

3 Supporting Information S2 Text

4 E.H. Bussell & N.J. Cunniffe

5 **S2 Approximate model and fitting details**

6 **S2.1 Approximate model**

7 The approximate model used to optimise the control via OCT is non-spatial, but in all other respects is identical
8 to the simulation model. The demographic parameters, such as age transition rates and birth rates, can be used
9 directly from the simulation model since the demographic dynamics are the same. The infection rates in the
10 non-spatial approximate model cannot be lifted from the simulation model, since the approximate model now
11 assumes infection comes from all infected hosts in the stand rather than through the spore deposition kernel.
12 The force of infection terms in the approximate model are therefore:

$$13 \quad \tilde{\Lambda}_{1,i} = \tilde{\beta}_{1,i} \sum_{j=1}^4 I_{1,j} + \tilde{\beta}_{12} I_2 \quad (1a)$$

$$14 \quad \tilde{\Lambda}_2 = \tilde{\beta}_{21} \sum_{j=1}^4 I_{1,j} + \tilde{\beta}_2 I_2 \quad (1b)$$

15
16 where $\tilde{\beta}$ indicate infection parameters that need to be fitted to the simulation model, and the dependence on
17 cell (x) of the states has been dropped since the model is non-spatial.

18 **S2.2 Fitting infection rates**

The seven infection rates $\tilde{\beta}$ are the only parameters that must be fitted to the simulation model. We use the
method of least squares to match the simulation and approximate models. To fit the parameters, the simulation
model is used to run a single trajectory with no control interventions. For the fitting process we use the same
initial conditions as described in the main text, with infection seeded in the centre of a 20 by 20 grid of cells.
The disease progress curves of this simulation run are then used as the baseline for fitting the approximate
model. For a trial set of $\tilde{\beta}$ parameters and an approximate model trajectory, we calculate the sum of squares as
the sum of squared deviations between the simulation and approximate disease progress curves for each age
class of tanoak, and for bay, at time points throughout the trajectory. Redwood is not included as it cannot be
infected. The $\tilde{\beta}$ parameters are then optimised by minimising this total summed squared error (SSE). For a set
of time points t_i , and where approximate model states are signified with a tilde, the equation for SSE is given
by:

$$\text{SSE} = \sum_i \left[\sum_{j=1}^4 (I_{1,j}(t_i) - \tilde{I}_{1,j}(t_i))^2 + (I_2(t_i) - \tilde{I}_2(t_i))^2 \right] \quad (2)$$

where the dependence on cell in the simulation terms has been dropped to indicate an average over all cells in the landscape, for example:

$$I_{1,j}(t) = \frac{\sum_x I_{1,j,x}(t)}{N_{\text{cells}}} . \quad (3)$$

19 An average is used so that the approximate model tracks stem density in the same units as the simulation
20 model.

21 In the simulation model, infectious pressure is dominated by sporulation from bay laurel. This makes
22 estimation of all ‘within-tanoak’ infection rates ($\beta_{1,i}$) difficult, as from the simulation data they are individually
23 unidentifiable. We therefore use a two stage fitting process. For the first stage, all infection rates in the
24 simulation model related to bay (β_2 , β_{12} , and β_{21}) are set to zero. This makes bay epidemiologically inactive,
25 but maintains the same demographic dynamics. The simulation model is run using these parameters, and the
26 SSE is minimised to find the within tanoak-infection rates.

27 In the second fitting stage all infection rates are fitted, with bay epidemiologically active again. The within-
28 tanoak rates relative to $\beta_{1,1}$ from the first stage are used as a constraint to ensure the identifiability of these
29 rates. This means a single within-tanoak rate is fitted in stage 2, with all other within-tanoak rates fixed relative
30 to this using the results from stage 1. This stage also fits the bay infection rate, and the cross-species infection
31 rates.

32 As can be seen in Figure 1B–1C in the main text, despite lacking any spatial component, the approximate
33 model can very closely capture the uncontrolled dynamics of the simulation model. However, the approximate
34 model should also fit as accurately as possible when control strategies are introduced. In Figure 1, the fit of the
35 approximate model is tested under constant control strategies using fixed control rates. It is clear that roguing
36 at the same rate is more effective in the approximate model. This is because of the difference in mixing between
37 the approximate and simulation models. The effect is small for thinning and protecting, but the same level of
38 roguing in the approximate and simulation models gives very different dynamics.

39 S2.2.1 Empirical parameterisation of control

40 Roguing is less effective in the simulation model because of an imposed spatial structure in the non-spatial
41 control strategy. Roguing in the simulation model removes infected hosts from the core and edge of the
42 spreading epidemic. Removal of hosts from the core has little effect on the rate of epidemic spread, since they
43 are not near the wavefront. In the non-spatial model however, all hosts are well-mixed, so removal of infected
44 hosts has a larger effect.

