

Bayesian modelling of wildlife disease across ecological scales

Anna Bush¹ and Dave Hodgson¹

¹University of Exeter

March 2, 2022

Abstract

Bayesian inference is a tool for treating uncertainty, combining data and prior information from multiple sources and formats into updateable models, often of considerable complexity. Attention is increasingly being paid to the use of Bayesian inference in the study of host-pathogen systems, in which complex networks of between-individual interactions operate across a hierarchy of ecological strata. Despite growing interest, the adoption of hierarchical systems-models by ecologists remains rudimentary. Bayesian inference has been applied to wildlife disease networks at a population level, and to epidemiological diagnostic regimes at an individual level, but there exist very few attempts to integrate models and data that link individual-, group-, population-, landscape- and assemblage-levels of wildlife disease systems. Furthermore, the use of Bayesian techniques at an individual level has been limited, yet this is vital for uncovering the fine-scale interactions and latent variables typical of disease networks. This review explores the use of Bayesian hierarchical models in the study of host-pathogen systems, identifying the future research required to achieve the desired “whole-system” approach. We argue that the complexities and uncertainties underlying disease processes are best described within a Bayesian framework, contending that although the infrastructure to craft complex Bayesian hierarchical models exists, the actual application of these methods is limited within wildlife disease research.

Introduction

Host-pathogen systems are characterised by the complex networks of interactions between an infectious agent and its host species. Examples of such host-pathogen systems are bovine tuberculosis (bTB), Severe Acute Respiratory Syndrome coronavirus (SARS-CoV) or the yet-unidentified wildlife host of SARS-CoV-2—the causative agent of COVID-19. Given the severity and economic impact of these and similar diseases, improved host-pathogen modelling capabilities will be of undoubted value to epidemiologists and ecologists working towards effective disease controls.

In comparison to their human counterparts, wildlife disease models contend with a data-poor environment. Animals tend to be hard to track and or trap, and infection-states can be hard to infer due to poor quality diagnostic tests. Accordingly, a common purpose of wildlife disease models is to describe known—or uncover unknown—latent relationships between the identified factors that make up a disease system, to better understand how disease spreads. Our synthesis is restricted to statistical rather than analytical models of disease processes because the parameters that may describe these processes vary stochastically, as well as in time and space; and are rarely, if ever, known.

Studying disease at a systems scale involves recognising the hierarchy of interacting levels over which disease dynamics persist—termed in this review as the ecological hierarchy (Figure 1) — and requires the adoption of hierarchical models. Hierarchical models start from the premise that a hierarchy of scale exists across ecological systems, and explore the nested relationships between differently scaled variables through sub-models, which link together to form a full model. For example, major progress in eliminating the *Sarcoptes scabiei* mite from bare-nosed wombat populations was facilitated by considering the disease statuses of wombat burrows at the metapopulation level as well as of the individual wombats; and consequently, both the burrows and the wombats were modelled as hosts. Hierarchical modelling has also enabled a database

of bat hibernation roost surveys to be analysed across time, space, and five species, to determine the latent disease severity of *Pseudogymnoascus destructans* infections—the causative agent of white-nose syndrome—at species and regional scales in North America .

We also draw attention to a second type of hierarchy typical of data associated with wildlife: the statistical hierarchy (Table 1). This hierarchy also exists across ecological levels and scales, and represents the hidden network of latent variables we wish to infer—such as being infected or being dead—that can only be measured by proxy, for instance, by live trapping, or by analysing the results of imperfect diagnostic tests. The statistical hierarchy is therefore a state-space representation of parameters that, when inferred, will help us to understand wildlife disease either directly or indirectly. We suggest that a “whole-system” approach to studying wildlife disease is essential because most mechanisms of wildlife disease transmission covary with other ecological parameters, are not fully understood, and are impossible to measure directly. For example, identifies substantial gaps in our knowledge of Chronic Wasting Disease (CWD) ecology, such as its unidentified reservoir species, and the biogeography of CWD transmission: a whole-system approach would be directly applicable to this problem.

