Ch. 13

Robust design[1]

"The investigator should have a robust faith—and yet not believe."
-- Claude Bernard

Questions to ponder:
- *Why is this model structure called 'the robust design'?*
- *How are closed and open models combined in this model structure?*
- *How can the model estimate emigration?*
- *How can I estimate the size of a super-population with the parameter estimates from my analyses?*

The ways of innovation

Advances in methods for parameter estimation allow new perspectives on the dynamics of populations. The 'spark' required for these advances requires basic understanding of the probability-based models that are present and some innovation on the part of the scientist. The robust design approach to mark-recapture analyses is an example of innovation and, simply, a good idea. Here is a quote from Kenneth Pollock, the 'father' of robust design as he explains the inspiration for his initial work (Pollock 1982):

> *"During the preparation of a review of capture-recapture methods….I realized that statisticians have drawn a sharp distinction between closed and open population models that is perhaps rather artificial. Here I describe a design for long-term studies that is robust to het-erogeneity and/or trap response. It allows an analysis that uses methodology from closed and open population models."*

The model structure that Pollock provided was innovative, and it was also prone to quick use by ecologists because it did not require (in most cases) a new approach to sampling. In fact, Pollock's approach uses the annual sampling approaches that are common to many fisheries or wildlife monitoring schemes—an annual repetition of several capture events that occur in a relatively short period of time. The robust design is so-named, because Pollock suggested that it was "robust" to the problem of unequal probability of capture—something that was assumed in contemporary models of his day for parameter estimation from either open or closed populations. Thus, the problem of trap-shyness or sampling efforts that varied through time could be alleviated with the new model structure.

---

[1] *With thanks for content to William Kendall, James Nichols, and Michael Conroy*

Pollock initially conceived his model as being useful to estimate a new parameter—birth rate. In subsequent work, Nichols and Pollock (1990) provided a mechanism to modify the robust design to estimate immigration. Eventually, Kendall et al. (1997) modified the method to allow estimation of emigration as well as immigration. The Robust design has undergone multiple recent innovations, including combining it with multi-state model structures, but we will focus on the basic robust design to provide a foundation to understand more complicated model structures covered in more advanced texts.

Robust design basics

In a typical CJS-type model (see Chapter 10), we assume that an animal with the capture history of 101 was at the study site during time 2, but was not captured. That is, the structure of CJS-type models does not allow for temporary emigration. We can imagine, of course, many situations where that assumption does not hold, and our estimate for the encounter probability is then biased low. Animals have home ranges or movements that take them, temporarily away from the area that we have selected as our study site. It is very conceivable that animals spend some time, during our study, away from the study site.
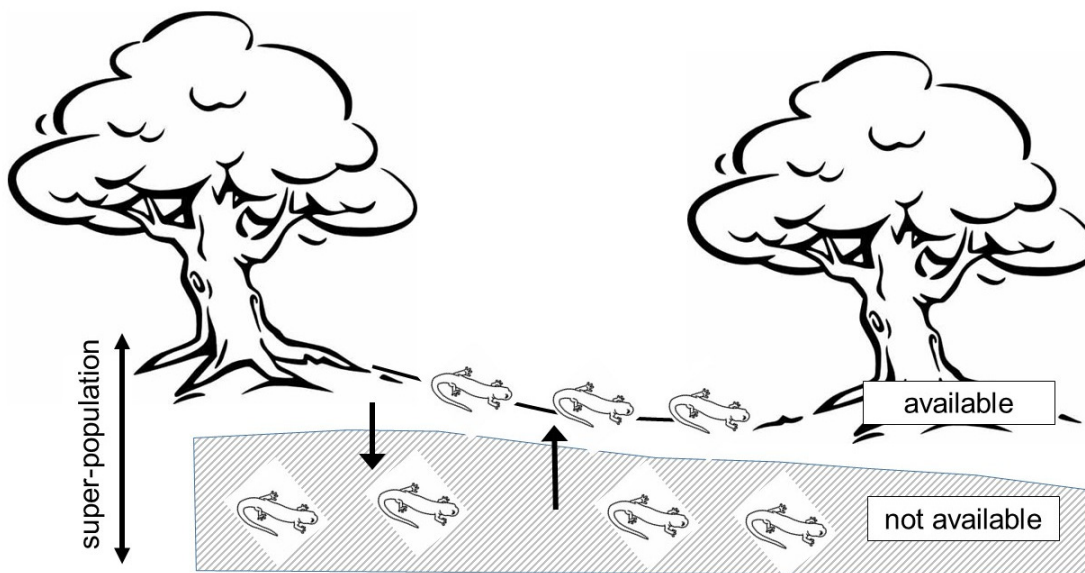
Figure 13.1: *Depiction of the super-population concept after Bailey et al. (2004). Terrestrial salamanders are available for counting by surveyors only when they are on the surface of the forest floor. The super-population consists of both available and unavailable animals.*

