quite obviously smaller than the CJS model (2804.750 versus 2859.939). We can confirm this using a LRT: the 2 age-class model fits significantly better than the CJS model (χ^2=65.19, df=5, P<0.001).
- What about the 2 age-class model with time dependence versus the 2 age-class model with constant survival? For the constant model, the deviance was 2792.804, which is larger than for the 2 age-class model with time-dependence. The question is, again, is this difference significant?
- For the constant model, there are only 2 parameters estimated for the survival rates, and 6 for recaptures (remember, with constant survival we can estimate all of the recapture rates in this model). Thus, the AIC for the model is 2808.804 (versus 2804.75 for the time-dependent model).
- Thus, there is a relatively small difference in AIC values for the time-dependent and constant age models (both are clearly better than the CJS model).
- What does the LRT show? The difference between these 2 models is significant (χ^2=20.054, df=8, P=0.010). Therefore, we reject the null hypothesis of no difference between these 2 models, and accept the more parameterized 2 age-class model with time-dependence as the most parsimonious.
- In biological terms, what have we shown: we have tested 2 major hypotheses: (1) there is age-dependence in survival, and (2) within age-class, there is time variation in survival.
- Now let's try 2 more examples, using the same data set: (1) using 3 age classes (spanning 1,2 and the remaining intervals respectively, and (2) imposing a linear constraint on both age classes in a 2 age-class time-dependent model.

More than 2 age classes

- Clearly, for both time-dependent, and constant age models, we may want to include more than 2 age classes in our model. We also might want some of the age classes to be time-dependent, and some to be constant. As mentioned, any model can be created by using a user-defined model. However, for the moment, let's see how we can use option "1" to add more age classes. In our first example, we had 2 age-classes - the first spanned one interval, and the second spanned the remaining intervals. In this example let's try a model which has 3 age classes: the first age class will span one interval. The second age class will span 2 intervals, while the third age class spans the rest. Since we are using option 1, survival is constant within age class.

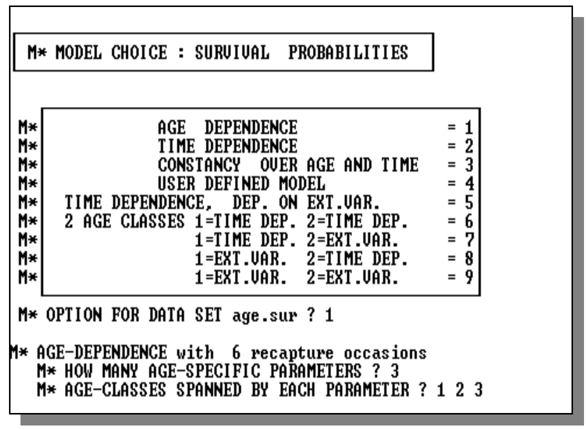
- This example differs from the previous one both in terms of the number of age classes included (3 versus 2), and also in terms of the number of intervals spanned by some of the age classes. The structure of this model would be:

cohort													
		ϕ_1		ϕ_2		ϕ_2		ϕ_3		ϕ_3		ϕ_3	
1	**1**	→	**2**	→	**3**	→	**4**	→	**5**	→	**6**	→	**7**
				ϕ_1		ϕ_2		ϕ_2		ϕ_3		ϕ_3	
2			**2**	→	**3**	→	**4**	→	**5**	→	**6**	→	**7**
						ϕ_1		ϕ_2		ϕ_2		ϕ_3	
3					**3**	→	**4**	→	**5**	→	**6**	→	**7**
								ϕ_1		ϕ_2		ϕ_2	
4							**4**	→	**5**	→	**6**	→	**7**
										ϕ_1		ϕ_2	
5									**5**	→	**6**	→	**7**
												ϕ_1	
6											**6**	→	**7**

- This model structure corresponds to:

1	2	2	3	3	3
	1	2	2	3	3
		1	2	2	3
			1	2	2
				1	2
					1

- Again, pay careful attention to the subscripting on each parameter - from left to right within each row we can see the number of age classes (3), and the "width" of each age class (1 interval, 2 intervals and then all remaining intervals in the final age class).
- Before we see how to implement this with SURGE, we need to think a minute as to what ϕ_2 really means: by constructing the model this way, we're essentially telling SURGE that the probability of an individual surviving from the second occasion after the marking occasion to the third occasion is the same as from the 3rd to the 4th.
- You specify this model by selecting choice "1" from the menu. SURGE will then ask you for the number of age classes. In this case, we answer "3".
- Then SURGE will ask you how many age-classes (i.e., intervals) are spanned by each parameter (i.e., by each age class)? Here, we enter "1 2 3", for 1, 2 and 3 years, respectively (6 total intervals).



```
M* MODEL CHOICE : SURVIVAL  PROBABILITIES

M*            AGE  DEPENDENCE                  = 1
M*            TIME DEPENDENCE                  = 2
M*            CONSTANCY  OVER AGE AND TIME     = 3
M*            USER DEFINED MODEL               = 4
M*  TIME DEPENDENCE,  DEP. ON EXT.VAR.         = 5
M*  2 AGE CLASSES 1=TIME DEP. 2=TIME DEP.      = 6
M*                1=TIME DEP. 2=EXT.VAR.       = 7
M*                1=EXT.VAR.  2=TIME DEP.      = 8
M*                1=EXT.VAR.  2=EXT.VAR.       = 9

M* OPTION FOR DATA SET age.sur ? 1

M* AGE-DEPENDENCE with  6 recapture occasions
  M* HOW MANY AGE-SPECIFIC PARAMETERS ? 3
  M* AGE-CLASSES SPANNED BY EACH PARAMETER ? 1 2 3
```

- Use the time-dependent model for recaptures (choice "2"), and proceed with the rest of the analysis.
- This model has a deviance of 2792.803, with 9 estimable parameters (since we added one more age class to our previous "constant" model, we've added one more parameter). Thus, the AIC for this model is 2810.803, which is slightly larger than the previous "constant" model with only 2 age-classes (2808.804).

Constraining an age model

- Consider the following case: 2 age-classes, with time-dependence in each age class. We have some reason to believe that there is a linear change over time in both age classes. How would we fit this model?
- Well, if you think about it, it is very similar in structure to a classical ANCOVA - except that here our two groups represent the two age classes. In other words, we want to compare the slopes of the linear trend in survival between young and old individuals. If the slopes are not significantly different, we could test against a model with a common slope for both age classes.
- First, we need to know the index values of the parameters we're going to constrain. For this example:

```
1   7   8   9   10  11
    2   8   9   10  11
        3   9   10  11
            4   10  11
                5   11
  survival          6
```

```
12  13  14  15  16  17
    13  14  15  16  17
        14  15  16  17
            15  16  17
                16  17
  recapture         17
```

- We are going to constrain survival. Thus for "juveniles", the parameters of interest are 1 → 6, and for "adults", the parameters are 7 → 11.
- So, let's build our PAR file. Again, we might as well call it LINEAR.PAR. Since we have 11 parameters total that we want to constrain, the LINEAR.PAR file would have the following structure:

```
11

1  2  3  4  5  6  7  8  9  10  11
```

- Now, what about the VAR file (we'll call it LINEAR.VAR)? Take a look back at Chapter 6. In the section where we discussed building VAR files to test for linear trend for 2 groups, you'll see that all we needed to do was construct a file with 1 column indicating group association, 1 column of increasing dummy variables to index the trend, and then a final column to account for a possible interaction of group and trend (see pp. 6-16 to 6-18).
- However, if you look back at the example in Chapter 6, you'll notice that, in that case, the two groups were "symmetrical" - had the same number of occasions for both groups. In our age example, however, this is not the case: we have 6 parameters for the juvenile age class (1 → 6) but only 5 parameters for the adult age class (7 → 11).
- Why is this? Well, the answer is obvious if you look back at the structure of our model - we obviously don't have an "adult" estimate over the first interval for the first cohort, since there were no "known" adults at that point in the study.
- In fact, this is not much of a problem - you could, of course, accommodate this "non-overlap" in our LINEAR.VAR file. For example, to account for the overlap, you could have 6 "1" values for the 6 juvenile parameters, but only 5 "0" values for the 5 adult parameters.
- How do we handle the vector of increasing values to handle the trend? Should we use 1 → 6 for juveniles, and 2 → 6 for adults (which corresponds to where the overlap occurs), or is it equivalent to use 1 → 6 and 1 → 5 respectively? As it turns out, it doesn't matter at all to SURGE - at least not mechanically (i.e., SURGE will run in either case). However, note that in the second case you would, in effect, be coding each time step differently for each group, which makes the intercepts no longer comparable. Thus, if you choose to accomodate the overlap, the first coding scheme (1 → 6 for juveniles, and 2 → 6 for adults in this example) is prefered.
- Another possiblity, of course, is to drop the first juvenile parameter altogether from our constraint. In other words, to include only those parameters which "overlap" (i.e., 2 → 6 and 7 → 11). If we do this, our PAR file would have only 10 parameters: 2 → 11. Since the difference in overlap here is only 1 interval, it is unlikely to make a significant difference.