Parameterising the double-hierarchy of whole-system disease models (ecological and statistical) is beyond the toolbox of classical, frequentist or maximum likelihood statistical techniques. For systems in which complexities prevent the definition of likelihood functions for observed data, analysts might look to Approximate Bayesian Computation or machine learning techniques to guide understanding of the system. Here, though, we examine how Bayesian inference is under-utilised in disease ecology yet is a powerful tool capable of quantifying unknown and or unobserved latent variables within ecological networks, by treating them as random variables described by probability distributions . Disease status and mortality status are latent variables largely unique to wildlife host-pathogen systems, and therefore are the variables we draw particular attention to throughout this review.

Improving the capability of statistical models to describe entire ecological systems is an important and desired advance in disease ecology, particularly since understanding wildlife health is critical to its management . Bayesian modelling is a pivotal and statistically rigorous approach with which to improve both the parameterisation of infectious disease processes , and learn the strength of whole-system drivers such as disease incidence. Noticeably, the application of Bayesian inference to disease modelling has been propelled by research relating to the current COVID-19 pandemic . In this synthesis, we refer to a whole-system model as one able to describe as many aspects of an ecological network as possible, across a hierarchy of ecological scales (Fig. 1).

We describe how Bayesian methods can help ecologists move towards a whole-system approach to studying disease networks. Our recommendations are unlikely to surprise statistical epidemiologists, but our finding is that Bayesian methods remain under-used in wildlife disease research. The review is structured sequentially as follows: first, we examine why Bayesian inference should be used to model wildlife disease; second, we show how Bayesian modelling has informed research into wildlife reservoirs of bTB; third, we discuss why Bayesian hierarchical modelling is essential to a whole-system approach; and finally, we clarify the importance of latent variables and individual heterogeneities to a whole-system model of wildlife disease. In these latter sections, we include a survey of the literature to reveal that Bayesian approaches to the modelling of wildlife disease are (a) relatively scarce and (b) tend to infer only limited subsets of a whole-system model.

Why use Bayesian inference to model wildlife disease?

A Bayesian model can be loosely defined as any model deriving its inference from a posterior probability distribution, which is acquired from a prior probability distribution and its associated likelihoods, using Bayes’ theorem and *any* available data . Bayesian inference allows the explicit modelling of both observed and unobserved *a priori* data, which act as model assumptions. Consequently, a disease ecologist may be able to combine all known ecological information relating to a host species into a single model. This method has already influenced our understanding of disease risks from invasive species , the potential for disease transmission , and vulnerabilities within livestock systems to foot and mouth disease . Overall, Bayesian

models are highly capable of representing multi-scale studies, making them critical tools for disease ecologists trying to understand host-pathogen systems.

We believe that ecologists should understand Bayesian modelling approaches, in order to fully realise and explain the causal relationships and interacting factors that make up disease systems . This need is being reflected in the scientific community by authors such as , who have recently updated the WAMBS (when to Worry and how to Avoid the Misuse of Bayesian Statistics) checklist originally proposed in 2017. The primary goal of statistical epidemiology is to infer the parameters most relevant to the understanding and management of epidemics, particularly infection prevalence, force and severity. The contemporary task of disease ecologists is to understand and differentiate among interactions and relationships within a complex host-pathogen system , yet there are multiple complications. For instance, when modelling disease systems, network complexity is known to add to network fragility , largely due to increasingly unpredictable ecological responses to perturbations . One such example is social perturbation, i.e. individual dispersal in response to interference, as observed during badger culling . A further complication is that the causative pathogens themselves may instigate unknown and potentially bi-directional behavioural alterations affecting transmission . Constructing a realistic host-pathogen network, including these fine-scale interactions such as individual behaviours, remains a key challenge to the development of a whole-system model.

