

# Using edmaps & Zonation to inform multi-pest early-detection surveillance designs

*Technical Report for CEBRA project 20121001*

James Camac<sup>1</sup>, John Baumgartner<sup>1</sup>, Susan Hester<sup>2</sup>, Ranjith Subasinghe<sup>3</sup>, and Susie Collins<sup>3</sup>

<sup>1</sup>The Centre of Excellence for Biosecurity Risk Analysis, The University of Melbourne

<sup>2</sup>UNE Business School, University of New England

<sup>3</sup>Department of Agriculture, Water and the Environment

August 2, 2021



**How to cite:**

Camac, J.S., Baumgartner, J.B., Hester, S., Subasinghe, R., Collins, S. (2021) Using edmaps & Zonation to inform multi-pest early-detection surveillance designs. Technical Report for CEBRA project 20121001.

# Contents

<b>1</b>	<b>Background</b>	<b>3</b>
1.1	Project objectives . . . . .	6
1.1.1	Outputs are preliminary not definitive . . . . .	6
1.1.2	Economic inputs & optimal surveillance designs . . . . .	7
1.1.3	Optimal surveillance effort and probabilities of absence . . . . .	7
1.2	Report structure . . . . .	8
<b>2</b>	<b>Changes &amp; enhancements to edmaps</b>	<b>9</b>
2.1	Changes in pathway leakage inputs . . . . .	9
2.2	New wind pathways . . . . .	13
2.3	Addition of native host input . . . . .	16
2.4	Modelling vectors and pathogens . . . . .	19
<b>3</b>	<b>Using Zonation &amp; edmaps outputs to inform multi-pest surveillance</b>	<b>20</b>
3.1	Estimating threat-specific establishment potential using edmaps . . . . .	20
3.2	Using Zonation to identify priority surveillance locations across multiple biosecurity threats . . . . .	31
3.3	Results: National Zonation outputs . . . . .	34
3.4	Results: Metropolitan Zonation outputs . . . . .	39
<b>4</b>	<b>Conclusions, Caveats &amp; Recommendations</b>	<b>46</b>
4.1	Conclusions . . . . .	46
4.2	Caveats & Recommendations . . . . .	47
	<b>Bibliography</b>	<b>52</b>
<b>A</b>	<b>Appendix A: Reference manual: edmaps R package</b>	<b>56</b>

# List of Figures

0.1	Top 100 surveillance locations for Sydney as determined by Core-Area Zonation (CAZ) multi-threat prioritisation. . . . .	2
1.1	edmaps is based on the principle that for a pest or disease to successfully establish it must overcome barriers associated with propagule pressure and the suitability of the abiotic and biotic environment. . . . .	4
2.1	Australian coastal segments exposed to northerly winds (i.e. winds originating from PNG and Indonesia). . . . .	15
2.2	Australian coastal segments exposed to Pacific winds (i.e. winds originating from Pacific islands) . . . . .	15
2.3	Australian coastal segments exposed to south-easterly winds (i.e. winds originating from New Zealand) . . . . .	16
2.4	Estimated distribution of native citrus species using points with 50km buffer (left) and alpha hull method (alpha = 130,000, right) on an Australian Albers raster . . . . .	18
2.5	Specifying threat groupings in edmaps parameters.xlsx. . . . .	19
3.1	$\log_{10}$ Estimated establishment likelihood of significant exotic plant pests, pathogens and pathogen vectors. . . . .	22
3.2	Estimated biotic suitability for significant exotic pests, pathogens or pathogen vectors. . . . .	26
3.3	Global climatic suitability for threats or threat vectors. . . . .	29
3.4	Estimated Australian climate suitability for significant exotic pests, pathogens or pathogen vectors. . . . .	30
3.5	The national top 5000 multi-threat priority surveillance locations as determined by Zonation's Core-Area prioritisation. . . . .	35
3.6	The national top 5000 multi-threat priority surveillance locations as determined by Zonation's Additive Benefit Function prioritisation. . . . .	36
3.7	Cumulative percentage of national establishment potential captured as surveillance coverage increases, based on individual and multi-threat prioritisation methods. . . . .	37
3.8	Percentage of national geographic overlap between the set of cells included in the optimal surveillance network when using threat-specific edmaps output, and the set of cells included when using either Core-Area Zonation (CAZ) or Additive Benefit Function (ABF) multi-threat prioritisation, with increasing area of surveillance coverage. . . . .	38
3.9	Top 100 surveillance locations for major Australian metropolitan areas as determined by Core-Area Zonation (CAZ) multi-threat prioritisation. . . . .	40
3.10	Top 100 surveillance locations for major Australian metropolitan areas as determined by Additive Benefit Function (ABF) multi-threat prioritisation. . . . .	41



3.11	Cumulative percentage of establishment potential captured in Sydney as surveillance coverage increases based on individual and multi-threat prioritisation methods. . . . .	42
3.12	Percentage of geographic overlap between the set of cells included in the optimal surveillance network for metropolitan Sydney when using threat-specific edmaps output, and the set of cells included when using either Core-Area Zonation (CAZ) or Additive Benefit Function (ABF) multi-threat prioritisation, with increasing area of surveillance coverage. . . . .	43
3.13	Cumulative percentage of establishment potential captured in Melbourne as surveillance coverage increases based on individual and multi-threat prioritisation methods. . . . .	44
3.14	Percentage of geographic overlap between the set of cells included in the optimal surveillance network for metropolitan Melbourne when using threat-specific edmaps output, and the set of cells included when using either Core-Area Zonation (CAZ) or Additive Benefit Function (ABF) multi-threat prioritisation, with increasing area of surveillance coverage. . . . .	45
4.1	Example of CAZ interactive map output highlighting priority surveillance locations within metropolitan Brisbane as well as threat-specific statistics for individual flagged locations. . . . .	49

# List of Tables

2.1	Average percent frequency of daily wind simulations crossing 500 km segments of the Australian coast for winds originating from Indonesia and PNG (North winds), Pacific islands (Pacific winds) and New Zealand (NZ winds). . . . .	14
3.1	Summary details about threats modelled in this study. . . . .	21
3.2	Summary of threat-specific leakage and viability bounds for pathways available in edmaps used in this analysis. . . . .	24
3.3	Source of occurrence records for threats with sufficient occurrence records requiring a climate suitability map . . . . .	27

# Executive summary

Post-border surveillance of exotic pests and diseases for the purposes of early detection is a critical component of any effective and efficient biosecurity system. However, surveillance is often designed on a threat-by-threat basis, often resulting in a portfolio of geographically diverse surveillance networks that encompass a wide variety of land uses with varying degrees of biosecurity risk. As such, surveillance is prone to inefficiency, and is frequently costly and labour intensive. Recently, CEBRA in collaboration with the Department of Agriculture, Water and the Environment (DAWE) developed a pragmatic and transparent framework for estimating post-border establishment potential of priority plant pests and diseases. This framework, now packaged into an R package called `edmaps`, estimates post-border establishment potential as a function of three fundamental barriers to establishment: 1) propagule pressure; 2) climate suitability; and 3) host availability. With the development of `edmaps`, an opportunity now exists to systematically identify areas of high establishment potential shared across multiple threats. This in turn can be used to examine whether the spatial configuration of early detection programs – currently determined on a threat-by-threat basis – can be consolidated to areas of shared high establishment potential, and thus, reduce logistical costs associated with surveillance.

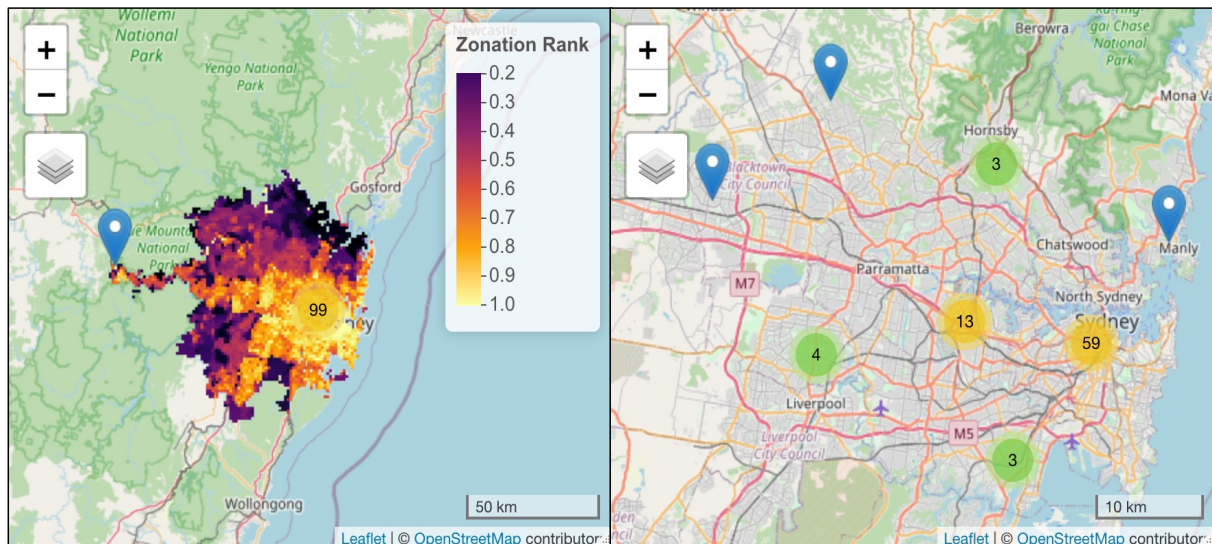
In this project, we examine how threat-specific `edmaps` output may be used in combination with the software `Zonation` to inform priority areas for early-detection surveillance across multiple threats. To achieve this we used `edmaps` to produce post-border establishment potential maps for 10 plant pests, 4 plant pathogens and 9 exotic pathogen vectors identified by DAWE as priority threats.