45 As a simple correction for this difference in roguing effectiveness, we investigate a simple scaling of the
46 roguing rate in the approximate model. To test the plausibility of a single scaling rate for all approximate
47 roguing controls, both models are run with constant roguing strategies. Roguing of small tanoak, large
48 tanoak and bay are all set to occur at the same rate for the whole simulation, and this rate is varied between
49 simulations. In the approximate model, the control rate is scaled by a single parameter which is also varied,
50 and we analyse the difference in the final number of healthy large tanoak after 100 years between the simulation
51 and approximation (Figure 2). To minimise the deviation across a range of control rates, the value of the scaling
52 parameter is optimised. We optimise the deviation in the number of large healthy tanoak, since this is the
53 primary objective of the control and must therefore be captured as accurately as possible. The optimisation
54 minimised the sum of squared errors (SSE) over the range of control rates.

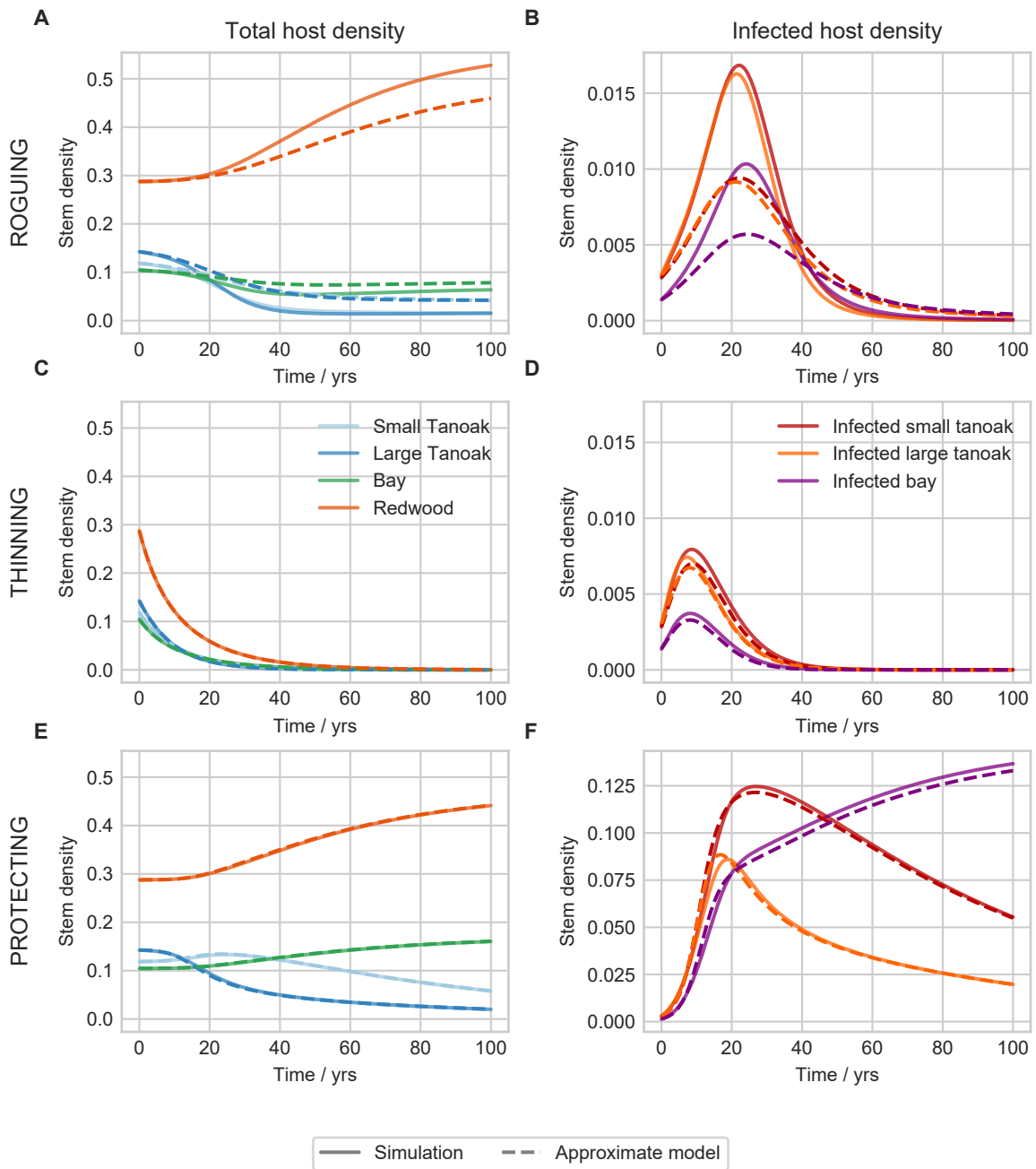


Figure 1: Testing of approximate model under constant control strategies. The rows from top to bottom show dynamics under constant roguing, thinning, and protecting strategies. The left plots show overall host dynamics, and the right plots show the infected host dynamics. The roguing and protecting strategies control at the maximum rates from Table 1 in S1 Text, whereas the thinning strategy controls at 10% of the maximum rate. The approximate model fits well under constant thinning and protecting strategies, but less well under a constant roguing strategy.