The many benefits of Bayesian approaches to inference, and the Monte-Carlo Markov Chain algorithms usually used to implement them, are described elsewhere . However in the context of wildlife-disease modelling, the benefits of adopting a Bayesian approach include but are not limited to: the inference of truth given data; the ability to introduce informative prior information when available; the flexibility to describe a hierarchy of states, processes and their noise in a single model; the ability to infer latent variables and parameters; the ability to integrate across multiple sources of data and multiple statistical processes; and the flexibility to work with a wider-than-usual range of likelihood functions . In contrast, the costs of adopting Bayesian methodologies include: the learning of new statistical concepts and software; the dropping of ingrained allegiances to tests of significance or information criteria ; the computational expense of running long, iterative chains of likelihood calculations; and the lack of consensus on how to judge the importance of rival models . Recent advances in computation, methodology, education, and software are already helping to minimise these apparent costs.

Bayesian inference is particularly useful to disease ecologists because field data from real world diseased or healthy wildlife populations is rare but can often be supplemented by expert prior knowledge. Therefore, a Bayesian modeller has the flexibility to combine both quantitative and qualitative data . Further, Bayesian hierarchical techniques can capture the intricacies of level, scale and hierarchy within ecosystems by accounting for their uncertainties, or if they are unobservable features, include them as random effects . Consequently, the uncertainty in latent variables such as an individual's infection status, which is often unknown and unobservable, can be both accounted for, and inferred. These characteristics allow Bayesian models to be used as proxies for diagnostic tests themselves—for example, to observe changes in expected biomarker patterns due to changes in individual disease status —or their associated accuracies .

Using Bayesian inference to research wildlife reservoirs of bTB

Bovine tuberculosis (bTB) infections—caused by zoonotic bacteria *Mycobacterium bovis*— are globally relevant, difficult to control, and scrutinised by disease ecologists across many host species. Research on mammals maintaining bTB reservoirs over wildlife-livestock boundaries dominate the literature, is high-profile and economically important. Yet evidence indicating that *no* capable bTB host should be disregarded in the search for bTB controls suggests that the current body of research on bTB infection reflects a tradeoff between economic and ecological importance. It is possible that Bayesian approaches could help bridge the data gaps between less and well-studied bTB hosts by enabling information on host ecology from non-disease studies to inform future epidemiological models.

In badger bTB research, a better understanding of transmissions within and among badger reservoirs, as well as between badgers and cattle, or other non-reservoir host species, is required. Like all disease systems, the

understanding of the badger-bTB system is constantly shifting with new pieces of information, which can act to better inform priors with expert knowledge and improve our beliefs. For example, the rapid serological Dual-Path Platform VetTB test has recently been validated for bTB testing in badgers and a new badger behaviour called super-ranging has been detected, which is potentially responsible for long-distance bTB transmissions . Consequently, these specific pieces of information *could* help provide updated estimates of disease transmission and disease progression within a badger bTB system .

Although the number of “how-to” papers describing the power of Bayesian inference in the context of wildlife epidemiology is increasing , research incorporating Bayesian modelling strategies specifically interested in the ante-mortem badger bTB system is limited. Consequently, we are only able to only touch on six specific examples of this following a comprehensive Web of Science search using combinations of the terms: “Bayesian”; “bovine”; “tuberculosis”; “model”; “badger”; “bTB”; and “testing”. Of these six publications , all used Bayesian methods to explore badger bTB transmissions on a “landscape-scale”, here defined as the geographical extent of Woodchester Park, Gloucestershire, UK, where the capture-mark-recapture data common to all six studies was collected. We consider that in the “data-poor” environment of wildlife disease modelling, our understanding of host-pathogen systems over a diversity of geographical scales is particularly limited.

