DEPARTMENT OF Plant Sciences

Department for Environment Food & Rural Affairs

Agri FoRwArdS EPSRC Centre for Doctoral Training in Agri-Food Robotics

Baselines for prioritization of epidemic control

Rachel Russell (University of Cambridge, Plant Sciences / Theoretical and Computational Epidemiology)
Nik Cunniffe (University of Cambridge, Plant Sciences / Theoretical and Computational Epidemiology)

The resources available to stakeholders managing disease epidemics across plant, animal and human systems are limited. Optimization techniques can guide stakeholders on where and when these resources should be allocated to maximize their impact but different optimisation approaches are demonstrated on different systems which makes it difficult determine the current state of the art for any new system of interest. This limits both progress in the field and applicability for managers. We propose that a wider range of simple heuristic controls should be considered as baselines for evaluation of more complex control approaches and the validity of the optimisation assumptions should be tested when feasible. The utility of simple baselines is demonstrated with an example evaluating continuous optimal control on a stochastic metapopulation model of geographical spread of a plant disease.

Baselines are important

"Baseline: A minimum or starting point used for comparisons." [Oxford Languages]

If we want to demonstrate that a new optimisation method represents genuine progress in the field, direct comparisons to older or simpler algorithms are essential. In computational fields such as machine learning, a combination of standard benchmarks, reference baseline implementations of common algorithms and open source code have enabled rapid progress and provided flagship results which encourage adoption [1][2]. In comparison, many papers optimising epidemic control present no baseline results or use a very simple baseline.

Without baselines

Confused policy maker...

Which approaches should I consider for my epidemic control problem

Confused new PhD student...

With baselines

I can see technique X does better than a simple heuristic for a problem like mine.

Technique Y is clearly the state of the art for solving this problem but I can do better!

Selecting baselines is difficult

The latest baseline for epidemic control is difficult to determine because of the diversity of problems:

- Different models
- Different parameterisations (pathosystem and costs)
- Different control actions modelled
- Stochastic vs deterministic assumptions
- Different measures of success
- Different assumptions about information available

Many analytical solutions can be shown to be theoretically optimal but this only holds when assumptions are met for the underlying model and when simplifying approximations are sufficiently accurate. Do the solutions still beat the baseline when subject to...

- Stochasticity
- Realistic control resolution (spatial / temporal)
- Imperfect control (delays, efficacy)
- Imperfect information (partial inspection, cryptic infection)
- Imperfect parameterisations

Which baselines should I consider?

Improving baseline quality

No baseline – appropriate for early theoretical work on abstract optimisations where assumptions / limitations of optimised model are clearly stated

No control – relevant when optimisation goal includes cost of control explicitly. Unlikely to be a strong baseline.

~50% of papers here (informal pilot review)

Random control - Randomisation of priorities or proportions of budget per control. Consider randomising once or randomising at intervals. [3]

Equal control – split budget equally between subpopulations / treatments. [4]

"What really happened" – relevant if case study is using past data or for an ongoing outbreak. Must be consistent with modelling assumptions. For example, not appropriate to test radius control on a metapopulation model or test an eradication focussed control on a model with cost assumptions which prevent eradication. [5]

~30% papers here

Improving baseline quality

Heuristics

Allocation over metapopulation or stratified population
Combinations of….

Characterisation metric:

Epidemic status (e.g. number of infected hosts, proportion of susceptible hosts)

Network based (e.g. centrality, degree)

Demographic based (e.g. no of hosts, host age)

Allocation procedure:

Act proportionally to metric (suitable where a balanced approach appropriate e.g. inspection, thinning)

Prioritisation list (do maximum feasible level of control on highest value and move down list until budget is exhausted) [6]

Knapsack allocation (completing two small tasks better than partially completing one big task) [7]

Propose that everyone can feasibly be here (for at least pilot level testing)

Allocation over raster grid (aka – graphical heuristics)
Examples: Controlling on wavefronts [8], radius control around infected hosts [9], "firebreaks" / host free barriers.