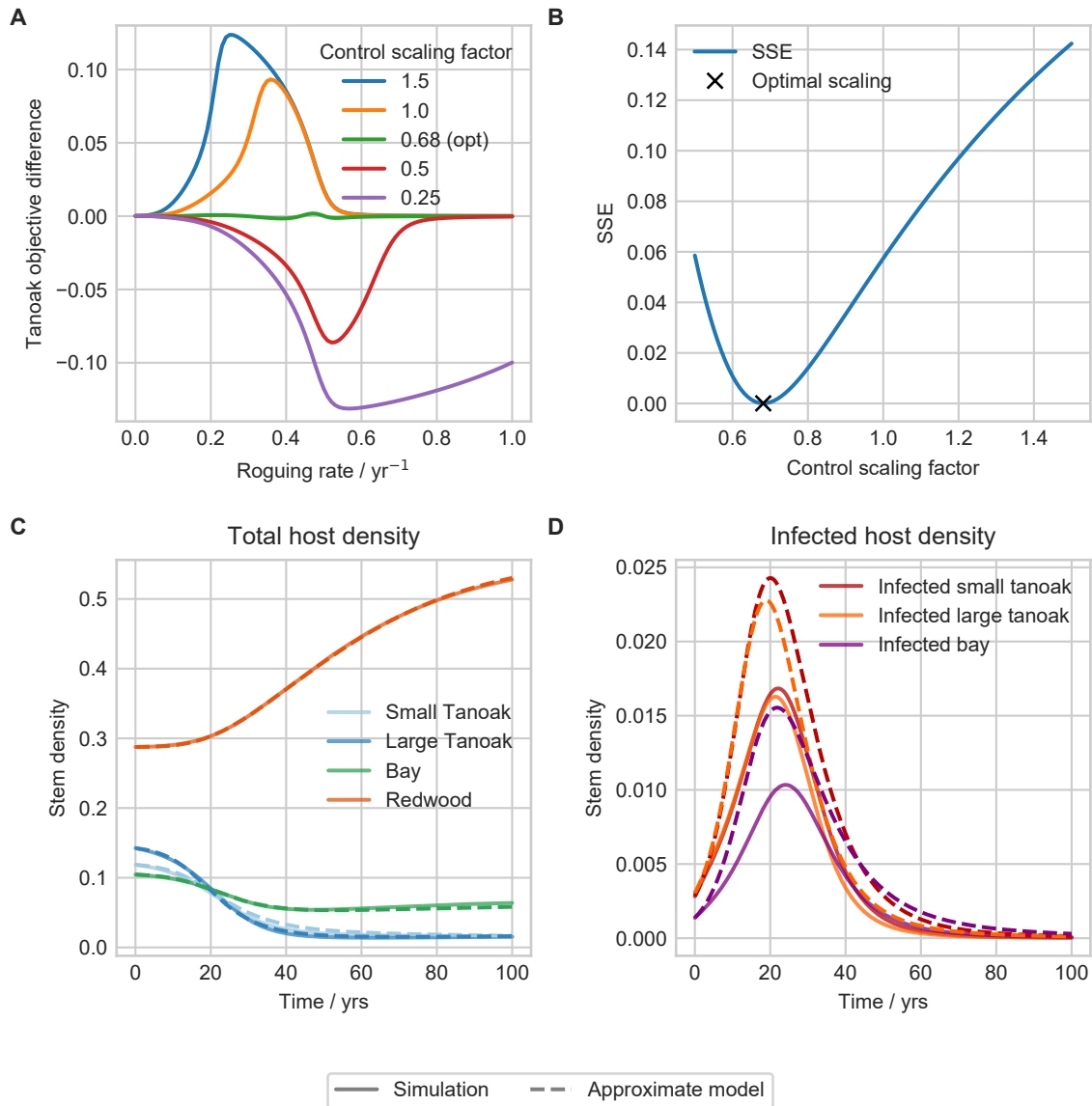


Figure 2: Empirical scaling of approximate model control rates to match simulation output. A: the difference in the final number of large healthy tanoak as a function of the constant roguing rate, for a number of control rate scaling factors. The optimal value minimises the sum of squared errors (SSE) over all rates, as shown in B, and is found to be 0.682 (3 s.f.). C and D show the dynamics with the newly scaled approximate control rates, under a constant roguing strategy. The approximate model now fits the overall host dynamics well, but to do this slightly overestimates the level of infection.