In South Island, New Zealand—where brushtail possum (*Trichosurus vulpecula*) were speculated to be the keystone reservoir species of bTB for *ca.* three decades (*Trichosurus vulpecula*)—recent Bayesian research has provided confirmation that its possum population is responsible for South Island’s bTB maintenance; rather than its cattle population. In the UK, although it has been confirmed via Bayesian Integrated Population Models *why* Woodchester Park badgers are an efficient bTB reservoir , the directionality of bTB transmissions between badgers and cattle remains debated, and it is suspected that badgers are responsible for roughly half of bTB infections in cattle within high cattle-bTB incidence areas . Another analysis concluded that badger to cattle transmissions were ~ 10.4 times more frequent than *vice versa* . The geographic scale of study is just one aspect of ecological hierarchy which requires consideration within a whole-system model. Yet until we have whole-system modelling capabilities, we must be careful with general comparisons between studies over varying geographic, spatial, temporal and or statistical scales because they inevitably suffer from the change-of-support problem : they all require complex inference of variables at values that have not been observed. One solution to the change-of-support problem is Bayesian hierarchical modelling.

A whole-system model of the bTB systems, capable of linking information throughout an ecological hierarchy, is required, and an example of what this model may look like, is presented in Figure 1. A particular limitation in the development of such a model is the ability to incorporate individual badger heterogeneities; individual traits are often ignored in disease models since detailed longitudinal datasets of individuals—such as the Woodchester Park dataset—within diseased populations are rare.

Within the badger bTB system, complex interactions among individual badger heterogeneities —such as between sex, inbreeding, disease and ageing —and among-individual variation in traits, act as proxies for infectiousness or “risk” , and are thought to drive fine-scale bTB dynamics. Fundamentally, an understanding of fine-scale disease processes in combination with Bayesian methodologies arms ecologists with the ability to parameterise previously unobservable processes, such as actuarial senescence . In addition to Hudson’s study, Bayesian analyses of the Woodchester badger population have contributed knowledge on the increased susceptibility of male badgers to bTB and on the diagnostic accuracies of badger bTB tests : information which improves our capability to model badger heterogeneities in the future. The importance of latent variables and individual heterogeneities to a “whole-system” model of wildlife disease is further discussed in the final section of this review.

Overall, the idea of achieving a better understanding of the bTB system in wildlife hosts using a whole-system method is not a new one. Analogous to the “whole-system” model suggested within this review, proposed the need for a novel modelling framework, and recommended a comprehensive epidemiological model. In addition, suggested that combining contact networks with Bayesian inference is the future direction for understanding wildlife epidemics. The inclusion of a hierarchy of scale within whole-system ecological models

in general have been recommended by several authors .

Why is Bayesian hierarchical modelling essential to a whole-system approach?

Within a whole-system model we need to analyse multiple scales of information, therefore modelling using hierarchical Bayesian methods is important. Bayesian hierarchical models can help describe Bayesian networks in which the data can be grouped into levels, and where parameters may be inferred for any grouped level . Since information may be incorporated across multiple observational scales, notes them to be especially useful for parameterising demographic infection processes, such as disease progression, mortality, and transmission.

Bayesian hierarchical models usefully combine both statistical and ecological hierarchies, so we first summarise the difference. A statistical hierarchy forms when a large and complex model is composed of simpler sub-models, where the parameters it infers may exist only at single levels of the statistical hierarchy, may apply to multiple levels, or may be aggregated at one level to inform a process at a higher or lower level . Therefore, a statistical hierarchy describes the uncertainty within nested parameters, and in a Bayesian framework, this nesting forms a hierarchical model.

The statistical hierarchy within a Bayesian hierarchical model can represent levels of an ecological hierarchy, depending on which disease processes or states are being measured. For instance, genetic susceptibility can be considered a within-individual or within-host trait, disease progression a within-host process, and disease prevalence a population-level parameter. In hierarchical models, the individual-level of a hierarchical model is defined as the smallest measurable unit of that system : for example genes, such as those coding for disease susceptibility, could be viewed as the individual unit; alternatively, very detailed models might consider individual pathogens; coarser models might start with the individual host and model within-host infection processes in the abstract. When Bayesian methods are used to infer latent and unmeasurable states, this abstraction allows truly binary or categorical states—such as dead or alive; infected or uninfected—to be inferred as probabilities, which better reflect their lack of direct measurement . Expanding on Figure 1, infection status or “being infected” is an example of a host-specific latent variable, which may inform processes at all levels of an ecological model. Likewise, actuarial senescence—a decline in fitness with advancing age—is an example of a demographic process that could reasonably exist on multiple levels of an ecological hierarchy .