We found that `Zonation` can be used with threat-specific early detection maps to inform multi-threat surveillance prioritisation. Moreover, based on preliminary pathway inputs supplied by DAWE, our analyses suggest that by moving from a threat-specific prioritisation (i.e. using individual threat early detection maps) to a multi-threat prioritisation, one could geographically consolidate surveillance with only a marginal reduction in the threat-specific establishment likelihood captured per unit area surveyed – especially if the geographic distribution of establishment likelihoods is similar across threats.

Outputs included in this report are example multi-threat prioritisation maps derived from two `Zonation` algorithms – core-area `Zonation` (CAZ) and Additive Benefit Function (ABF). These maps are constructed at both national and metropolitan extents. An example of the metropolitan maps can be seen below (Figure 0.1).

Also accompanying this report as a separate file is an example interactive map of a core-area `Zonation` multi-threat prioritisation (`top_caz_national.html`). This file

allows readers to interactively examine an output of zonation that has identified the top 5,000 ranked priority locations across Australia for multi-threat surveillance. Users can zoom in on particular flagged locations and see the break down of threat-specific establishment likelihood captured and priority ranking locations. Also included as a separate file is the `edmaps` input file (`parameters.xlsx`) used to construct the threat-specific establishment likelihood maps. Raw output files can be supplied on request.



**Figure 0.1.:** Top 100 surveillance locations for major Australian metropolitan areas as determined by Core-Area Zonation (CAZ) multi-threat prioritisation. Left panels for each metropolitan area show the geographic extent of the metropolitan area and distribution of CAZ rankings, scaled to the range 0–1. Right panels show the geographic distribution of high priority surveillance locations. Warmer colours = higher priority rank (left panels) or higher number of priority sites in the vicinity (right panels), blue markers signify single locations. Note that flagged locations are approximate.

## Report Recommendations

1. Establish an expert task force for model interrogation, pathway parameter estimation and model implementation/use in surveillance;
2. Conduct regular reviews of pathway leakage parameters and pathway model assumptions;
3. Examine utility of incorporating economic considerations (e.g. differential threat impacts, logistical constraints) to inform optimal surveillance effort.
4. When deriving multi-threat priority maps always examine the proportion of establishment potential captured relative to using an individual threat map to examine threat representation in priority locations.

# 1. Background

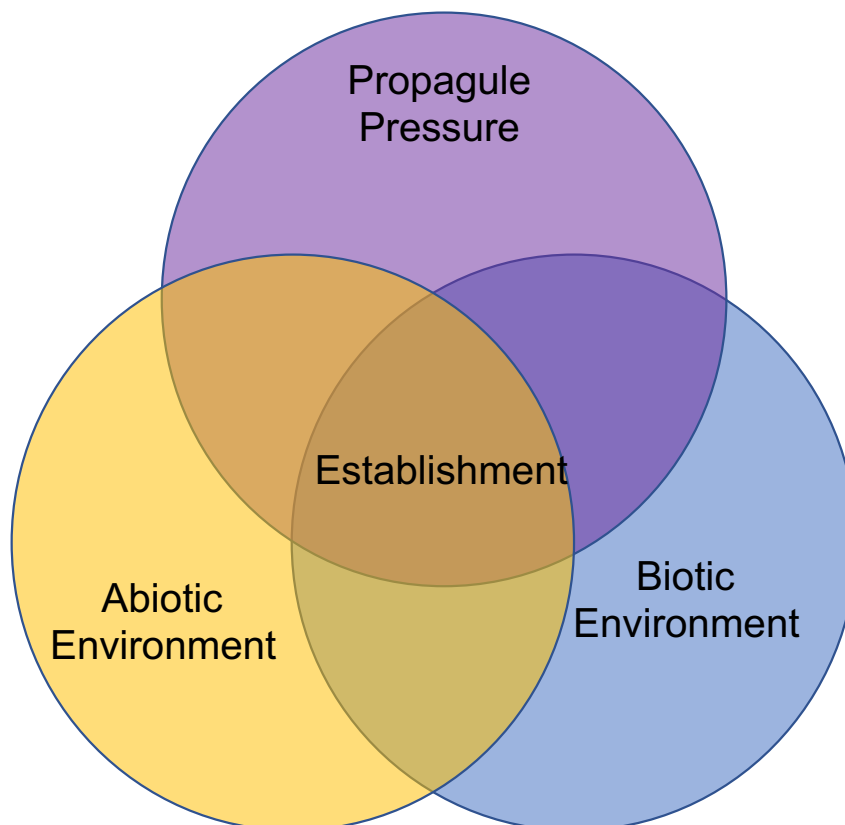
Early detection of high impact plant pests and diseases is critical for their containment, control and eradication, which in turn, minimises the magnitude of economic, environmental and social consequences. However, implementing effective and efficient early detection surveillance programs is becoming increasingly difficult worldwide, as countries become more exposed to novel threats as a result of changing climate coupled with increasing globalisation of human movement and trade.

In Australia, post-border early detection surveillance programs for exotic plant pests and diseases are designed on a threat-by-threat basis. While this process generally incorporates the best available knowledge and data on high risk points of entry and biology of specific pests, the underlying models used to inform where to conduct surveillance often vary among jurisdictions and are not always transparent or well justified.

The lack of a consistent and transparent framework can also introduce inefficiencies when conducting surveillance across multiple threats. For example, without a framework for identifying areas with the highest likelihood of incursion and establishment, biosecurity practitioners will be unable to identify high risk areas shared across multiple threats, and consequently, where surveillance effort can be consolidated beyond points of entry. Ultimately, this leads to biosecurity agencies implementing a large portfolio of threat-specific surveillance programs that are geographically diverse, and thus, logistically costly to maintain.

The Centre of Excellence for Biosecurity Risk Analysis (CEBRA), in collaboration with the Australian Department of Agriculture, Water and the Environment (DAWE), have recently focused on improving the robustness, transparency and reproducibility of establishment likelihood models underlying early detection surveillance programs for exotic plant pests and diseases. This work has resulted in the development of an internally available R package called `edmaps` ([Camac \*et al.\*, 2021](#)). `edmaps` uses a pragmatic (given data constraints) and reproducible workflow to integrate a variety of public and department-held data sources to estimate and map establishment likelihoods as a function of three fundamental factors influencing exotic pest/disease establishment outlined in [Catford \*et al.\* \(2009\)](#), namely: propagule pressure, abiotic suitability and biotic suitability (Figure 1.1).

A major advantage of edmaps is that it contains an easy-to-use interface (an Excel spreadsheet of parameter inputs) and requires minimal coding, statistical or GIS skill to implement<sup>1</sup>. Moreover, it is designed to be flexible such that it can be used on most exotic plant pests and diseases. But perhaps its greatest strength is that it allows one to identify knowledge and data gaps, while also providing the quantitative foundations necessary to incorporate additional biological, pathway or threat-specific complexity as data becomes available. With the development of edmaps, an opportunity now exists to systematically identify areas of high establishment potential shared across multiple threats. This in turn can be used to examine whether the spatial configuration of early detection programs – currently based on a threat-by-threat basis – can be consolidated to areas of shared high establishment potential. However, identifying and prioritising areas for multi-threat surveillance is not as simple as ranking the sum of the threat-specific maps of establishment likelihood. Doing so would likely bias surveillance towards a small number of pests with high establishment potential, while providing little to no coverage of other pests that may have different geographic distributions of establishment potential. As such, more sophisticated spatial prioritisation tools are needed.



**Figure 1.1.:** edmaps is based on the principle that for a pest or disease to successfully establish it must first overcome three geographic barriers related to propagule pressure (i.e. Can it get to the location of interest?), abiotic suitability (i.e. Is the climate suitable?) and biotic suitability (e.g. Is there a suitable host or food source available?). Figure is adapted from [Catford et al. \(2009\)](#).

<sup>1</sup>Though it is important to have a firm understanding of the data inputs and model assumptions.

Since the early 1980s scientists have developed a variety of sophisticated spatial prioritisation tools for identifying locations for conservation investment (Wintle *et al.*, 2019; Wintle, 2008). These tools have increasingly been used to identify priority areas not only for biodiversity and protected area establishment, but also for investment in a variety of other conservation activities such as fire management, invasive species control, biodiversity offset prioritisation and land restoration (Wintle *et al.*, 2019; Kukkala & Moilanen, 2013; Wilson *et al.*, 2009). One such tool that has shown great promise is *Zonation* (Moilanen *et al.*, 2005, 2009).

*Zonation* is a spatial prioritisation tool for optimising resource allocation to maximise a defined objective function while observing decision constraints and, if available, associated costs. It is commonly applied to conservation reserve design problems, whereby habitat quality maps for a set of threatened species are used to identify locations across the landscape that are most important to include in a network of reserves. *Zonation* works by taking these maps and producing a hierarchical prioritisation of the landscape, ranking the importance of grid cells by iteratively removing those that provide the least value to the reserve design. In this situation, value is defined with respect to habitat quality, and, if information is available, can be balanced against spatial costs such as the varying cost of land parcels. *Zonation* can also account for complementarity in its cell removal algorithms, where cells will be ranked highly if they contain high quality habitat for a species that is not represented elsewhere in the landscape.

The surveillance planning problem is analogous: biosecurity agencies seek to identify the optimal spatial arrangement of early-detection surveillance when the establishment potential varies both among threats and across space. They also want to ensure all target threats are appropriately represented in the surveillance program. By valuing complementarity, *Zonation* can ensure that a minimum threshold level of surveillance is achieved for each species.



## 1.1. Project objectives

### Primary objective: Using edmaps & Zonation to inform multi-pest surveillance designs

The primary objective of this report is to *illustrate* how edmaps and Zonation can be used to identify areas of high establishment potential shared across multiple threats. In doing so, we compare the total establishment potential captured by a threat-by-threat approach, whereby surveillance is optimised for each species independently, to that captured by an approach based on multi-threat priority areas identified by Zonation. We conduct this analysis at both the national scale, as well as for the major metropolitan centres of Sydney, Melbourne, Hobart, Adelaide, Perth, Darwin, Cairns and Brisbane.