- For our trend analysis, we'll simply use the following LINEAR.VAR file, which includes survival for the first age class over the first interval

```
11  3
$
$
1  1  1
1  2  2
1  3  3
1  4  4
1  5  5
1  6  6
0  2  0
0  3  0
0  4  0
0  5  0
0  6  0
```

- All you need to do now is run SURGE, and apply the constraint to the underlying model (2 age-classes with time-dependence for survival, simple time-dependence for recaptures). And, as discussed in detail in Chapter 6, by varying which columns of the VAR file you use in the constraint, you can test hypotheses concerning equivalence of slope between age classes, equivalence of intercepts, and so forth. You could also test for additivity (parallelism) in survival for a time-dependent model. **Everything** you learned in Chapter 6 in terms of constraining an underlying CJS model applies equally well to age models (or cohort models, or ANY models) **- *all you need to do is know the parameter structure of the underlying model you want to constrain, and then build the appropriate PAR and VAR files*!**

Using data where both young and adults are marked

- Everything we have covered concerning age models has to this point assumed that we were dealing with individuals all marked as young (or, at the least), at a common age.
- However, very commonly when we take samples from populations we sample individuals from several age classes, and individually mark any unmarked individuals. For example, if our sample contains both newborns and adults, we will mark both age classes. Of course, we could choose to just mark the young, but this is often not desirable.
- Analytically, the question becomes - how to deal with data from both groups of newly marked individuals? In fact, the use of the word "group" was intentional - the solution is simply to treat both types of birds as different groups with SURGE.
- Previously, we have compared males and females, or controls and treatments, or good and poor colonies. In this case we simply have "marked as young" versus "not marked as young".
- In this case, we are testing the hypothesis that survival within an age-class over a given interval doesn't depend upon the age at which the individual was marked.
- How do we do this? User-defined models! Next Chapter!

Cohort models

- As first discussed in Chapter 4, and at length in Lebreton *et al*. (1992), the time-dependent CJS model assumes that survival and recapture rates are constant over cohorts. In other words, neither survival nor recapture are affected by cohort.
- Is this a reasonable assumption? Under many instances, it may very well be reasonable. However, there may be good biological reasons to expect cohorts to differ. Suppose, for example, that animals newly marked on occasion 3 were present in the population on occasion 1, but were simply "missed" from the marking sample. Perhaps there is a "reason" why these animals were missed - and this reason might influence subsequent survival or (perhaps more likely) recapture rate.

- Another common example is that cohorts differ in the environmental conditions experienced by the individuals during marking, such that subsequent survival, or recapture, or both, are affected.
- In essence then, a cohort model is simply one where survival and/or recapture rates differ as a function of the cohort an animal is first released into.
- In it's simplest form, a cohort model can be represented as:

cohort	
1	$\mathbf{1} \xrightarrow{\phi_1} \mathbf{2} \xrightarrow{\phi_1} \mathbf{3} \xrightarrow{\phi_1} \mathbf{4} \xrightarrow{\phi_1} \mathbf{5}$
2	$\mathbf{2} \xrightarrow{\phi_2} \mathbf{3} \xrightarrow{\phi_2} \mathbf{4} \xrightarrow{\phi_2} \mathbf{5}$
3	$\mathbf{3} \xrightarrow{\phi_3} \mathbf{4} \xrightarrow{\phi_3} \mathbf{5}$
4	$\mathbf{4} \xrightarrow{\phi_4} \mathbf{5}$

survival

cohort	
1	$\mathbf{1} \rightarrow \underset{p_2}{\mathbf{2}} \rightarrow \underset{p_2}{\mathbf{3}} \rightarrow \underset{p_2}{\mathbf{4}} \rightarrow \underset{p_2}{\mathbf{5}}$
2	$\mathbf{2} \rightarrow \underset{p_3}{\mathbf{3}} \rightarrow \underset{p_3}{\mathbf{4}} \rightarrow \underset{p_3}{\mathbf{5}}$
3	$\mathbf{3} \rightarrow \underset{p_4}{\mathbf{4}} \rightarrow \underset{p_4}{\mathbf{5}}$
4	$\mathbf{4} \rightarrow \underset{p_5}{\mathbf{5}}$

recapture

- In this case, survival and recapture are constant over time, but differ among cohorts.