Both infection status and senescence have obvious implications for disease transmission since individual state is a key driver of physiology, ageing , and ultimately time-to-death. Yet utilising information about a latent state or process can be tricky. For example, senescence can be quantified as an—albeit very uncertain—individual-level process, and equally, as a demographic parameter; yet it is unmeasurable except at the population level. Further, the biology of senescence is complex: for example, in marmot individuals, senescence is more likely to occur in favourable conditions . Moreover, senescence theory itself remains disputed: the traditional viewpoint of senescence is that fitness declines with ageing , but this is challenged by other perspectives, for example that individual state dictates senescence in a heterogeneous fashion . Importantly, there are many ways to introduce complex latent variables such as senescence into whole-system models. What matters are reasonable and justifiable *a priori* assumptions regarding the proxy data available.

To date, Bayesian hierarchical methods have been applied only sparingly to wildlife disease problems. Here we can only cite eight examples (Table 2) following a comprehensive Web of Science search using combinations of the terms: “Bayesian”; “hierarchical”; “model”; “wildlife”; “animal”; “disease”; “infection”; “system”. These citations do not provide an exhaustive list of all papers that use Bayesian hierarchical methods to model disease but serve to present its limited application. Bayesian hierarchical models seem to have been applied to the field of wildlife disease epidemics in a rather *ad hoc* manner, in terms of methodology and terminology as well as across disease system, level of ecological hierarchy, or process to be inferred. The potential and power of Bayesian hierarchical methods is described above, yet they have rarely been used to explore individual, group and population hierarchies within the same model.

Table 2 indicates that the current applications of Bayesian hierarchical models are useful for generalising

large within-population and or landscape-scale processes, and that they are broadly applicable across disease systems. Yet, studies across broad ecological levels and scales are rare. Table 2 also suggests that studies purely investigating disease spatially, or spatially and temporally, are less common than those that have a temporal investigation alone. Equally, as demonstrated by our bTB case study, Bayesian hierarchical analyses of evolving longitudinal datasets are also rare, but likely integral to the discovery of fine-scale ecological interactions pertinent to understanding disease processes.

Modelling latent variables is essential to the whole-system model.

Bayesian state-space models are a form of Bayesian hierarchical model that allow Bayesian networks to easily distinguish dynamic biological processes such as changes through time from unavoidable errors due to the imperfect detection of disease. Accordingly, approaches such as those used by McDonald or Wells may be described as state-space methods since they quantify the randomness and variability of ecological processes (Table 3). Specifically, they do this by accounting for whether a parameter is unobserved or observed, as well as any associated sampling error. This means that Bayesian state-space models are especially good at, for example, teasing apart demographic stochasticity and sampling error.

The inferred latent variables and ecological findings of the Bayesian state-space models presented in Table 3 demonstrate that disease parameters are mainly studied on group or population levels even though most use individual serology data to inform models. Table 3 illustrates that the application of Bayesian state-space models within wildlife disease epidemiology is limited, but in our searches we also encountered problems caused by vague or inconsistent model terminologies. Underpinning this observation are three search methods referenced within Table 3. The first is a comprehensive Web of Science search using combinations of the terms: “Bayesian”; “state-space”; “disease”; “wildlife”, which only yielded four relevant studies. A further relevant study was found using the search topic “state-space model” when filtering by the Web of Science category “ecology”. Two additional relevant examples were found incidentally in the absence of either “state” or “space” as a keyword.