To achieve the above objective, we used edmaps to estimate and map the national post-border establishment potential for 10 plant pests, 4 plant pathogens and 9 exotic pathogen vectors (representing 12 individual pests/diseases and 2 aggregated pathogen groups, i.e. Xylella, and bee mites) identified by the Department of Agriculture, Water and the Environment (DAWE) as national priority threats. We outline the process and findings of this work in Chapter 3.

### Auxiliary objective: Enhancements to edmaps

While changes to edmaps infrastructure were initially considered out-of-scope in this project, DAWE had recently changed how annual pathway leakage rates were estimated and reported for priority plant threats. To incorporate this change into edmaps some modifications were required. We outline these changes, as well as other enhancements made to edmaps in Chapter 2 of this report.

## What was considered out-of-scope for this report

### 1.1.1. Outputs are preliminary not definitive

At the time of writing, critical pathway inputs (i.e. pathway leakage and viability bounds) required by edmaps to estimate viable threat propagule pressure and consequently establishment potential, were still being estimated by the Department of Agriculture, Water and the Environment. As a consequence, the focus of the project was to *illustrate* how edmaps and Zonation can be used together to inform multi-threat surveillance. To achieve this, we used preliminary approximations of pathway inputs supplied by DAWE to illustrate the work flow and outputs available to practitioners. Prior to use in informing surveillance designs, model inputs, outputs and assumptions should undergo rigorous independent examination by taxon and surveillance experts. Providing such expert endorsed outputs was considered out-of-scope for this project,



however, we provide details about how such a review process could be undertaken to further improve the rigour of models and associated outputs in Chapter 4.2.

### 1.1.2. Economic inputs & optimal surveillance designs

This report provides an examination of where to prioritise early-detection surveillance solely as a function of estimated establishment potential derived from edmaps. Surveillance costs or the differential potential economic, environmental and social impacts among selected national priority threats (which could be used to weight threats differently) are not considered in this report. However, we envisage that the framework outlined here could be incorporated into future economic analyses that would explicitly consider the optimal resource allocation (e.g. [Nguyen \*et al.\*, 2021](#); [Camac \*et al.\*, 2020](#); [Kompas \*et al.\*, 2017](#); [Kompas & Che, 2009](#)) and the marginal benefits gained from early detection of each pest using both single pest and multi-pest surveillance designs.

### 1.1.3. Optimal surveillance effort and probabilities of absence

This report does not attempt to quantify the optimal level of surveillance effort either within a location or across a region. It also does not attempt to estimate location-specific probabilities of absence as a function of surveillance effort and establishment potential ([Camac \*et al.\*, 2019](#); [Stanaway \*et al.\*, 2011](#)). To determine each requires knowing budget constraints, surveillance costs, potential impacts of threats as well as details about the threat-specific surveillance sensitivity (e.g. [Camac \*et al.\*, 2020](#)). The framework outlined in this report is focused on identifying *where* to prioritise early detection surveillance across multiple threats. We illustrate this by identifying the top 5000 1km grid cells across the Nation, or the top 100 1km grid cells within each metropolitan area. A subsequent analysis would be required to determine the number of priority locations that should be sampled, and how much sampling should be done within each to meet a desired level of confidence that the surveillance system would detect incursions should they occur.

## 1.2. Report structure

In this report we examine the utility of using `edmaps` & `Zonation` to inform multi-pest early-detection surveillance designs. To meet this aim, we have structured the report into four distinct chapters.

1. Chapter 1 provides a detailed background of the problem and the scope of this project.
2. Chapter 2 outlines changes and extensions to `edmaps` that were required to: 1) cater for a larger suite of priority pests, pathogens and/or pathogen vectors; 2) accommodate changes in leakage rate inputs provided by DAWE; and 3) include new pathways and options for incorporating native host ranges.
3. Chapter 3 provides a summary of how threat-specific maps of establishment potential were derived from `edmaps` and how they were used in combination with the prioritisation software `Zonation` to identifying priority areas in which early-detection surveillance may be undertaken for multiple threats.
4. Chapter 4 provides a summary of key findings accompanied by associated caveats and recommendations when using this approach.

The Appendix of this report (Appendix A) provides the latest `edmaps` R package manual with all changes incorporated as part of this project. The latest version of `edmaps` can be downloaded from the following github repository: <https://github.com/jscamac/edmaps>.

## 2. Changes & enhancements to edmaps

Due to changes in how DAWE estimates and reports pathway leakage rates of plant pests and diseases, some modifications to edmaps were deemed necessary. The latest version of edmaps can be readily downloaded at <https://github.com/jscamac/edmaps>. In the following sections we outline these changes as well as additional enhancements made as part of the project.

### 2.1. Changes in pathway leakage inputs

The biggest change to edmaps was triggered by a change in how DAWE estimated and reported leakage parameters. In the original model (Camac *et al.*, 2021), edmaps used annual point estimates of the expected number of leakage events that would occur in a given year for each pathway of risk. These point estimates were derived from DAWE's own Risk Return Resource Allocation (RRRA) model. The plant division of DAWE has recently attempted to add additional biological complexity to this process, whereby two sets of independently estimated parameters are provided for each pathway deemed *high risk* to a specific plant pest or disease. These are described below.

#### Parameter set 1: Annual expected number of post-border leakage events

The first set of parameters are the lower and upper bounds (2.5th and 97.5th percentiles) of the expected number of leakage entries for a given pest and pathway combination per year ( $\text{Entry}_{\text{low}}$ ,  $\text{Entry}_{\text{high}}$ ). These Entry estimates are conceptualised as the lower and upper expected number of consignments that could constitute a border leakage event, irrespective of whether an establishment event could occur. Examples include a passenger carrying a disease on undeclared plant material or snails on a container.

These bounds are obtained using a combination of interception data and expert opinion methods (e.g. Hemming *et al.*, 2018), which are then used to determine the 95% confidence limits that would encompass the potential number of entry events in a given year.

## Parameter set 2: Pathway probabilities of establishment potential

The second parameter set supplied by DAWE describes the threat-specific lower and upper probabilities an entry event could be viable for establishment ( $p_{low}, p_{high}$ ). These parameters are estimated using expert elicitation, where experts are asked to propose 95% probability bounds that encapsulated details such as whether a contaminated consignment is likely to contain a viable population size for establishment, the survivability along the pathway, and whether likely end-point destinations will contain suitable host material. This additional complexity was deemed necessary for informing post-border surveillance because while some pathways may contribute to post-border leakage events, high mortality rates and/or low expected population sizes of these events may make establishment highly unlikely to occur.

## Integration into edmaps

Incorporating this additional complexity into edmaps required a substantial change in how risk was estimated and dispersed. In the following sections we briefly summarise how edmaps previously estimated post-border establishment potential (the “old approach”; more details can be found in [Camac et al. \(2021\)](#)). We then detail the modifications required to cater for the change in leakage inputs (the “new approach”).

### The old approach

In the original model ([Camac et al., 2021](#)), edmaps approximated the expected number of contamination events in raster cell  $i$ ,  $N_i$ , across  $k$  pathways of interest as:

$$N_i = \sum_{k=1}^n \left( w_{i,k} \times \text{Total pathway units}_k \times \text{Pr}(\text{Contamination}_k) \right). \quad (2.1)$$

Here,  $\text{Total pathway units}_k$  is the known total incoming number of pathway units (e.g. number of tourists, number of containers), for pathway  $k$ , in a representative year.  $w_{i,k}$  is a pathway-specific weight ranging from 0 to 1 for raster cell  $i$ , used to proportionally allocate  $\text{Total pathway units}_k$  across raster cells. These pathway-specific weights are based either on spatially explicit historical data (e.g. container destination data, NRM fertiliser usage statistics), or on spatially explicit continuous factors thought to be correlated with post-border movements (e.g. population density, tourist accommodation, distance from airports etc.) for each pathway ([Camac et al., 2021](#)). Note that  $w_{i,k}$  is also equal to the probability that a single pathway unit for pathway  $k$  arrives at cell  $i$ .  $\text{Pr}(\text{Contamination}_k)$  is the per unit probability of contamination for pathway  $k$ .  $\text{Pr}(\text{Contamination}_k)$  is calculated by dividing the point estimates of leakage rates provided by DAWE’s RRRRA model,  $\text{Entry}_k$ , for the  $k$  pathways of interest, by the respective number of units entering the country in a given year,  $\text{Total pathway units}_k$ :

$$\Pr(\text{Contamination}_k) = \frac{\text{Entry}_k}{\text{Total pathway units}_k}. \quad (2.2)$$

Using the old approach, edmaps would then derive a relative establishment score,  $N_{est_i}$ , by weighting the expected number of contaminations arriving in raster cell,  $N_i$ , by the associated biotic (i.e. host availability) and abiotic (i.e. climate) suitability:

$$N_{est_i} = N_i \times \text{Abiotic suitability}_i \times \text{Biotic suitability}_i, \quad (2.3)$$

where both abiotic and biotic suitability components are on a scale of 0 to 1, and thus, assumed to be equally weighted.