- How would this be represented by SURGE?

1	1	1	1
	2	2	2
		3	3
survival			4

5	5	5	5
	6	6	6
		7	7
recapture			8

- Of course, we could add time-dependence within-cohort, but this will substantially increase the number of parameters.
- For example, consider the following cohort-time model for survival - estimates differ over time within cohort, but over a given interval, they differ among cohorts.

"survival"
cohort and time

cohort	
1	$\mathbf{1} \xrightarrow{\phi_1} \mathbf{2} \xrightarrow{\phi_2} \mathbf{3} \xrightarrow{\phi_3} \mathbf{4} \xrightarrow{\phi_4} \mathbf{5}$
2	$\mathbf{2} \xrightarrow{\phi_5} \mathbf{3} \xrightarrow{\phi_6} \mathbf{4} \xrightarrow{\phi_7} \mathbf{5}$
3	$\mathbf{3} \xrightarrow{\phi_8} \mathbf{4} \xrightarrow{\phi_9} \mathbf{5}$
4	$\mathbf{4} \xrightarrow{\phi_{10}} \mathbf{5}$

- Not only is the number of parameters increased significantly, but the number of estimable and non-estimable parameters is also changed. For example, suppose that the preceding model is applied to survival, but simple time-dependence for recaptures.

“recapture”
time

cohort	**1**	→	**2**	→	**3**	→	**4**	→	**5**
1			p_2		p_3		p_4		p_5
			2	→	**3**	→	**4**	→	**5**
2					p_3		p_4		p_5
					3	→	**4**	→	**5**
3							p_4		p_5
							4	→	**5**
4									p_5

- How many estimable parameters would there be? As discussed in Chapter 4, one way to count the number of estimable parameters for any single-group model is to write out the probability statements for each of the “saturated” capture histories.
- For this example, the saturated histories, and their corresponding probability statements are:

11111	$\phi_1 p_2 \phi_2 p_3 \phi_3 p_4 \phi_4 p_5$
01111	$\phi_5 p_3 \phi_6 p_4 \phi_7 p_5$
00111	$\phi_8 p_4 \phi_9 p_5$
00011	$\phi_{10} p_5$

- How many unique, identifiable parameters are there? Clearly, the key to answering this question lies in the terminal product terms (i.e., $\phi_4 p_5$, $\phi_7 p_5$...). Each of these β terms contains p_5. Is p_5 estimable? No - not without more information! As such, we have 4 different β terms - and thus 13 parameters.
- In contrast, how many estimable parameters would there have been if *both* survival and recapture models were both cohort AND time-dependent? In this case, the saturated histories and the associated probabilities would look like:

11111	$\phi_1 p_2 \phi_2 p_3 \phi_3 p_4 \phi_4 p_5$
01111	$\phi_5 p_6 \phi_6 p_7 \phi_7 p_8$
00111	$\phi_8 p_9 \phi_9 p_{10}$
00011	$\phi_{10} p_{11}$

- In this case, we see a clear difference in the terminal products: there is no common term among cohorts. Thus, as you might expect, none of these product terms are individually identifiable - they are all β terms. Thus, in this example, we have 16 individually identifiable parameters (out of a total of 20 parameters in the model).
- As you can see, with progressively complex models, it can be more work to count the number of parameters. And, clearly, knowing the number of parameters is essential for model selection.
- How do you build cohort models in SURGE? Well, the answer is - once again - user-defined models! So clearly, this topic is very important, since it allows you great flexibility in building models of arbitrary design.

That is the end of Chapter 7. We have considerably expanded the range of “underlying” models we can fit to mark-recapture data, using just the options SURGE provides in the model specification menus. We have also seen how constraints can be applied to these models as easily as we did with CJS models.

In the next chapter - user-defined models! A very important concept that gives SURGE considerable flexibility as an analytical tool. And, very easy to apply as well. If you’ve understood everything we’ve done to this point (of course!).