An example provided by Bailey et al. (2004) involves terrestrial salamanders (*Plethodontidae*) that temporarily emigrate away from the surface of the forest where biologists monitor them. The problem was that population size estimates were thought to be constantly lower than the actual population size, because only the salamanders on the surface were captured. And, in fact, Bailey et al.'s (2004) analysis confirmed that "significant proportions of terrestrial salamander populations are subterranean." In fact, the biologists estimated that 87% of the population was unavailable, and thus the animals that could be sampled on the surface was a small fraction of

the entire population. The robust design is a good approach when a large portion of the population is not available for encounter during a given sample period.

In robust design terms, we refer to the set of animals that may be sampled at some point during our study as the **super-population**. The super-population, by definition, is larger than the sub-set of the population that is "available" for encounter during a short period of study because other individuals are not available. But, those not-available individuals in the super-population may become available at some time in the future, or they were available at some time in the past.

When a biologist attempts to estimate "population size" for a population, they often are interested in the size of the super-population—not just the size of the small group of animals currently inhabiting their study site. And, this is a good reason to use the methods of robust design.

Combining open and closed structures
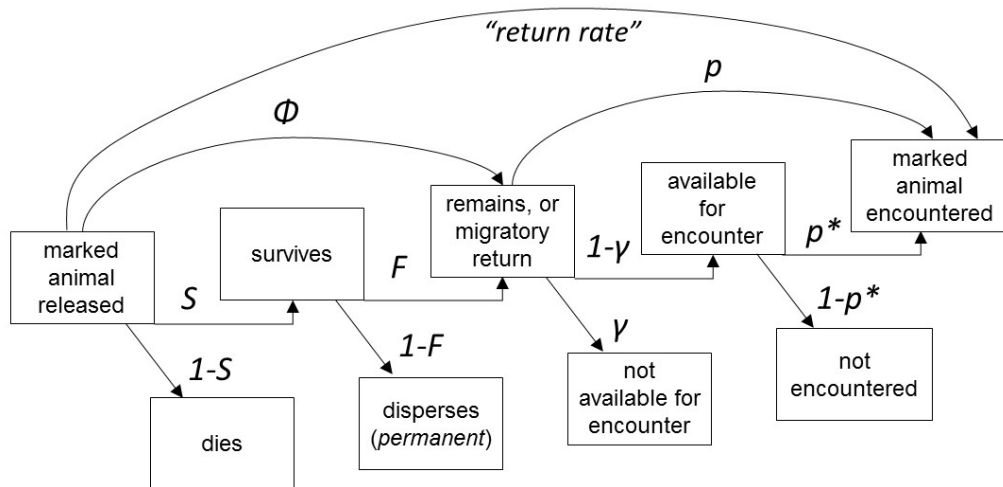
Figure 13.2: *Description of potential fates of a marked animal released in time i and potentially encountered again in time i+1 (after Kendall 2014).*

The robust design allows the estimation of many key parameters that describe dynamics of populations. It is nice to know population size and probability of survival, but in addition, knowledge of movement patterns is critical if we want to understand the complicated dynamics of a population.