Three key conclusions can be drawn from the examples within Table 3. First, that observed serological data is common to all studies: presumably because most disease states in wildlife remain latent following visual surveillance. Second, it is encouraging that ecological stochasticity is modelled in the dimensions of space and time, often within the same study. And most importantly, the conclusions drawn by all examples found in Table 3 only regard population- or species- levels. Based on these conclusions, we observe that state-space models often span two levels of a hierarchy but rarely multiple latent variables. An example of this is demonstrated within previous work on the badger bTB system by , which used state-space models to infer the latent variable “alive” but ignored uncertainty in diagnostic test outcomes.

When combined with Bayesian state-space methods, Table 3 demonstrates that serological data collected over space and time can yield powerful conclusions about high-level disease parameters. But, state-space models are also an obvious tool for filling in any unknown relationships between individual disease states. For example, frequentist multi-state modelling revealed that epidemiological and demographic parameters vary between disease states in badgers , yet a vast number of unknown complexities still exist within this relationship which cannot be quantified without Bayesian methods. There is a fundamental need to parameterise processes within disease models more rigorously by applying Bayesian state-space theory. Although the need for good epidemiological parameter inference has been apparent for over a decade , the potential of state-space models has not yet been realised: they can help define the mutable nature of disease across *any* level of the “whole-system” model, inclusive of space, as well as time.

Including individual heterogeneities is essential, but difficult.

Studying disease on the individual level is difficult because single transmission events are not just impossible to observe in the wild, but also associated with a wide variety of host characteristics, such as behaviour, antibodies, age, or senescence: measuring the relevant predictors can be challenging. As explained above, there are few examples of how individual states have been modelled using Bayesian hierarchical models, including state-space techniques.

Wildlife diseases are often studied using data on antibodies or general pathological observations . Consequently, many epidemiological state-space models are based on serology records . Although the seropositive statuses of most wild host species are unknown , it is becoming increasingly important to look at the information that serology records provide, to reveal individual heterogeneities. For example, in *Eidolon helvum* fruit bats, seropositive thresholds were used to distinguish between the genetic and acquired immunities to Lagos bat virus and African henipavirus, using Bayesian mixture-models . Here, Bayesian inference determined that immunity relied on patterns in disease transmission , suggesting that serological data is an invaluable way to measure individual heterogeneities. In turn, this suggests that estimating seroprevalence is a good proxy for inferring the probability of infection in the absence of reliable testing. Yet even though the state-space models described in Table 3 are based on individual-level data, the outcome is still on a population level, with highly generalised disease processes , illustrating the difficulty in disaggregating the individualistic characteristics of disease processes.

A further difficulty in representing the individual state within state-space models is the complexity of the data involved. For wildlife systems, ageing is a latent individual process with a limited understanding particularly in terms of its relationship with disease. Serological data has been directly associated with age to infer infection rate, the probability of antibody loss and recovery rates in brucellosis-infected Elk . Yet to infer these parameters, the authors adopted Approximate Bayesian Computation methods due to the difficulty in writing closed form likelihood functions for the study parameters, and the associated difficulty in implementing them within a standard Monte-Carlo Markov Chain algorithm . This is an example of where the usefulness of Bayesian state-space modelling is currently limited in terms of its accessibility to disease ecologists. Although it is likely that the latent parameter “ageing” is intrinsically linked to disease via a host of known and unknown latent variables: within the whole-system model, this additional complexity must also be accounted for. Bayesian methods are a practical tool of choice for modelling complex systems, but realistically, the modelling of whole-systems using the Bayesian hierarchical approaches described within this review will rely on stronger collaborations between statisticians, epidemiologists and ecologists.

Conclusion

Modelling host-pathogen systems can be considered a “wicked problem” : its success is dependent on multidisciplinary thinking between statisticians, epidemiologists, and ecologists; and there is a balance between accepting over-simplified solutions and being overwhelmed by overly complex ones . Moreover, any solution involves balancing conflicting and fluid temporal and spatial ecological scales . In addition to space and time, the environmental processes that describe host movement—such as climate or seasons—are an often a disregarded, yet essential, third dimension required to model disease systems .