## The new approach

The new approach shifts focus from the user manually calculating the pathway-specific probability of contamination as defined in equation 2.2, to one where the user simply inputs the reported lower and upper bounds of both the expected number of entries and establishment probabilities for each pathway.

edmaps uses the entry bounds ( $\text{Entry}_{low_k}$ ,  $\text{Entry}_{high_k}$ ) for pathway  $k$  to estimate the mean and standard deviation of a log-normal distribution, from which the rate parameter  $\lambda_k$  can be estimated:

$$\log(\mu_k) = \frac{\log(\text{Entry}_{low_k}) + \log(\text{Entry}_{high_k})}{2} \quad (2.4)$$

$$\log(\sigma_k) = \frac{\log(\mu_k) - \log(\text{Entry}_{low_k})}{1.96} \quad (2.5)$$

$$\lambda_k \sim \text{Log Normal}(\log(\mu_k), \log(\sigma_k)) \quad (2.6)$$

The discrete expected number of entry events,  $n$ , in a given year can then be simulated from a Poisson distribution using the rate parameter,  $\lambda$ :

$$n_k \sim \text{Poisson}(\lambda_k) \quad (2.7)$$

Using a similar approach, the lower and upper probability bounds,  $(p_{low}, p_{high})$  for pathway  $k$  can be used to estimate the mean and standard deviation of a logit-normal distribution, from which the likelihood of establishment viability,  $p_k$ , can be simulated:

$$\text{logit}(\mu_{p_k}) = \frac{\text{logit}(p_{low_k}) + \text{logit}(p_{high_k})}{2} \quad (2.8)$$

$$\text{logit}(\sigma_{p_k}) = \frac{\text{logit}(\mu_{p_k}) - \text{logit}(p_{low_k})}{1.96} \quad (2.9)$$

$$\text{logit}(p_k) \sim \text{Normal}(\text{logit}(\mu_{p_k}), \text{logit}(\sigma_{p_k})) \quad (2.10)$$

$$p_k = \text{invlogit}(\text{logit}(p_k)) \quad (2.11)$$

edmaps then uses the simulated distributions of  $n_k$  and  $p_k$  to estimate the expected number of viable incursion events for each pathway,  $k$ , using a Binomial distribution:

$$N_{\text{viable}_k} \sim \text{Binomial}(n_k, p_k) \quad (2.12)$$

$N_{\text{viable}_k}$  is then used to create a discrete probability mass function,  $p_{\text{viable}_{N_k}}$ , that defines the discrete likelihood that  $N$  viable incursions will occur in a given year, attributable to pathway  $k$ .

edmaps then uses  $p_{\text{viable}_{N_k}}$  to estimate the probability raster cell  $i$  receives one or more viable incursions from pathway  $k$ , as:

$$\Pr(p_{\text{viable}_{i_k}}) = 1 - \sum_{N=0}^n \left( p_{\text{viable}_{N_k}} \times (1 - w_{i,k})^N \right), \quad (2.13)$$

where  $w_{i,k}$  is a pathway-specific weight that estimates the likelihood that a pathway unit for pathway  $k$  will arrive in cell  $i$ , as described earlier. For example, for the mail pathway, the geographic distribution of mail items is assumed to follow that of human population density, hence  $w_{i,k}$  for that pathway is defined as the population density at cell  $i$ , divided by the sum of population density for the entire study area (Australia). The complement,  $1 - w_{i,k}$ , seen in Equation 2.13, gives the likelihood that a given pathway unit for pathway  $k$  does *not* arrive in cell  $i$ .

The probability of one or more viable incursions at cell  $i$  across pathways,  $\Pr(p_{\text{viable}_i})$ , then becomes:

$$\Pr(p_{\text{viable}_i}) = 1 - \prod_{k=1}^n (1 - p_{\text{viable}_{i_k}}) \quad (2.14)$$

`edmaps` then estimates the probability of establishment by weighting  $\Pr(p_{\text{viable}_i})$  by the estimated biotic and abiotic suitability of the location:

$$\Pr(p_{\text{establishment}_i}) = \Pr(p_{\text{viable}_i}) \times \text{Abiotic suitability}_i \times \text{Biotic suitability}_i. \quad (2.15)$$

Like the original approach, biotic and abiotic suitability estimates are on a scale of 0 to 1, and thus, assumed to be equally weighted.

A key difference from the old approach is that `edmaps` now estimates the weighted probability an establishment event will occur, whereas before it estimated the weighted expected number of such events. We believe that this new approach not only better encapsulates the inputs DAWE is now estimating, but it also allows `edmaps` outputs to be directly used in other applications. Example applications might include seed-ing spread models (Bradhurst *et al.*, 2021, 2020), integrating with surveillance data to inform probabilities of absence (Camac *et al.*, 2019) or facilitating development of Bayesian hierarchical models that simultaneously estimate both incursion and detection processes (Stanaway *et al.*, 2011).

## 2.2. New wind pathways

Discussions with DAWE and relevant stakeholders (e.g. the National Citrus Surveillance Steering Committee) revealed a need to incorporate wind-dispersal pathways into `edmaps`. Long distance wind dispersal events have been identified as a significant pathway of entry into Australia for as much as 18% of the 3000+ horticultural pests and diseases (Finlay *et al.*, 2018), including Asiatic citrus psyllid, examined in this report.

The Plant Biosecurity CRC (PBCRC) recently used air-parcel trajectory modelling coupled with wet deposition simulations to determine the frequency of potentially threat-carrying wind events crossing Australian coasts. Specifically, this modelling estimated the frequency of winds originating from three sources – the north (Southern Indonesia & Papua New Guinea), the Pacific (i.e. Solomon Islands, Vanuatu, New Caledonia and Norfolk Island) and New Zealand – crossing sixteen 500 km segments of the northern and eastern Australian coastlines (Finlay *et al.*, 2018). For all wind pathways, organisms were assumed to be able to stay in the air-column for at least 200 hours ( 8 days). For more details about this modelling approach please consult Finlay *et al.* (2018).

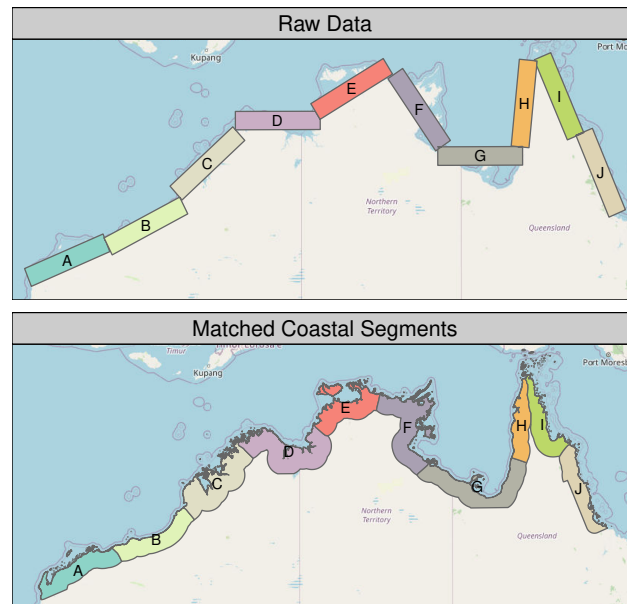
At the request of DAWE project leaders and stakeholders, CEBRA have incorporated the modelled wind frequencies presented in [Finlay et al. \(2018, Table 2.1\)](#) as three new wind pathways in `edmaps`. Polygons associated with these three wind pathways – Northern winds, Pacific winds and New Zealand winds – were mapped to the strip of coastline through a Voronoi tessellation of the landscape (Figures 2.1–2.3). To achieve this, for each primary wind direction we: (1) identify the subset of the Australian coastline adjacent to the [Finlay et al. \(2018\)](#) input polygons (by applying a spatial buffer to the polygons and intersecting with the coastline polyline); (2) convert this subset of coastline to points; (3) additionally, buffer the coastline subset by a user-specified distance (representing the distance inland permeated by coastal winds) and intersect with a polygon of Australia; (4) perform a Voronoi tessellation of the polygon resulting from step 3, with respect to the points generated at step 2; (5) use a spatial join to transfer the original [Finlay et al. \(2018\)](#) segment IDs from the points to the polygons; and (6) dissolve the Voronoi polygons, grouping by segment ID. This entire procedure is performed by the `rasterize_wind()` function within the `edmaps` R package. For each wind pathway, the resulting spatial distribution of deposition frequencies is used by `edmaps` as a weight layer, defining the relative frequencies with which wind-dispersed threats arrive at each wind region.

As with other pathways, users are required to input the bounds associated with both the entry rate ( $\text{Entry}_{low_k}, \text{Entry}_{high_k}$ ) and likelihood of viable establishment ( $p_{low}, p_{high}$ ). However, users are also required to enter a threat-specific distance parameter, that describes how far inland the influence of wind should extend. As we have no information as to how the propagule pressure associated with these wind events changes with distance from coast, the model assumes the risk in a given coastal segment is uniform across the influenced region. Users can specify these wind pathways within the species-specific parameters tab of the user input file, `parameters.xlsx`.

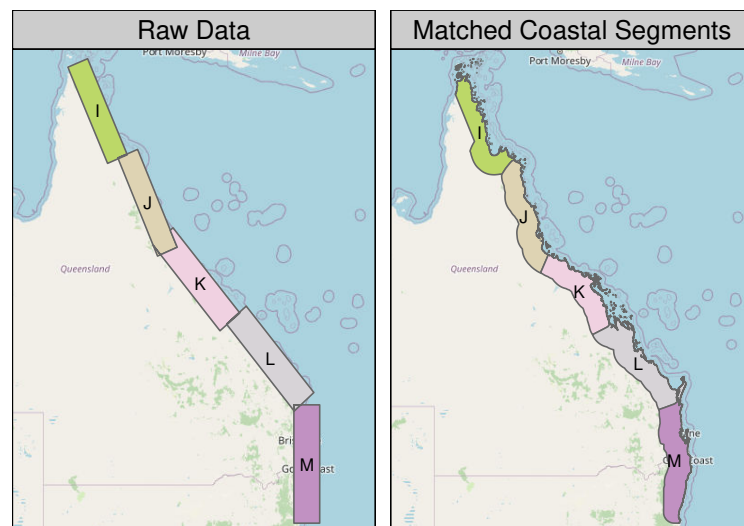
**Table 2.1.:** Average percent frequency of daily wind simulations crossing 500 km segments of the Australian coast for winds originating from the north (i.e. PNG, Indonesia), the Pacific (i.e. Pacific islands) and New Zealand. Based on 10-year (North winds) and 5-year (NZ and Pacific winds) daily wind trajectory and wet deposition modelling by [Finlay et al. \(2018\)](#).

