Appendix S1: Rationale for using only sites uninfested at t and t-1 as target sites

This appendix addresses a question that looks simple but is not. When is a site that is observed to be uninfested at t and t-1 more likely to be truly uninfested at t than a site that is observed to be uninfested at t (regardless of its observed status at t-1)?

We use the following notation:

P_{tu}	:	Probability of a site to be truly uninfested at t-1
P_d	:	Probability of detecting an infestation, given the site is truly infested
P_e	:	Probability of a site to change from truly infested to truly uninfested within one time step
P_c	:	Probability of a site to change from truly uninfested to truly infested within one time step
u	:	uninfested (state of a site)
i	:	infested (state of a site)
S_1	:	set of sites observed uninfested at t
S_2	:	set of sites observed uninfested at t and t-1 ($S_2 \subseteq S_1$)

Possible combinations and their probabilities of true and observed states of all sites observed uninfested at *t*, assuming independence between detection of infestation and change of true state:

True state at		Observed state at		Probability for combination	Abbreviation for this probability
t-1	*t*	*t*-1	*t*		
u	*u*	*u*	*u*	$P_{tu}\,(1\text{-}P_c)$	*a*
u	*i*	*u*	*u*	$P_{tu}\,P_c\,(1\text{-}P_d)$	*b*
i	*u*	*i*	*u*	$(1\text{-}P_{tu})\,P_e\,P_d$	*c*
		u	*u*	$(1\text{-}P_{tu})\,P_e\,(1\text{-}P_d)$	*d*
i	*i*	*i*	*u*	$(1\text{-}P_{tu})\,(1\text{-}P_e)\,P_d\,(1\text{-}P_d)$	*e*
		u	*u*	$(1\text{-}P_{tu})\,(1\text{-}P_e)\,(1\text{-}P_d)^2$	*f*

The condition that the proportion of truly uninfested sites at survey *t* should be greater in set S_2 than in set S_1 can now be stated as:

$$\frac{a+d}{a+b+d+f} > \frac{a+c+d}{a+b+c+d+e+f}$$

Multiplying both sides by both denominators yields:

$$(a+d)(a+b+c+d+e+f) > (a+c+d)(a+b+d+f)$$

and after cancelling:

$$(a+d)(c+e) > c(a+b+d+f)$$

and after another round of cancelling:

$$e(a+d) > c(b+f)$$

Replacing these probabilities by the combinations of parameters that they stand for yields:

$$(1\text{-}P_{tu})\,(1\text{-}P_e)\,P_d\,(1\text{-}P_d)\,[P_{tu}\,(1\text{-}P_c)+(1\text{-}P_{tu})\,P_e\,(1\text{-}P_d)]$$

$$> (1\text{-}P_{tu})\,P_e\,P_d\,[P_{tu}\,P_c\,(1\text{-}P_d)+(1\text{-}P_{tu})\,(1\text{-}P_e)\,(1\text{-}P_d)^2]$$

Subtracting the left side of the above inequality from both sides and expanding the products yields:

$$P_d P_{tu} - P_c P_d P_{tu} - P_d^2 P_{tu} + P_c P_d^2 P_{tu} - P_d P_e P_{tu} + P_d^2 P_e P_{tu} - P_d P_{tu}^2 + P_c P_d P_{tu}^2 + P_d^2 P_{tu}^2 - P_c P_d^2 P_{tu}^2 + P_d P_e P_{tu}^2 - P_d^2 P_e P_{tu}^2 > 0$$

and after cancelling and collecting terms of P_d and P_{tu}:

$$(1\text{-}P_d)\,P_d\,(1\text{-}P_e\text{-}P_c)\,(1\text{-}P_{tu})\,P_{tu} > 0 \qquad \Leftrightarrow \qquad P_e + P_c < 1 \qquad (1)$$

assuming both P_d and P_{tu} are nonzero and not equal to 1.

Sites observed uninfested at t-1 and t are therefore less likely to be infested at t than sites observed uninfested at t (without further restriction) if the sum of the *true* extinction and establishment probabilities is less than one.

The conditional probability, P_{oui}, of a site to change from *observed* uninfested at t-1 to *observed* infested at t, given it is *observed* uninfested at t-1 equals:

$$P_{oui} = \frac{P_{tu}P_cP_d + (1-P_{tu})(1-P_d)(1-P_e)P_d}{P_{tu} + (1-P_{tu})(1-P_d)}$$

Similarly the conditional probability, P_{oiu}, of a site to change from *observed* infested at t-1 to *observed* uninfested at t given it is *observed* infested at t-1 equals:

$$P_{oiu} = P_e + (1-P_d)(1-P_e)$$

Let $\Sigma_{true} = P_e + P_c$ and $\Sigma_{obs} = P_{oui} + P_{oiu}$. We will now show that if $\Sigma_{true} > 1$ then $\Sigma_{obs} \geq 1$. Define $D = \Sigma_{obs} - \Sigma_{true}$. Substituting in the definitions,

$$D = \Sigma_{obs} - \Sigma_{true} = \frac{P_{tu}P_cP_d + (1-P_{tu})(1-P_d)(1-P_e)P_d}{P_{tu} + (1-P_{tu})(1-P_d)} + (1-P_d)(1-P_e) - P_c$$

and using *Mathematica 4.2* to simplify gives

$$D = (1 - P_c - P_e)\frac{1 - P_d}{1 - P_d(1 - P_{tu})} = D'\frac{1 - P_d}{1 - P_d(1 - P_{tu})}$$

where D' denotes $1 - P_e - P_c = 1 - \Sigma_{true}$. Again assuming both P_d and P_{tu} are nonzero and not equal to 1,

$$0 \le \frac{1 - P_d}{1 - P_d(1 - P_{tu})} \le 1 \ ,$$

hence

$|D| \le |D'|$. If $\Sigma_{true} > 1$, then by definition $D < 0$ and $D' < 0$, and therefore

$\Sigma_{true} = |D| + \Sigma_{obs}$ and $\Sigma_{true} = 1 + |D'|$ and therefore

$$\Sigma_{obs} = 1 + (|D'| - |D|) \ge 1 .$$

It is therefore not possible to have $\Sigma_{true} = P_e + P_c > 1$ and $\Sigma_{obs} = P_{oui} + P_{oiu} < 1$. However, because of sampling variability, the observed transition frequencies used to estimate Σ_{obs} could still be less than one even though $\Sigma_{true} > 1$. Nevertheless, we used the sum of the observed transition frequencies as an estimate for Σ_{obs}.

The average extinction probability, estimated as the total number of transitions from observed infested at *t* to observed uninfested at *t*+1 divided by the total number of sites observed infested equals 0.62. The average establishment probability, estimated as the total number of transitions from observed uninfested at *t* to observed infested at *t*+1 divided by the total number of sites observed uninfested equals 0.09. The sum of these probabilities equals 0.71 and is therefore less than one. Hence including only sites that were uninfested at *t* and *t*-1 increases the chance of counting sites truly uninfested at *t*.