55 The results in Figure 2 show that a single scaling factor can largely eliminate the deviation under constant
 56 roguing strategies. We do not expect this scaling to ensure that the approximate model is always closely aligned
 57 to the simulation model, particularly once control strategies are time-varying. The approximate model simply
 58 cannot capture the heterogeneities in host mixing. This scaling does however go some way to ensuring that
 59 control strategies from the approximate model perform well on the simulation model. Figure 3 shows the
 60 open-loop and MPC control strategies and host dynamics when the roguing rate is not rescaled. Whilst OCT
 61 finds similar strategies, the error in the approximate model under open-loop is larger than when the roguing
 62 rate is scaled.

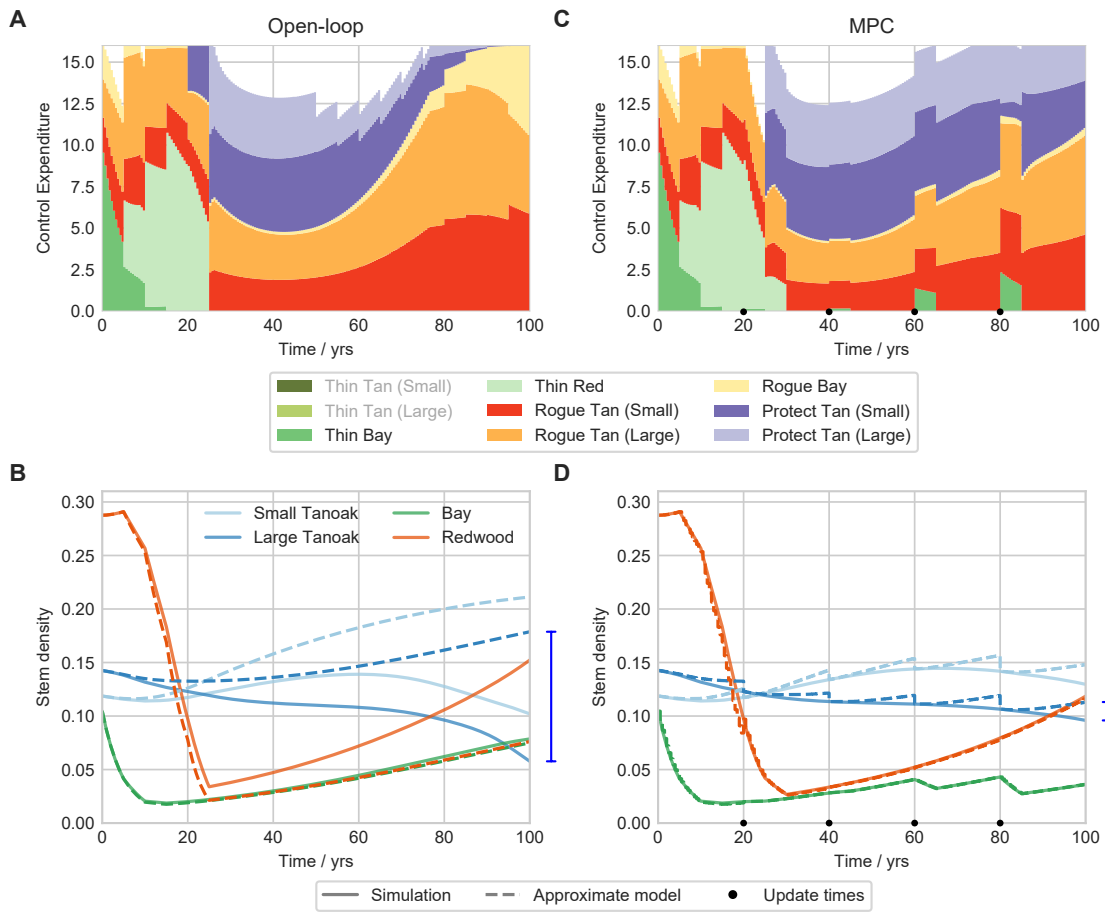


Figure 3: Control strategies and host dynamics under open-loop (A and B) and MPC (C and D) when the roguing rate is not scaled. Similar strategies are found but the error in the approximate model (blue bar) is much larger under open-loop than with the rescaling (compare with Figure 2 in the main text).