Allocation over time
Examples – switching control on and off at fixed times, switching control on and off at particular thresholds (with and without hysteresis) [10]

Sweeps - Full optimisation across all parameters of a control strategy are often not feasible but single value sweeps may be more tractable. For example control radii or switching thresholds. Can potentially optimise levels of one control leaving the other at a default value or unused. [9]

~20% papers here

Static / immediate optimisations
e.g. Minimising R0, immediate infection pressure, connectivity metrics [11]

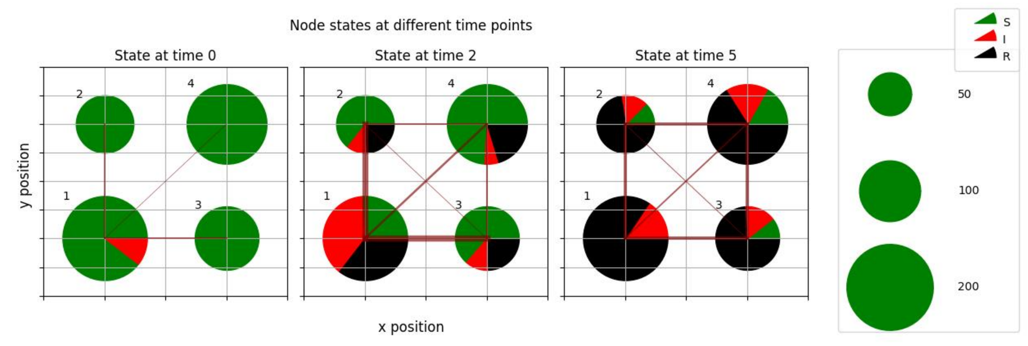
We should aim to be here

"The last best thing" – most convincing candidates from literature review. Can be challenging if epidemic model used for assessment is not identical or code is not available.

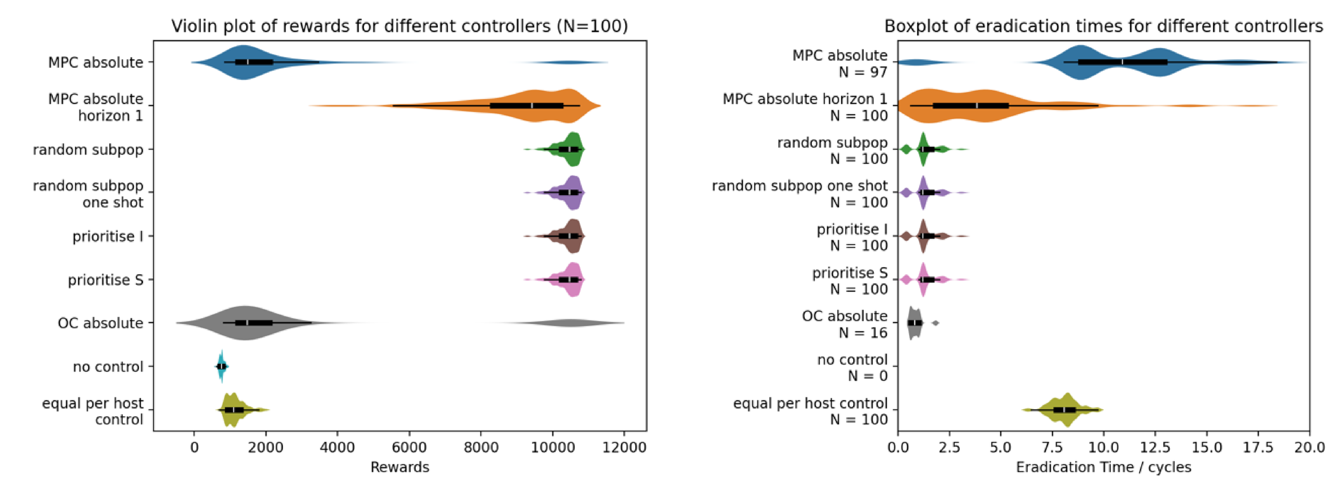
Ablation study for more complex algorithms – are all the parts of your algorithm important? How is the performance affected if you remove parts? [12]

Routine standard for other fields (e.g. ML)