Figure 13.2 shows all possible fates of an individual after it is marked. The animal is either going to survive or die during the next time period. If the animal survives, it may disperse from the super-population and never return—we label this **permanent emigration** and will distinguish it from **temporary emigration** later.

We previously discussed the permanent emigration problem when we described "apparent survival" in the Cormack-Jolly-Seber model structure. As we will soon see, the robust design is also prone to the apparent survival problem. That is, the basic model structure does not allow

us to estimate **fidelity** (F), or permanent emigration (1-F). We will continue to distinguish "true" survival, S, and apparent survival, ϕ. Thus, the robust design model structure (e.g., Figure 13.2) will use ϕ as a reference to apparent survival.

If an animal remains on the study area (or returns to the study area after seasonal migration), it is possible that it may be **available** for encounter or it may be unavailable. Animals that are unavailable might be below ground (as in our salamander example) or outside of the boundaries of the area subject to netting or trapping. Terminology of the robust design defines the probability of being "not available" as **temporary emigration**, γ.

Last, if an animal is available, it may either be encountered (p^*, trapped, sighted, etc.) or not. And, we note that the CJS-type encounter probability, p, is a function of both availability (1- γ) and true encounter probability. In parallel fashion to "apparent survival", we might refer to CJS-type encounter probabilities, p, as "apparent encounter probability" because they do not account for the lack of availability. And, we can re-define p for robust design as the probability of encounter for a member of the super-population during a given time period. Thus, we see the relationship:

$$p = (1\text{-}\gamma)p^*$$

Before we leave the discussion of the structure behind the robust design model, we should also point out the relationship to "return rate" (discussed in Chapter 10). The crude "return rate" is usually defined as the proportion of animals released that return to a study site—and this has a deep history in waterfowl management in particular. Here, we see that **return rate** is actually the product of apparent survival, ϕ, and apparent encounter probability, p (Figure 13.2). And, each of these parameters is a function of other parameters: ϕ is a function of survival, S, and fidelity, F, and p is a function of temporary immigration, γ, and encounter probability, p^*. So, return rate, R, can be defined as

$$R = SF(1\text{-}\gamma)p^*$$
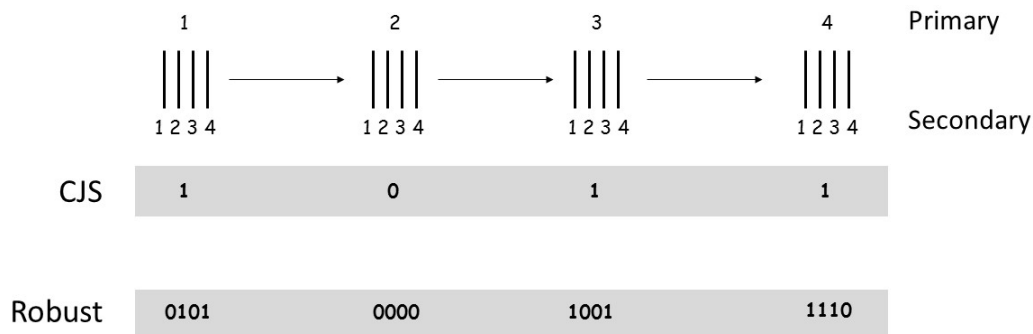
Figure 13.3: *The 'robust design' for mark-recapture including primary and secondary observation periods. Hypothetical capture histories are shown for all secondary periods under the robust design; capture success can be summarized into a simpler capture history for the primary periods used in Cormack-Jolly-Seber model structures.*

Model structure and capture histories

The basic robust design capture history, as described by Pollock (1982) is divided into "primary" and "secondary" sampling periods (Figure 13.3). The secondary sampling periods are considered closed—the time-frame, of course (see Chapter 8), depends on your study animal. But, it is within these secondary periods that we will be estimating closed-form population size, and we can obtain an estimate of N for each secondary period.