Coastal ID	Reference name	North winds	Pacific winds	NZ winds
A	Exmouth	0.3	0.0	0.0
B	Broome	2.6	0.0	0.0
C	Derby	8.7	0.0	0.0
D	Kununurra	16.2	0.0	0.0
E	Darwin	35.4	0.0	0.0
F	Arnhem/Groyte	38.5	0.0	0.0
G	Gulf Carp-South	15.1	0.0	0.0
H	Weipa	34.9	0.0	0.0
I	Iron Range	34.0	24.8	0.0
J	Townsville	6.8	17.0	0.0
K	Whitsunday	0.0	10.6	4.2
L	Gladstone	0.0	8.4	5.5
M	Bryon Bay	0.0	5.6	8.2
N	Newcastle/Sydney	0.0	0.0	5.6
O	Mallacoota	0.0	0.0	3.8
P	Tasmania	0.0	0.0	2.7

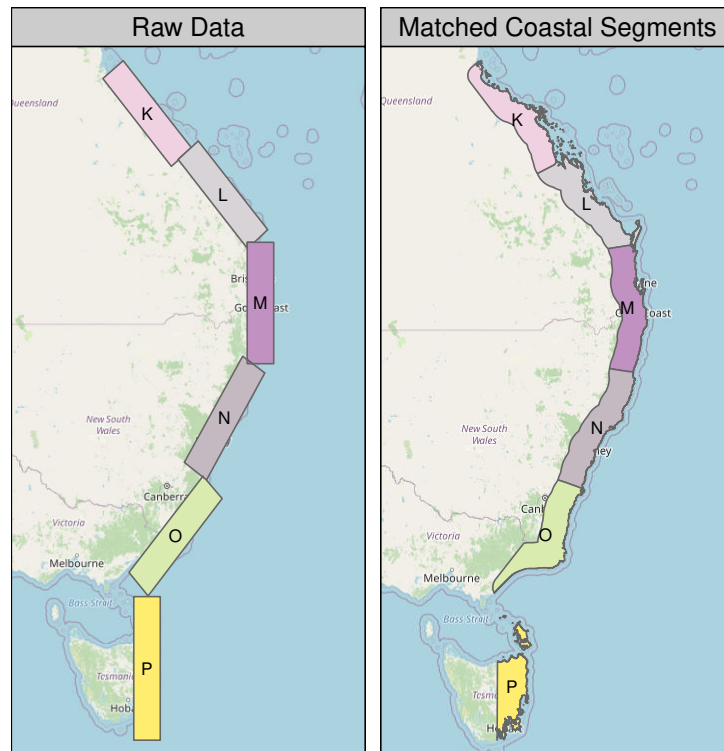




**Figure 2.1.:** Australian coastal segments exposed to northerly winds (i.e. winds originating from PNG and Indonesia). Top panel shows the original polygons constructed by [Finlay et al. \(2018\)](#). Bottom panel are the polygons mapped onto the Australian coastline using Voronoi tessellation.



**Figure 2.2.:** Australian coastal segments exposed to Pacific winds (i.e. winds originating from Pacific islands). Left panel shows the original polygons constructed by [Finlay et al. \(2018\)](#). Right panel are the polygons mapped onto the Australian coastline using Voronoi tessellation.



**Figure 2.3.:** Australian coastal segments exposed to south-easterly winds (i.e. winds originating from New Zealand). Left panel shows the original polygons constructed by [Finlay et al. \(2018\)](#). Right panel are the polygons mapped onto the Australian coastline using Voronoi tessellation.

## 2.3. Addition of native host input

DAWE and associated stakeholders requested the ability for users to either supply their own native host distribution layer, or provide a mechanism through which edmaps could approximate it. The reasoning for this enhancement was that in some cases, native hosts pose a significant source of host material that is otherwise not readily captured by either the National Vegetation Information System (NVIS) or the Australian Land Use and Management Classification (ALUMC) rasters. Estimated establishment potential would therefore be underestimated in regions where these native species are the predominant host material.

To incorporate this additional complexity into edmaps we have provided users with the ability to supply a binary-valued GeoTiff file defining the native host distribution (0 = absent, 1 = present). The path to this raster can be specified within the species-specific parameters tab of the user input file, `parameters.xlsx` and can be used to supplement specified host ranges defined by ALUMC or NVIS land codes.

To facilitate users in constructing a raster object of native host distribution, the function `rasterize_range` has been added to edmaps. This function allows users to supply a `.csv`, `SpatialPoints` or `sf` object containing coordinates of native species

occurrences. Such occurrences can be obtained from local experts, GBIF ([gbif.org](https://gbif.org)) or the Atlas of Living Australia (ALA).<sup>1</sup> The function then uses these coordinates in combination with a template raster, defining the extent and resolution of output, to construct a native host distribution raster that can be used in `edmaps`. `rasterize_range` provides two methods for constructing a native host distribution raster: 1) directly transfer coordinates to raster; or 2) construct alpha hulls (Pateiro-Lopez *et al.*, 2019) around occurrence records.

Transferring occurrence records directly to a raster is the simplest and most conservative approach. Effectively, this approach means that the occurrence records used, accurately and completely define all the raster cells that contain native host material. That is, only raster cells containing an occurrence record are defined to have suitable native hosts, and all others are not. This is unlikely, as vegetation surveys and reporting varies considerably across geographic space (e.g. Zizka *et al.*, 2021). `rasterize_range` can somewhat alleviate this strong assumption by allowing users to specify a buffer radius, whereby all raster cells within that distance of occurrence records are deemed to have native hosts.

An alternative approach also available in `rasterize_range` is to construct alpha shapes (i.e. alpha hulls) around the supplied occurrence records. While still susceptible to biases in occurrence record sampling, this approach is less conservative in that it estimates hulls of native distribution, whereby cells that fall within these alpha hulls are deemed to contain native hosts, and those outside them it do not. To use the alpha hull method, users are also required to specify an alpha parameter that is relevant to the template raster projection and resolution. This often involves some trial and error until an appropriate alpha parameter is defined. For example, for an Australian Albers projection (EPSG:3577; which is in meters), users might want to specify the alpha parameter somewhere between  $10^4$  to  $10^6$ . By contrast, if the template raster was unprojected (i.e. decimal degrees, EPSG:4326), then an alpha value between 0 and 10 is likely to be appropriate.

Irrespective of the method used, we strongly recommend that when incorporating native host ranges into `edmaps`, these are cross-verified by taxon experts. As a consequence, users are required create a native host raster and include its file path in the `parameters.xlsx` spreadsheet under the column name "Host distribution raster file". It is important to note, that the above methods result in a binary native host layer. The consequence of using binary layers is that the users are making the strong assumption that there is zero probability of establishment in regions that contain 0's. A continuous alternative that could be used to weight establishment likelihood is for users to use a Species Distribution Model (SDM) to approximate native host range. If using a SDM method, the main criteria is that the output raster must be constrained between 0 and 1, and that the assumption will be that the establishment likelihood is scaled proportionally to the supplied layer.

---

<sup>1</sup>`edmaps` has a function called `get_gbif_records` that can be used to query GBIF. As GBIF records can contain many errors, we recommend that retrieved records are carefully scrutinized and cleaning routines from R packages such as `CoordinateCleaner` (Zizka *et al.*, 2019) are used. Note that `edmaps` uses `CoordinateCleaner` automatically to clean GBIF records used to estimate abiotic suitability

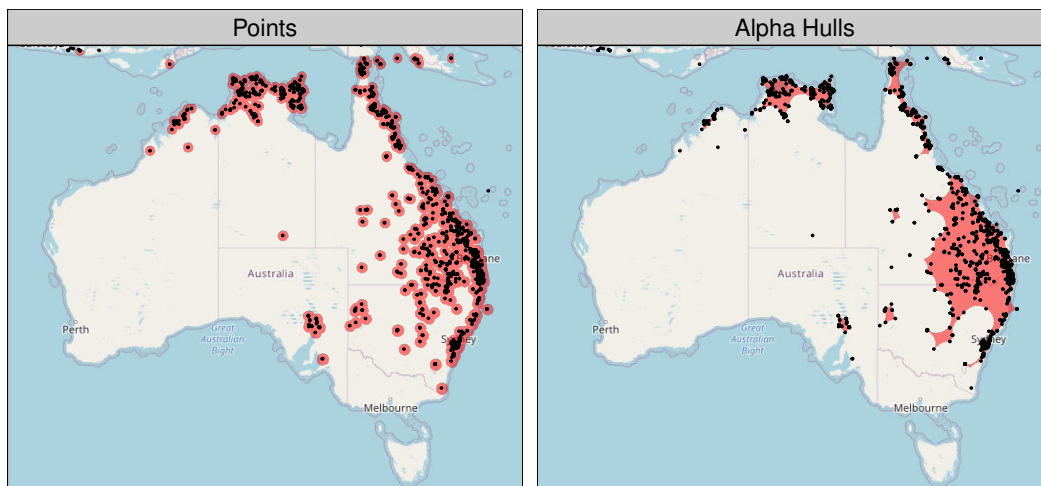
Below we provide a brief example of using `rasterize_range` in R, and an example of the resulting distribution in Figure 2.4.

```
# Create a 1km Australian Albers template raster
# for the extent of Australia
template <- raster::raster(raster::extent(-1888000, 2122000, -4847000, -1010000),
                           res=1000,
                           crs="+init=epsg:3577")

# Load occurrence records into R
occ <- readRDS("data/native_citrus.rds")

# Use points method with buffer of 50km
pt <- edmaps::rasterize_range(xy = occ,
                             method = "points",
                             template = template,
                             point_buffer=50000)

# Use alpha hull method with alpha set at 130,000
hull <- edmaps::rasterize_range(xy = occ,
                                method = "alphahull",
                                template = template,
                                alpha=130000)
```



**Figure 2.4.:** Estimated distribution of native citrus species using points with 50km buffer (left) and alpha hull method (alpha = 130,000, right) on an Australian Albers raster. Black dots signify occurrence records, derived from GBIF, for 9 Australian native citrus species (*Murraya paniculata*, *Murraya koenigii*, *Citrus glauca*, *Citrus australasica*, *Citrus australis*, *Citrus inodora*, *Citrus garrawayi*, *Citrus maideniana*, *Citrus gracilis*). Red shaded areas give the estimated extent of native host distribution.