A deeper forensic approach is required to better understand and parameterise complex host-pathogen systems, and the Bayesian toolkit provides a good starting place for this. Overall, future studies of host-pathogen systems require a better representation of ecological hierarchy and scale. Scale needs to be examined in terms of applying suitable Bayesian methods (statistical hierarchy) and by paying attention to the complexity of the system that is being analysed (ecological hierarchy). The join-up between these scales has rarely been studied within ecological systems, and has never been completed for a single system, yet is essential to providing a whole-system model.

Our survey of the wildlife disease literature shows that the current application of Bayesian networks to disease problems is limited: in particular, there is a paucity of hierarchical analyses that infer truly latent parameters or individual heterogeneities across ecological scales. Bayesian methods are now being used in several wildlife disease systems but usually only to tackle standard hypotheses at a single level of the ecological hierarchy or, at most, spanning two levels of the ecological or statistical hierarchies.

By developing Bayesian hierarchical modelling methods and integrating them with real-world empirical data that is not exclusively serological, the potential exists for ecologists to create whole-system models that can provide unique insights into the epidemiology of wildlife disease networks. And the first step towards the whole-system model is to develop a Bayesian hierarchical model that spans the state-space nature of each

level of the host-pathogen ecological hierarchy.

Acknowledgements

AFB and DJH conducted this work within the scope of a part-time and self-funded PhD programme supervised by DJH, supported by the University of Exeter. We are grateful to Mario Recker and ?? anonymous reviewers for their valuable critique of draft manuscripts.

References

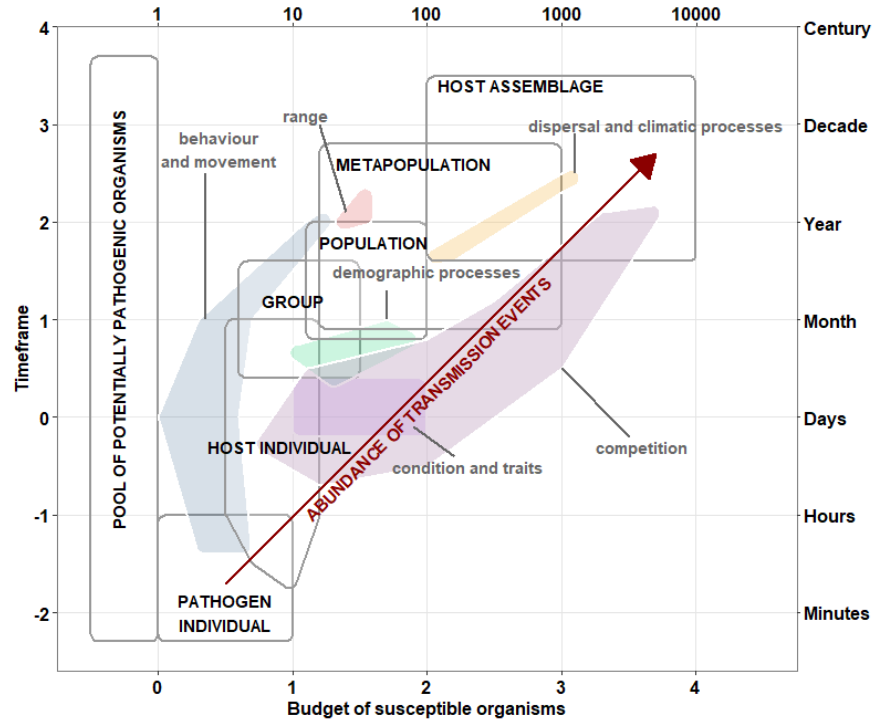


Fig. 1. A conceptual diagram illustrating an “ecological hierarchy”. Hierarchical ecological information, defined by the grey outlines (spanning from a single pathogen to a host assemblage) may contribute to a whole-system model of disease flow within a host-pathogen system. As the time since an initial infection increases on the y-axis, transmission events may also increase, and the budget of susceptible organisms represented on the x-axis that could be exposed to infection, increase. The ecological scale shown constitutes both the structures and processes acting throughout an ecosystem. Consequently, scale can be parameterised by system dynamics, such as host or population behaviour(s) or group demographic processes, which describe expected fluctuations around ecological equilibria. Coloured polygons therefore propose example locations for these latent dynamics in the context of a “whole-system” model.