The primary periods are created by grouping the secondary periods (Figure 13.3). Modern software typically allow the length of time of the primary periods (the gaps between the secondary periods) to vary, although you must account for it in the model structure. The population is considered "open" between the primary periods.

You should be able to see that you could revert to a simple CJS-type analysis if you were to summarize the data from secondary periods into either encountered or not-encountered for each primary period. In this example, we would have a 4-occasion CJS-type capture history.

Similarly, we could conduct 4 separate closed-analyses of population size from our data. This is often done by biologists. However, Pollock's (1982) inspiration was to realize that more information could be obtained from a combined model structure—information about change in population size due to movements.
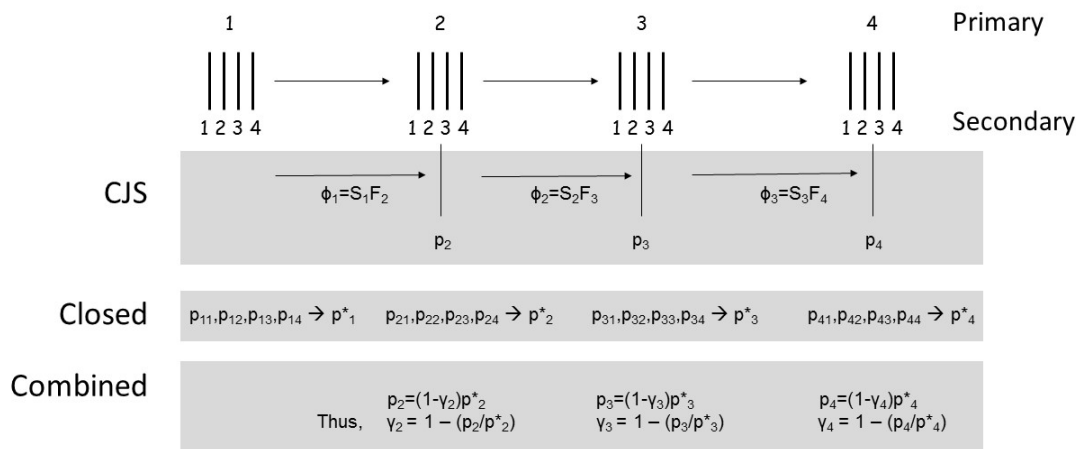

Figure 13.4: *Relationship of key parameters in the robust design (after Kendall 2014).*

What would this look like in the real world?

Imagine you had a 4-year study in which you conducted trapping of a lizard during only 5 days (of the same week in early July) each year. Certainly, we'd consider the lizard population to be closed during the 5-day period. Your primary and secondary periods might look like this:

Primary:	Year 1	Year 2	Year 3	Year 4	Year 5
Secondary:	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5

How does the robust design estimate temporary emigration? It may help to clearly lay out the parameters (Figure 13.4) that are estimated by the open (CJS-type) and closed estimators, so that we can see the manner in which the combined information is used to estimate availability ($1-\gamma$, the complement of temporary emigration).

First, we can see that in the closed population model structure, we can obtain a time-specific estimate for p. Next, we can calculate the overall probability for capture, p^*, during the first primary period (Figure 13.4) as:

$$p_1^* = 1-((1-p_{11})(1-p_{12})(1-p_{13})(1-p_{14}))$$

> NOTE: We use the subscripts, p_{ij}, to indicate the primary period, i, and the occasion of the secondary sampling period, j.

Our model structure also allows the estimation of the "apparent encounter probability", p_i, for the primary period from the open-portion of the model structure (CJS; Figure 13.4). And, we can see that if all animals were available ($\gamma =0$, or no temporary emigration), p_i should equal p^*. Because of this, we can estimate γ as a function of the relative difference of p and p^*:

$$\gamma_i = 1 - \left(\frac{p_i}{p_i^*} \right)$$

This is a fairly intuitive, and very important concept that is at the heart of the robust design. It's that simple!