## 2.4. Modelling vectors and pathogens

The `edmaps` framework was designed to estimate the establishment potential of individual threats entering Australia. However, for some plant diseases, entry may occur via multiple vectors along varying pathways, and host preferences and distributions of climate suitability may differ. To assess the post-border establishment potential of a plant disease with multiple vectors such as *Xylella fastidiosa* and Huanglongbing disease (HLB), we modified `edmaps` to not only estimate the vector-specific establishment likelihood, but also the likelihood that one or more disease-carrying vectors may establish at a location.

Species	Species group	Pathways
Xylella vector <i>Cicadella viridis</i>	Xylella	Goods, Residents, Vessels
Xylella vector <i>Graphocephala atropunctata</i>	Xylella	Goods, Residents, Vessels
Xylella vector <i>Homalodisca vitripennis</i>	Xylella	Goods, Residents, Vessels
Xylella vector <i>Philaenus spumarius</i>	Xylella	Goods, Residents, Vessels
Xylella pathogen	Xylella	Nursery, Residents

**Figure 2.5.:** Specifying threat groupings in `edmaps parameters.xlsx`.

`edmaps` now allows users to assign individual threats to a threat group. This is done within the species tab of the `parameters.xlsx` file (Figure 2.5).

`edmaps` then uses this input to independently estimate the probability of disease establishment attributable to each vector,  $v$ , and each location,  $i$ :

$$\Pr(p_{\text{establishment}_{i,v}}) = \Pr(p_{\text{viable}_{i,v}}) \times \text{Abiotic suitability}_{i,v} \times \text{Biotic suitability}_{i,v}. \quad (2.16)$$

The conditional probability of establishment given that one or more vectors arrives and establishes is then estimated as:

$$\Pr(p_{\text{establishment}_i}) = 1 - \prod_{v=1}^n \left(1 - \Pr(p_{\text{establishment}_{i,v}})\right). \quad (2.17)$$

## 3. Using Zonation & edmaps outputs to inform multi-pest surveillance

### 3.1. Estimating threat-specific establishment potential using edmaps

We used the enhanced version of `edmaps` described in Chapter 2 to estimate 1 km resolution post-border establishment potential maps for 10 plant pests, 4 pathogens and 9 exotic pathogen vectors identified by the Department of Agriculture, Water and the Environment (DAWE) as national priority plant threats (Table 3.1, Figure 3.1).

In the following sections we provide a brief overview of the threat-specific pathway leakage and viability parameters used, and how abiotic and biotic suitability were estimated in `edmaps`. For more detail as to how `edmaps` works please consult [Camac \*et al.\* \(2021\)](#). For more comprehensive details of the parameters used in this analysis please examine the attached `parameters.xlsx` input file used in this analysis.

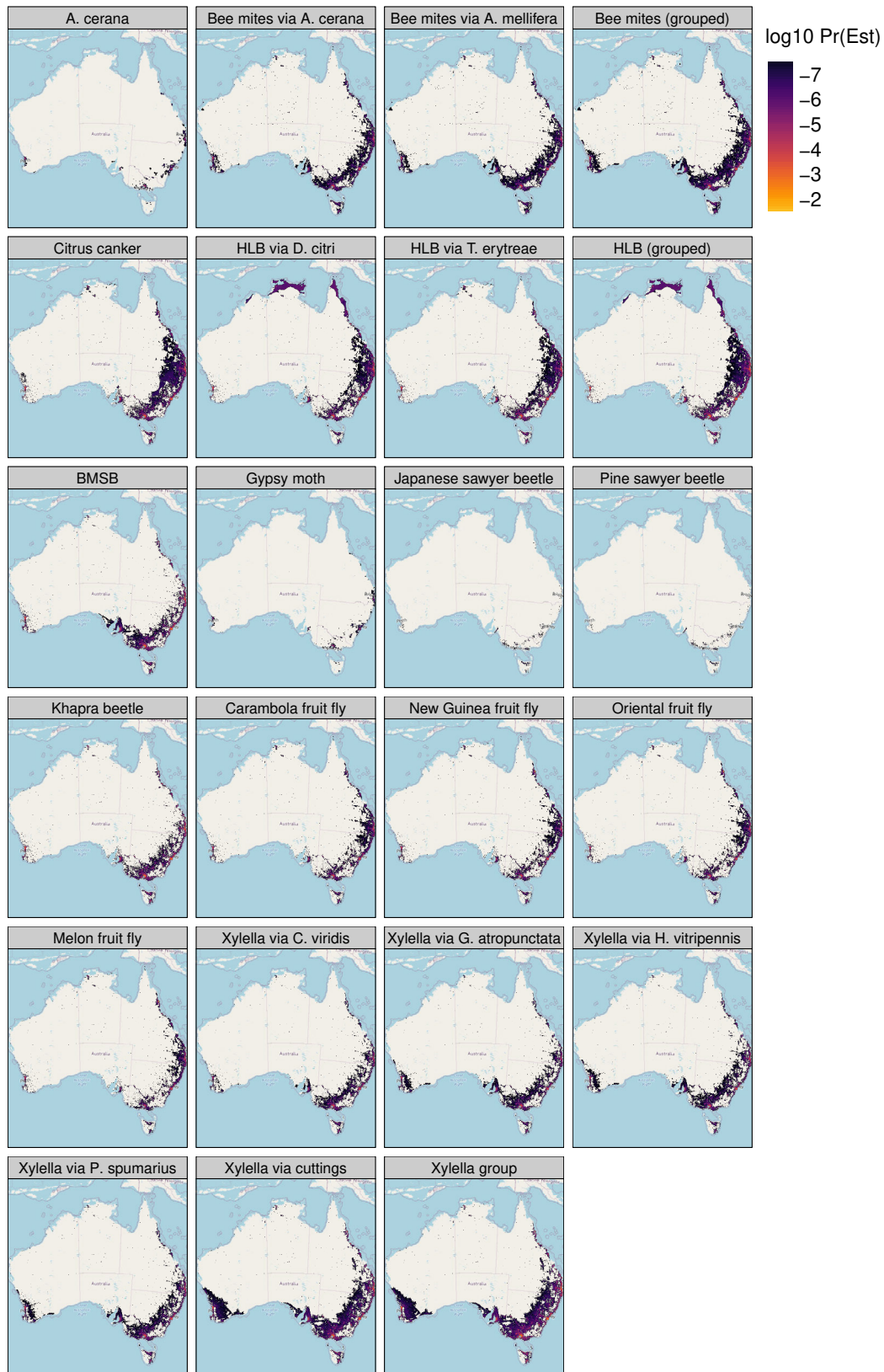
#### Pathway parameter inputs

`edmaps` currently allows for 13 pathways of entry: Air passengers (tourists, Torres Strait, returning residents), international vessels and mail, containers, imported nursery stock, machinery, fertiliser and food, and wind dispersal associated with northern winds, Pacific winds, New Zealand winds. For a comprehensive description of these pathways please consult [Camac \*et al.\* \(2021\)](#) and section 2.2 of this report. To distribute post-border risk associated with these pathways, `edmaps` requires threat-specific pathway bounds for both the annual entry ( $\text{Entry}_{low_k}$ ,  $\text{Entry}_{high_k}$ ) and the pathway viability ( $p_{low}$ ,  $p_{high}$ ; see Chapter 2). At the time of writing, DAWE was still in the process of estimating and verifying these bounds. As such, this analysis utilises preliminary estimates reported in Table 3.2. These estimates should be treated as illustrative only and not used to inform biosecurity risk management or surveillance design.

**Table 3.1.:** Summary details about threats modelled in this study. Climatic suitability for threats with sufficient occurrence records (sourced from GBIF or supplied by experts) was modelled using `edmaps` default range bagging model algorithm. Biotic suitability can be estimated using a combination of data sources including: the Australian Landuse and Management Classification raster (ALUMC), the National Vegetation Information System (NVIS), Normalized difference vegetation index (NDVI; a measure of vegetation productivity), and/or a user's own defined host layer (e.g. a native host layer as seen in Figure 2.4).

Threat	Threat type	Climate suitability	Biotic suitability
Asian honey Bee ( <i>Apis cerana</i> )	Exotic pest	Yes	NDVI
New Guinea fruit fly ( <i>Bactrocera trivialis</i> )	Exotic pest	Insufficient data	ALUMC, NDVI, Native citrus hosts
Bee mites via Asian honey bee ( <i>Apis cerana</i> )	Exotic bee mite vector	Indirectly by vector climate suitability	NDVI (measure of bee suitability)
Bee mites via European honey bee ( <i>Apis mellifera</i> )	Exotic bee mite vector	Indirectly by vector climate suitability	NDVI (measure of bee suitability)
Brown marmorated stink bug ( <i>Halyomorpha halys</i> )	Exotic pest	Yes	ALUMC, NDVI
Carambola fruit fly ( <i>Bactrocera carambolae</i> )	Exotic pest	No (insufficient data)	ALUMC, NDVI, Native citrus hosts
Citrus canker ( <i>Xanthomonas citri</i> )	Exotic disease	No (insufficient data)	ALUMC
Gypsy moth ( <i>Lymantria dispar</i> )	Exotic pest	Yes	ALUMC, NDVI, NVIS
Huanglongbing via Asiatic citrus psyllid ( <i>Diaphorina citri</i> )	Exotic disease vector	Yes	ALUMC (HLB hosts), Native citrus hosts
Huanglongbing via African citrus psyllid ( <i>Trioza erytreae</i> )	Exotic disease vector	No (insufficient data)	ALUMC (HLB hosts), Native citrus hosts
Japanese sawyer beetle ( <i>Monochamus alternatus</i> )	Exotic pest	No (insufficient data)	ALUMC, NDVI
Khapra beetle ( <i>Trogoderma granarium</i> )	Exotic pest	No (pest of stored goods)	ALUMC
Melon fly ( <i>Bactrocera cucurbitae</i> )	Exotic pest	Yes	ALUMC, NDVI, Native citrus hosts
Oriental fruit fly ( <i>Bactrocera dorsalis</i> )	Exotic pest	Yes	ALUMC, NDVI, Native citrus hosts
Pine sawyer beetle ( <i>Monochamus galloprovincialis</i> )	Exotic pest	Yes	ALUMC, NDVI
Xylella via green leafhopper ( <i>Cicadella viridis</i> )	Exotic disease vector	Indirectly by vector climate suitability	ALUMC (Xylella hosts)
Xylella via blue-green sharpshooter ( <i>Graphocephala atropunctata</i> )	Exotic disease vector	Indirectly by vector climate suitability	ALUMC (Xylella hosts)
Xylella via glassy-winged sharpshooter ( <i>Homalodisca vitripennis</i> )	Exotic disease vector	Indirectly by vector climate suitability	ALUMC (Xylella hosts)
Xylella via meadow frog hopper ( <i>Philaenus spumarius</i> )	Exotic disease vector	Indirectly by vector climate suitability	ALUMC (Xylella hosts)
Xylella via plant cuttings	Exotic disease	No	ALUMC (Xylella hosts)