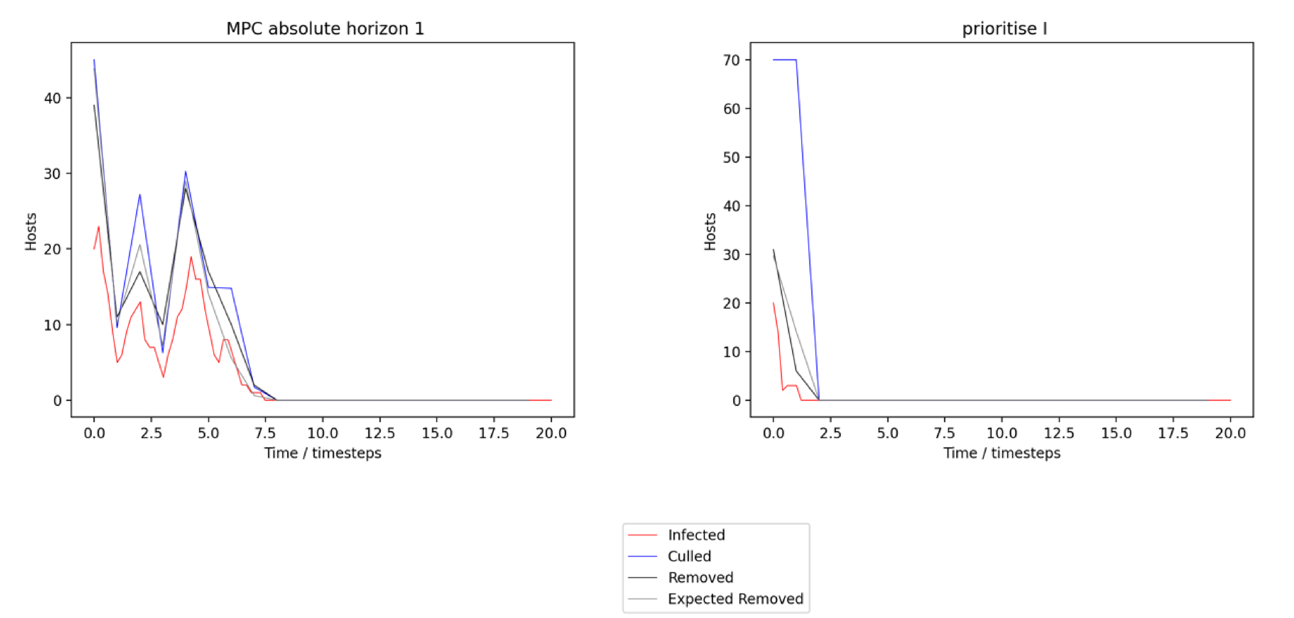
Case study – deterministic optimal control



Target system is a simple stochastic metapopulation SIR model with spatial spread across nodes in an nxn grid (shown for 2x2). The control actions to be optimised are the rates roguing and thinning and there is a limited budget per year. Budget was set to be generous such that eradication was possible. Management goal was to maximise the integral number of susceptible hosts over a fixed horizon..