Out, and back in?
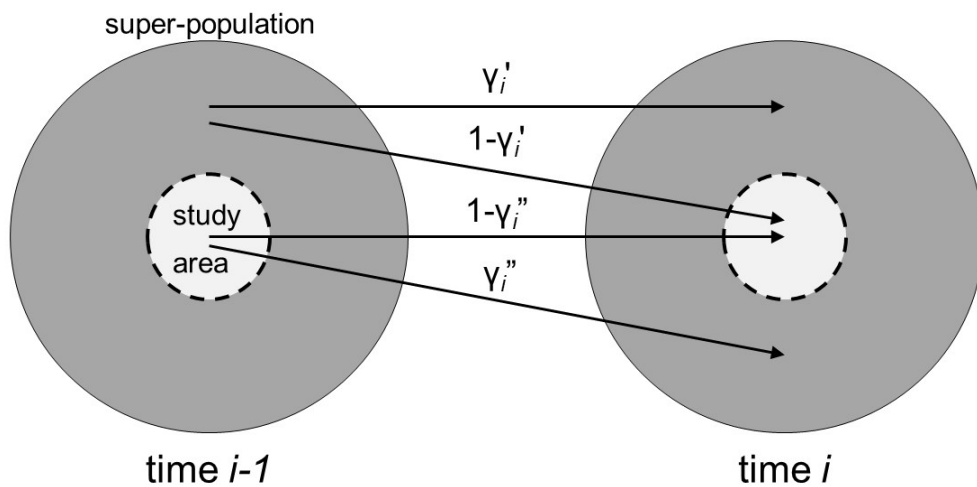
Figure 13.5: *Relationships between γ' and γ''. The super-population is represented by the larger circle, and the portion of the super-population available for encounter in the study area is represented by the inner circle. Animals temporarily immigrated with a probability of γ'' and are then not available for encounter. Animals return to the study area with probability of $1-\gamma'$.*

Kendall et al. (1997) provided one additional modification to the basic robust design, which forms the structure of the robust design model that is now typically available in most software packages. Kendall et al. (1997) realized that we could also be interested, over time, in what happens to animals once they leave—that is, we're interested in temporary emigration, but we are also interested in the rate of immigration back into our study site (Figure 13.5).

So, we now have a modification to the simple γ in our previous descriptions:

> **γ'**: for animals that are away from the study site (but in the super-population) the probability of remaining away from the study site.
>
> > Defined by Kendall et al. (1997) as the probability of being off the study area, unavailable for capture during primary trapping session i, given that the animal was not present on the study area during primary trapping session i-1, and survives to trapping session i.
>
> **γ''**: for animals on the study site, the probability of moving away from the study site.
>
> > Defined by Kendall et al. (1997) as the probability of being off the study area, unavailable for capture during primary trapping session i, given that the animal was present on the study area during primary trapping session i-1, and survives to trapping session i.

Movement models

Current analyses with the robust design model, subsequent to Kendall et al. (1997), typically consider three movement models as competing models of reality:

> **No movement (null):** $\gamma' = \gamma'' = 0$; all animals <u>available</u> for capture.
>
> **Random:** $\gamma' = \gamma''$; chance of being 'out' is the same for 'in' and 'out' animals (that is, every animal in the super-population has the same probability of being outside the study area during the next time period)
>
> **Markovian:** the animal's location in a previous time period affects γ' and γ''; chance of being 'out' during the next time period is different for 'in' and 'out' animals
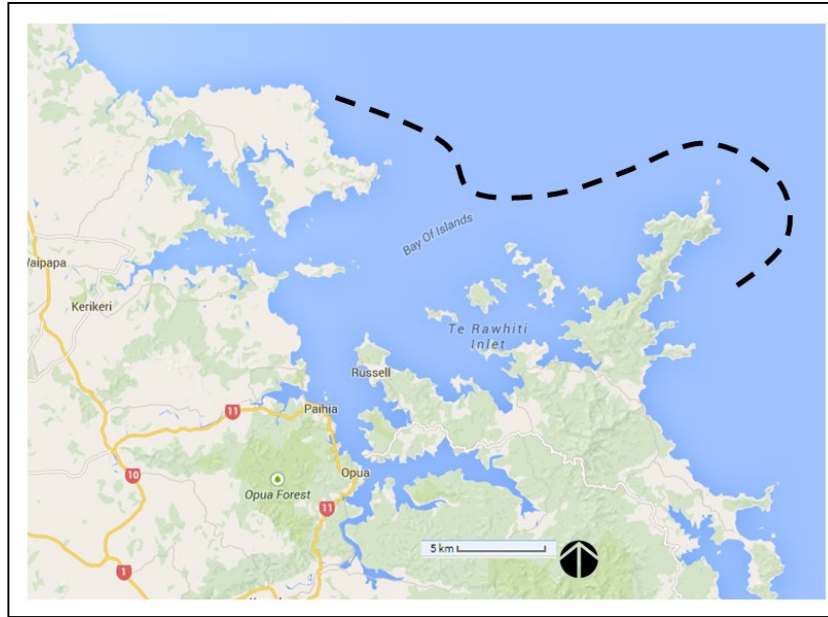
Figure 13.6: *The study site of Tezanos-Pinto et al. (2013) where dolphins were sampled in a bay from tourist vessels using photos. The dotted line shown represents a hypothetical extent of the sampling area, with the super-population extending beyond the sampling area.*

An example: bottlenose dolphins in a bay

Tezanos-Pinto et al. (2013) used the robust design to estimate population size of a coastal population of bottlenose dolphins (*Tursiops truncatus*) in northern New Zealand (Figure 13.6). The population was uniquely suited for the use of the robust design because:

- Super-populations of dolphins exists on the coasts of New Zealand. The populations do not emigrate permanently away from the coasts, and the various super-populations are known to be genetically isolated (animals do not move from one super-population to another).

- The specific super-population in the study was sampled through photos taken by tourism vessels and other vessels in the Bay of Islands. The super-population extends beyond the bay, so only a fraction of the super-population was available to be encountered during specific, short time intervals.

- Biologists had a goal of assessing population trends through time, which required estimates of N that were not biased by temporary emigration and immigration.

Figure 13.7: *Bottlenose dolphin (Tursiops truncatus) dorsal fins with distinctive natural markings. Copyrighted photo courtesy of Gabriela Tezanos-Pinto.*