**Figure 3.1.:**  $\log_{10}$  Estimated establishment likelihood of significant exotic plant pests, pathogens and pathogen vectors. Lighter colours = higher establishment likelihood, white = very low (<  $1e-8$ ) likelihoods of establishment. Establishment likelihood accounts for viable pathway leakage rates and the suitability of the abiotic and biotic environment. Pathogen groups are the aggregated establishment potential across vectors.



For the purposes of this analysis, we were required to make several assumptions about pathway post-border movement distances from points of entry. These assumptions were:

1. that the northern wind pathway provided an entry risk up to 100 km inland from the coast;
2. that 50% of pest/pathogen entries via vessels were likely to be dispersed (whether naturally or via hitch hiking on disembarked goods/passengers) within 1 km of a marine port;
3. that 50% of Torres Strait air passengers into Cairns would disperse within 10 km of the airport; and
4. that 50% of tourists would disperse within 200 km of an international airport within the first few days of arrival – when the risk of threat entry is greatest.

These distances are combined with other weighting functions (e.g. population density or tourist accommodation) and are approximations of post-border movements. We recommend DAWE utilises passenger card data to better parametrise these distance functions (Camac *et al.*, 2021), we also recommend additional modelling be conducted to determine how the risk associated with wind pathways is likely to disperse inland from the coast. For more detail about how these pathways distribute risk please consult both Camac *et al.* (2021) and Chapter 2. For full details of pathway-specific parameters used in this analysis please consult the user input file, `parameters.xlsx`.

Pathway:	Mail/Returning Resident		Tourists		Torres Strait passengers into Cairns		Machinery		Food		Goods	
Pathway description:	Risk attributable to returning air passengers or international mail (smuggled or unintentional)		Risk attributable to short term visitors. (Used if tourist risk different to returning residents)		Risk attributable to air passengers from Torres Strait into Cairns		Risk attributable to imported new and used cars & non-agricultural machinery		Risk attributable to imported fresh food		Risk attributable to imported goods	
Pathway distributed by:	Population density		Tourist Accommodation density & distance from Int. airport		Population density & distance from Cairns airport		Population density		Population density		Population density	
Threat	Entry bounds	Viability bounds	Entry bounds	Viability bounds	Entry bounds	Viability bounds	Entry bounds	Viability bounds	Entry bounds	Viability bounds	Entry bounds	Viability bounds
Asian honey Bee ( <i>Apis cerana</i> )												
New Guinea fruit fly ( <i>Bactrocera trivialis</i> )	3, 33	0.001, 0.01	7, 67	0.001, 0.01	1, 100	0.001, 0.01			1.489, 7.987	0.005, 0.05	0.5, 5	0.001, 0.01
Bee mites via Asian honey bee ( <i>Apis cerana</i> )	0.087, 0.365	0.1, 0.9							0.007, 0.04	0.001, 0.01		
Bee mites via European honey bee ( <i>Apis mellifera</i> )	0.087, 0.365	0.1, 0.9							0.007, 0.04	0.001, 0.01		
Brown marmorated stink bug ( <i>Halyomorpha halys</i> )	2, 20	0.005, 0.05					3, 30	0.01, 0.1			0.1, 1	0.01, 0.1
Carambola fruit fly ( <i>Bactrocera carambolae</i> )	3, 33	0.001, 0.01	7, 67	0.001, 0.01					1.489, 7.987	0.005, 0.05	0.5, 5	0.001, 0.01
Citrus canker ( <i>Xanthomonas citri</i> )	1.595, 6.426	0.02, 0.2							20, 200	1e-05, 1e-04		
Gypsy moth ( <i>Lymantria dispar</i> )							0.01, 0.1	0.002, 0.02				
Huanglongbing via Asiatic citrus psyllid ( <i>Diaphorina citri</i> )	0.892, 4.798	0.03, 0.3							5, 50	0.001, 0.01		
Huanglongbing via African citrus psyllid ( <i>Trioza erytreae</i> )	0.892, 4.798	0.03, 0.3							5, 50	0.001, 0.01		
Japanese sawyer beetle ( <i>Monochamus alternatus</i> )	1.488, 7.988	0.001, 0.01									15.692, 108.202	5e-04, 0.005
Khapra beetle ( <i>Tragoderma granarium</i> )	6.012, 51.169	0.002, 0.02							1.037, 5.639	0.01, 0.1	1.765, 7.164	0.005, 0.05
Melon fly ( <i>Bactrocera cucurbitae</i> )	3, 33	0.001, 0.01	7, 67	0.001, 0.01	1, 100	0.001, 0.01			1.489, 7.987	0.005, 0.05	0.5, 5	0.001, 0.01
Oriental fruit fly ( <i>Bactrocera dorsalis</i> )	3, 33	0.001, 0.01	7, 67	0.001, 0.01	1, 100	0.001, 0.01			1.489, 7.987	0.005, 0.05	0.5, 5	0.001, 0.01
Pine sawyer beetle ( <i>Monochamus galloprovincialis</i> )	1.488, 7.988	0.001, 0.01									15.692, 108.202	5e-04, 0.005
Xylella via green leafhopper ( <i>Cicadella viridis</i> )	1, 10	0.005, 0.05									3.608, 16.791	0.005, 0.05
Xylella via blue-green sharpshooter ( <i>Graphocephala atropunctata</i> )	1, 10	0.005, 0.05									3.608, 16.791	0.005, 0.05
Xylella via glassy-winged sharpshooter ( <i>Homalodisca vitripennis</i> )	1, 10	0.005, 0.05									3.608, 16.791	0.005, 0.05
Xylella via meadow frog hopper ( <i>Philaenus spumarius</i> )	1, 10	0.005, 0.05									3.608, 16.791	0.005, 0.05
Xylella via plant cuttings	1.888, 6.938	0.02, 0.2										

Pathway:	Fertiliser		Nursery		Vessels		Containers		North Wind	
Pathway description:	Risk attributable to imported fertiliser or agricultural equipment		Risk attributable to imported plants or plant parts		Risk attributable to international vessels docking at ports		Risk attributable to shipping containers		Risk attributable to winds from Indonesia and PNG	
Pathway distributed by:	Fertiliser usage statistics (Higher in production zones)		Population density		Port activity & distance from port		Container destinations		Wind frequencies from PNG and Indonesia	
Threat	Entry bounds	Viability bounds	Entry bounds	Viability bounds	Entry bounds	Viability bounds	Entry bounds	Viability bounds	Entry bounds	Viability bounds
Asian honey Bee ( <i>Apis cerana</i> )					0.04, 0.4	0.05, 0.5	0.03, 0.16	0.05, 0.5		
New Guinea fruit fly ( <i>Bactrocera trivialis</i> )										
Bee mites via Asian honey bee ( <i>Apis cerana</i> )					0.04, 0.4	0.05, 0.5	0.028, 0.172	0.05, 0.5		
Bee mites via European honey bee ( <i>Apis mellifera</i> )					0.04, 0.4	0.05, 0.5	0.028, 0.172	0.05, 0.5		
Brown marmorated stink bug ( <i>Halyomorpha halys</i> )					3, 30	0.01, 0.1	1, 10	0.01, 0.1		
Carambola fruit fly ( <i>Bactrocera carambolae</i> )										
Citrus canker ( <i>Xanthomonas citri</i> )	20, 200	1e-05, 1e-04								
Gypsy moth ( <i>Lymantria dispar</i> )					0.05, 0.5	0.01, 0.1	0.165, 0.774	0.005, 0.05		
Huanglongbing via Asiatic citrus psyllid ( <i>Diaphorina citri</i> )			0.3, 3	0.03, 0.3					0.3, 3	0.01, 0.1
Huanglongbing via African citrus psyllid ( <i>Trioza erytreae</i> )			0.3, 3	0.03, 0.3						
Japanese sawyer beetle ( <i>Monochamus alternatus</i> )							10, 100	1e-04, 0.001		
Khapra beetle ( <i>Tragoderma granarium</i> )							14.872, 79.898	5e-04, 0.005		
Melon fly ( <i>Bactrocera cucurbitae</i> )										
Oriental fruit fly ( <i>Bactrocera dorsalis</i> )										
Pine sawyer beetle ( <i>Monochamus galloprovincialis</i> )							10, 100	1e-04, 0.001		
Xylella via green leafhopper ( <i>Cicadella viridis</i> )					0.1, 1	0.01, 0.1				
Xylella via blue-green sharpshooter ( <i>Graphocephala atropunctata</i> )					0.1, 1	0.01, 0.1				
Xylella via glassy-winged sharpshooter ( <i>Homalodisca vitripennis</i> )					0.1, 1	0.01, 0.1				
Xylella via meadow frog hopper ( <i>Philaenus spumarius</i> )					0.1, 1	0.01, 0.1				
Xylella via plant cuttings			0.2, 2	0.02, 0.2						