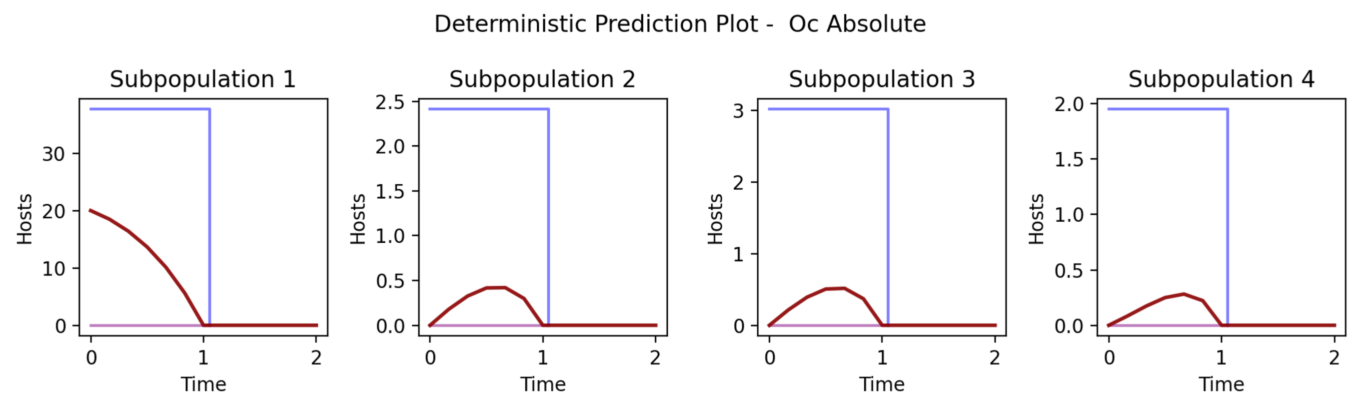


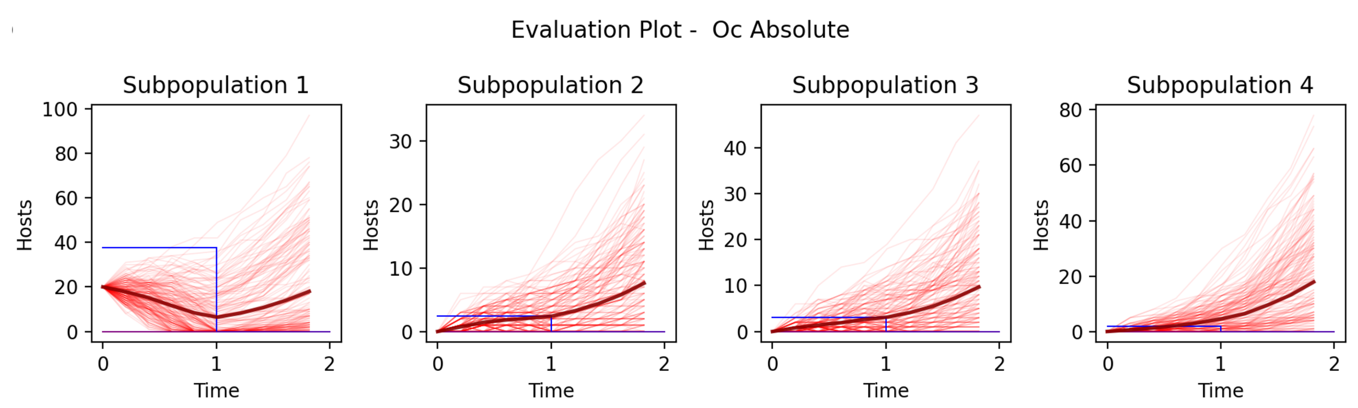
When eradication is feasible, the continuous optimal control can be outperformed by simple prioritisation-based baselines.



The optimal control is constrained by the assumptions of the continuous model to only remove the number of infected hosts that it expects.

In a stochastic system, this can always be outperformed by a strategy which spends all of the available budget.





Comparing the behaviour expected by the optimal control algorithm to the behaviour of the stochastic system when the optimised controls are applied. The expected behaviour removes all infected hosts and no further control is required. In the stochastic system, removing "just enough" leaves infected hosts in a significant proportion of simulations

References

[1] Boulesteix et al, PLoS One 8, e61562 (2013)
[2] Henderson et al, arXiv:1709.06560 [cs, stat] (2019)
[3] Andersen et al, Phytopathology 109, 1519 (2019)
[4] Cairns, Mathematical Medicine and Biology 6, 137 (1989)
[5] Alam-Khan et al, Scientific Programming 2020, e7627290 (2020)
[6] Cristancho-Fajardo et al, J. R. Soc. Interface 19, 20210744 (2022)
[7] Dangerfield et al, Bulletin of Mathematical Biology 81, 1731 (2019)
[8] Cunniffe et al, Proceedings. National Academy of Sciences 113, 5640 (2016)
[9] Cunniffe et al, PLoS Computational Biology 11, e1004211 (2015)
[10] Carli et al, Annual Reviews in Control 50, 373 (2020)
[11] Bussell et al, Philos. Trans. R. Soc. B, 374, 20180284 (2019)
[12] Meirom et al, Proceedings. ICML 38, 7565 (2021)

Engineering and Physical Sciences Research Council

UNIVERSITY OF CAMBRIDGE DEPARTMENT OF ENGINEERING

University of East Anglia

UNIVERSITY OF LINCOLN