The team of scientists scoured photos taken by scientists and by tourists in dolphin-watch vessels in the region. Dolphins were identified from photos by unique nicks and cuts on the trailing edge of the dorsal fin (Figure 13.7; Tezanos-Pinto et al. 2013). The biologists then used their estimates of gamma (γ) to estimate the size of the super-population (all coastal dolphins in the region; Figure 13.8).
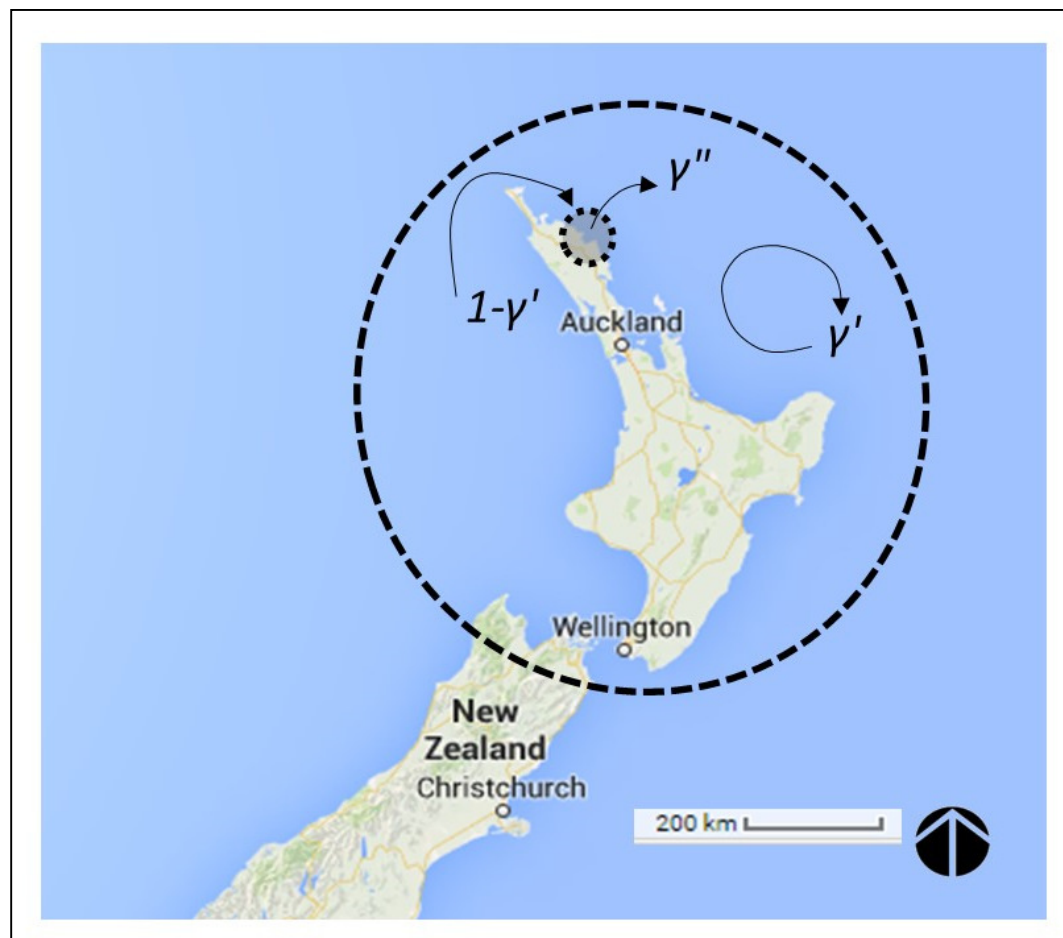
Figure 13.8: *The hypothetical borders of the study site (small circle) and super-population (large circle) of Tezanos-Pinto et al (2013) shown in the context of the robust design structure used in their analysis. Dolphins in the study area during a given time period could temporarily emigrate (γ") before the next time period. Animals outside the study area in a given time period could return to the study area (1- γ') or remain outside of the study area (γ').*

Estimating N_{sp}

As a final comment, let us examine the three basic movement models to describe how we could estimate N_{sp}—the size of the super-population. Again, we start with the fact that our robust design analysis will provide us with estimates of N for each primary sampling period. We will also receive, from the most general model, estimates of γ for each time period.

No movement. This is the easiest result to interpret, but we might predict it does not occur very often, because animals move—the basic reason for using the robust design model!.

> *If $\gamma' = \gamma''=0$, there is no temporary emigration, and animals remain on the study site. If animals move, they remain available during all primary sampling periods—even if they may move outside the sampling area.*

$$N_{sp} = N$$

Random movement. The estimation of N_{sp} is fairly straightforward under this model, which makes this model desirable for the convenience of the research biologist.

> *If $\gamma' = \gamma''$: A constant proportion of the super-population is found outside the sampling zone between periods. Indeed, we can view γ as the proportion of the population of the super-population found outside the study area.*

$$N_{sp} = \frac{N}{1-\gamma}$$

As an example, under random movement with $\gamma' = \gamma'' = 0.13$ and N(.)=100

$$N_{sp} = 100/0.87 = 114$$

The dolphin biologists in our example, above, found that the random emigration model was chosen over the no-movement model, and their estimates of γ ranged from $\gamma' = \gamma'' = 0.182$ to $\gamma' = \gamma'' = 0.820$ for each year in their study. Thus, approximately 20-80% of the dolphin super-population was not available for sampling during a given time period.