Table 1. Examples of latent parameters and their proxy measures at each stratum of a typical ecosystem. A broad statistical method used to infer the respective latent parameter is also suggested. As the strata increase with respect to the number of organisms present, the statistical methods that may be used to infer common latent parameters become broader. Consequently, a large variety of spatio-temporal methods may be used to infer latent parameters above the host level using proxy data. As the strata represent more complex latent data, the need for spatial temporal modelling techniques—such as Bayesian state-space models—that can account for this complexity increase, and the number of commonly reported statistical methodologies to infer common latent parameters decrease. A whole-system model would be a spatio-temporal modelling technique able to infer latent parameters within any stratum of an epidemiological system, which may themselves be

dependent on latent parameters in other strata.

Layer of ecological network	Latent parameter to infer	Proxy measure in host	Strata
Pathogen	Virulence	Observation for symptoms (physical/behavioural) in host	Local
	Basic Reproduction Number	Serosurveys or behavioural surveys in host	Local
	Presence/absence	Laboratory culture of host serosurvey data	Local
Host	Location	Telemetry	Home range
	Infection status	Capture-mark-recapture	Metapopulation
	Alive/dead	Environmental DNA	Dynamics
Group	Membership	Social co-dynamics	Spatial
	Social Contact Network	Social co-dynamics	Spatial
Population	Temporal/spatial abundance	Camera trapping (presence/absence)	Spatial
	Population size	Resource abundance	Spatial
Metapopulation	Disease prevalence	Observation for symptoms (physical/behavioural) in host	Local
	Connectivity	Topography	Spatial
	Colonisation	Telemetry	Spatial
Assemblage	Gene flow	CMR	Spatial
	Species interactions	Prey kills in wild	Spatial
	Species distribution	Camera trapping (presence/absence)	Spatial
	Species richness	Stable isotopes	Spatial

Table 2. Examples of wild host-pathogen systems that have been investigated using Bayesian hierarchical modelling (S = spatial and T = temporal).

Host-pathogen system
bTB in European badgers
Devil Facial Tumour Disease (DFTD) in Tasmanian devils (<i>Sarcophilus harrisii</i>)
CWD in white-tailed deer (<i>Odocoileus virginianus</i>)
CWD in North American elk (<i>Cervus elaphus nelsoni</i>)
Influenza A in captive mallard (<i>Anas platyrhynchos</i>) and lesser snow geese (<i>Chen caerulescens</i>); <i>Yersinia pestis</i> in coyotes
Brucellosis (<i>Brucella abortus</i>) in wild elk and livestock herds
CWD in mule deer (<i>Odocoileus hemionus</i>)
Five pathogens (porcine reproductive and respiratory syndrome virus, pseudorabies virus, Influenza A virus, Hepatitis E virus)

Table 3. Examples of inferred wildlife disease parameters investigated using Bayesian state-space models.

Host-pathogen system	The Bayesian state-space model	The Bayesian state-space model
	Observed time series	Observed time series
	Serology	Other data
Swine Influenza in domesticated Chinese swine		Virological
CDV in lions (<i>Panthera leo</i>) and domesticated dogs (<i>Canis lupus familiaris</i>)		regional variation
Morogoro virus in multimammate mice (<i>Mastomys natalensis</i>)		weight; in
Hantavirus in striped field mice (<i>Apodemus agrarius</i>)		NA
CDV in grizzly bears (<i>Ursus arctos</i>) and wolves (<i>Canis lupus</i>)		NA
Fox (<i>Vulpes vulpes</i>) rabies		NA
Avian malaria (<i>Plasmodium relictum</i>) in Hawaiian honeycreeper species		NA

* Web of Science one-term and one-topic search; + Web of Science five-term search; ++ad hoc search