Markovian movement. This result tells you that your population has complicated dynamics of movement, which are not unexpected in nature. But, unfortunately, estimation of N_{sp} is difficult.

> **If $\gamma' \neq \gamma''$ and if $\gamma', \gamma'' \neq 0$:** estimating the size of the super-population is not straightforward. See Kendall et al. (1997) for more thoughts on this complicated dynamic.

Robust design model assumptions

As a combination of closed and open methods, robust design shares **assumptions** with simpler models discussed in previous chapters. In addition:

- the population is assumed closed to additions and deletions across each secondary sampling session within a primary period
- survival rates are assumed to be the same for all animals in the population, regardless of availability for capture
- permanent emigration out of the super-population influences the survival estimate (incorporated as mortality)
- individual marks are required; thus, marks must not fall off and must be recognizable

We caution you to consider the number of parameters estimated by this model structure as you design your study. Tezanos-Pinto et al. (2013, dolphin example above) reported 188 parameters (including, for example, 28 estimates for N and 27 estimates for γ for specific time periods during the study) estimated for their top-ranked model. One of the lower-ranked models had 531 parameters!

How can you determine how many animals to tag? Adequate sampling necessitates the ability to make a good guess (at a minimum) as to the movement dynamics of your population—if a large number of animals are not available for capture during a given time period, you will need to tag more animals to obtain a useful sample of captured (and recaptured) animals! We encourage you to review the literature—look at the sample sizes, and the precision of estimates used by others with similar research designs.

Conclusion

The typical use for robust design methods is for monitoring or sampling schemes with annual, or long-term periodic sampling and short bursts of *effort* during the year. The method is useful for situations in which the sampling gear may not 'cover' the entire area of interest (e.g., the super-population's full extent). And, the method is used by biologists who are interested in closed and open parameters (movement, survival, N).

We note, in this last chapter dedicated to mark-recapture analyses of survival, that many model structures are available for survival estimation. The method selected by the biologist should depend on the design of the project in terms of **types of tags** (e.g., is known fate possible?), the **timing of sampling** (e.g., is robust design possible?), and the **design of sampling** (e.g., is multistate possible?).

We hope that you now have a basic understanding of these model structures. At this point, we also hope that you can see the creative 'sport' that exists as you think about the parameters of interest and how to get the information you need to describe the dynamics of your population!

References

Bailey, L. L., T. R. Simons, and K. H. Pollock. 2004. Estimating detection probability parameters for Plethodon salamanders using the robust capture-recapture design. Journal of Wildlife Management 68: 1-13.

Kendall, W. L., J. D. Nichols, and J. E. Hines. 1997. Estimating temporary emigration using capture-recapture data with Pollock's robust design. Ecology 78: 563-578.

Pollock, K. H. 1982. A capture-recapture design robust to unequal probability of capture. The Journal of Wildlife Management 46: 752-757.

Tezanos-Pinto, G., R. Constantine, L. Brooks, J. A. Jackson, F. Mourão, S. Wells, and C. Scott Baker. 2013. Decline in local abundance of bottlenose dolphins (*Tursiops truncatus*) in the Bay of Islands, New Zealand. Marine Mammal Science 29: E390-E410.

For more information on topics in this chapter

Conroy, M. J., and J. P. Carroll. 2009. Quantitative Conservation of Vertebrates. Wiley-Blackwell: Sussex, UK.

Kendall, W. 2014. Chapter 15: The 'robust design' *In* Program MARK: a gentle introduction, 12th edition, Cooch, E. and G. White, eds. Online: http://www.phidot.org/software/mark/docs/book/pdf/chap15.pdf

Williams, B. K., J. D. Nichols, and M. J. Conroy. 2002. Analysis and management of animal populations. Academic Press, San Diego.

Citing this primer

Powell, L. A., and G. A. Gale. 2015. Estimation of Parameters for Animal Populations: a primer for the rest of us. Caught Napping Publications: Lincoln, NE.