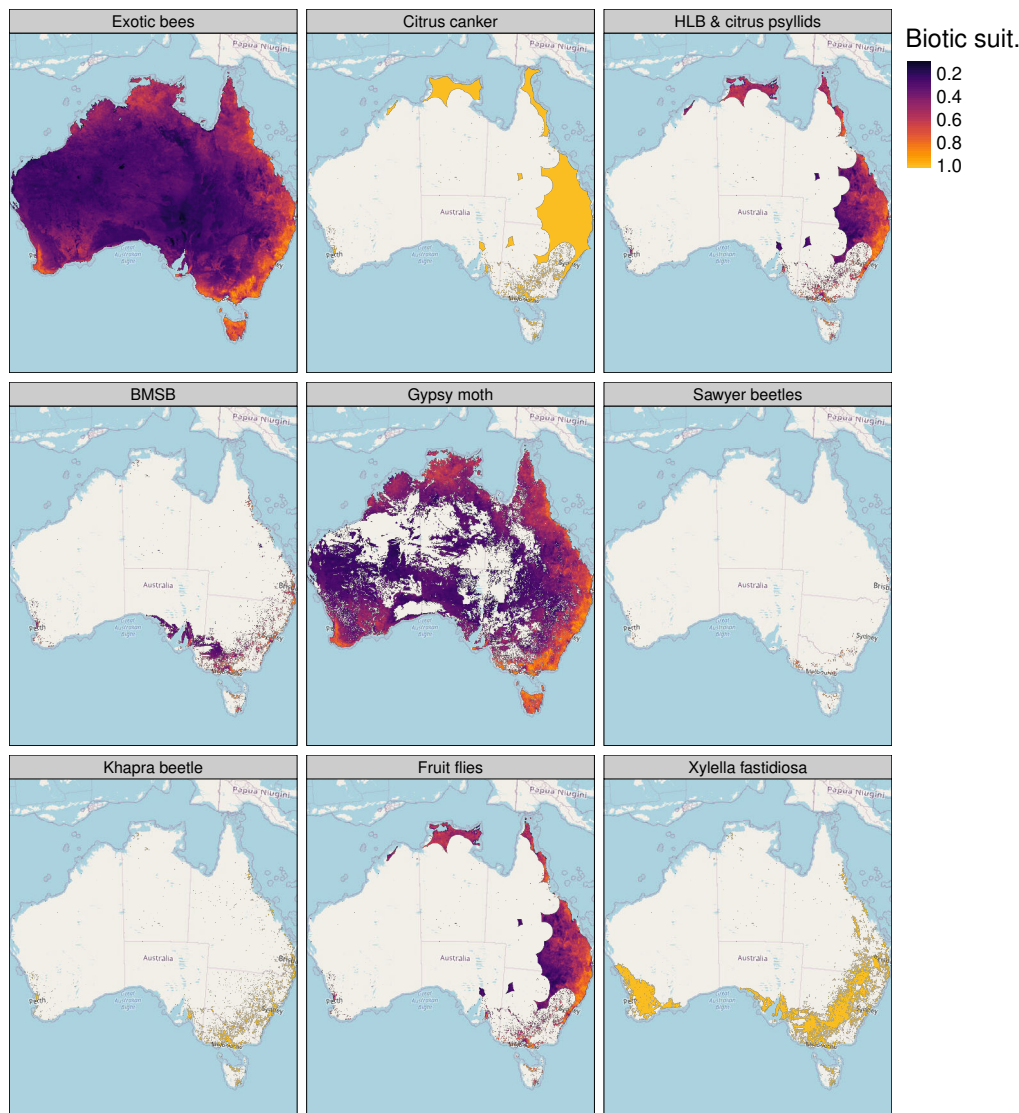
**Table 3.2.:** Summary of threat-specific leakage and viability bounds for pathways available in edmaps used in this analysis. For more details on how pathway risk is distributed post-border please consult [Camac et al. \(2021\)](#). The aggregated establishment potential of exotic pests, pathogens or pathogen vectors is estimated by combining individual vector edmaps using the method described in Section 2.4. **Important note:** The bounds included are preliminary (non-endorsed) estimates supplied by DAWE.

## Biotic suitability

The presence or absence of host material or suitable habitat is a major factor influencing the successful establishment of a pest or disease (Catford et al., 2009). In edmaps, the geographic distribution of host material can be informed using a variety of datasets. In this analysis, we utilised four:

1. Australian Land Use and Management Classification (ALUMC version 8; ABARES, 2019). ALUMC is the most comprehensive and up-to-date Australian landuse layer (current layer was compiled in December 2018) that classifies the dominant landuse of every  $50 \times 50$  m grid cell across Australia, at three levels of information resolution.
2. National Vegetation Information System; NVIS. NVIS is a  $5 \times 5$  km raster that maps the extent and distribution of vegetation types across Australia.
3. 6-monthly average Normalised Difference Vegetation Index; NDVI). NDVI provides a measure of the amount of live green vegetation, and can be coupled with other spatial host layers.
4. Australian occurrence records for 9 native citrus species (*Murraya paniculata*, *Murraya koenigii*, *Citrus glauca*, *Citrus australasica*, *Citrus australis*, *Citrus inodora*, *Citrus garrawayi*, *Citrus maideniana*, *Citrus gracilis*) derived from the Global Biodiversity Information Facility (GBIF). These data were then used to construct alpha hulls (see Figure 2.4) of native host distributions to supplement host distributions derived from ALUMC and NDVI for case studies of exotic fruit fly and citrus psyllids. Native species were proposed by the National Citrus Steering Committee.

A summary of which data sources were used to inform host distribution can be found in Table 3.1. For more specific detail about which landuse and NVIS codes were used, please consult the `parameters.xlsx` user input spreadsheet. Estimated biotic suitability maps for plant pests and disease vectors can be found in Figure 3.2.



**Figure 3.2.:** Estimated biotic suitability for significant exotic pests, pathogens or pathogen vectors. Lighter colours = higher biotic suitability, darker colours = low biotic suitability, white = areas of no biotic suitability (i.e. no hosts). Notes: 1) biotic suitability of disease vectors is based on the pathogen's susceptible host range; 2) Examination of European honey bee occurrence records highlighted that its distribution spanning most major Australian vegetation types. As such we used NDVI (i.e. a measure of vegetation greenness) to approximate host abundance, and thus biotic suitability; 3) While minor differences in host range exist among fruit fly species, these are indistinguishable at the national scale, thus we only show the host distribution for Oriental fruit fly.

## Abiotic suitability

If a climate suitability model is required (defined in `parameters.xlsx`), and the user does not supply a pre-built raster, `edmaps` will estimate climate suitability using range bagging (Drake, 2015) with dimensionality set to 2 (i.e. two covariates are fitted at a time), the number of bootstrapped models set to 100 and the proportion of occurrence records used per model set to 0.5. Unless otherwise specified, the range bagging algorithm will sample from all 19 WorldClim version 2 (Fick & Hijmans, 2017) bioclimatic parameters (i.e. BIO01 to BIO19) derived from the published 10 minute (approximately 20 km resolution) raster layers. The use of these “simple” two-dimensional combinations minimises biases associated with model over-fitting (see Camac et al. 2021).

In this analysis, DAWE did not supply CEBRA with any pre-built climate suitability maps. As such, where data was deemed sufficient (i.e. > 100 occurrences) we used range bagging for all threats requiring a climate suitability map (Table 3.1). Threat occurrence records were obtained from a variety of sources (Table 3.3) including expert curated published (e.g. Wang et al., 2019; Hill et al., 2017; Kriticos et al., 2017) and unpublished datasets (e.g. DAWE, NSW DPI) and the Global Biodiversity Information Facility (GBIF).

**Table 3.3.:** Source of occurrence records for threats with sufficient occurrence records requiring a climate suitability map

Threat	Data source
Asian honey Bee ( <i>Apis cerana</i> )	DAWE
European honey bee ( <i>Apis mellifera</i> )	GBIF
Brown marmorated stink bug ( <i>Halyomorpha halys</i> )	Kriticos et al. 2017
Carambola fruit fly ( <i>Bactrocera carambolae</i> )	GBIF
Gypsy moth ( <i>Lymantria dispar</i> )	GBIF
Asiatic citrus psyllid ( <i>Diaphorina citri</i> )	GBIF, NSW DPI, Wang et al. 2019
Melon fly ( <i>Zeugodacus cucurbitae</i> )	GBIF
Oriental fruit fly ( <i>Bactrocera dorsalis</i> )	GBIF, Hill et al. 2017
Pine sawyer beetle ( <i>Monochamus galloprovincialis</i> )	GBIF
Green leafhopper ( <i>Cicadella viridis</i> )	GBIF
<i>Graphocephala atropunctata</i> ( <i>Xylella</i> vector)	GBIF
<i>Homalodisca vitripennis</i> ( <i>Xylella</i> vector)	GBIF
<i>Philaenus spumarius</i> ( <i>Xylella</i> vector)	GBIF

Prior to running the range bagging algorithm, `edmaps` first cleans the occurrence records by using cleaning routines provided by the R package `CoordinateCleaner` (Zizka *et al.*, 2019). Specifically, `edmaps` removes records that:

1. have equal latitude and longitudes or are within 0.5 degrees radius of coordinates 0,0;
2. are within a 5 km radius of a capital city centroid;<sup>1</sup>
3. are within a 10 km radius of either a country or province centroid;
4. are within a 1 degree radius around the GBIF headquarters in Copenhagen, Denmark;
5. are within a 100 m radius around the centroids of known biodiversity institutions;
6. are located in the ocean; or
7. contain mismatches between the country specified in the record and the country corresponding to the record's coordinates.

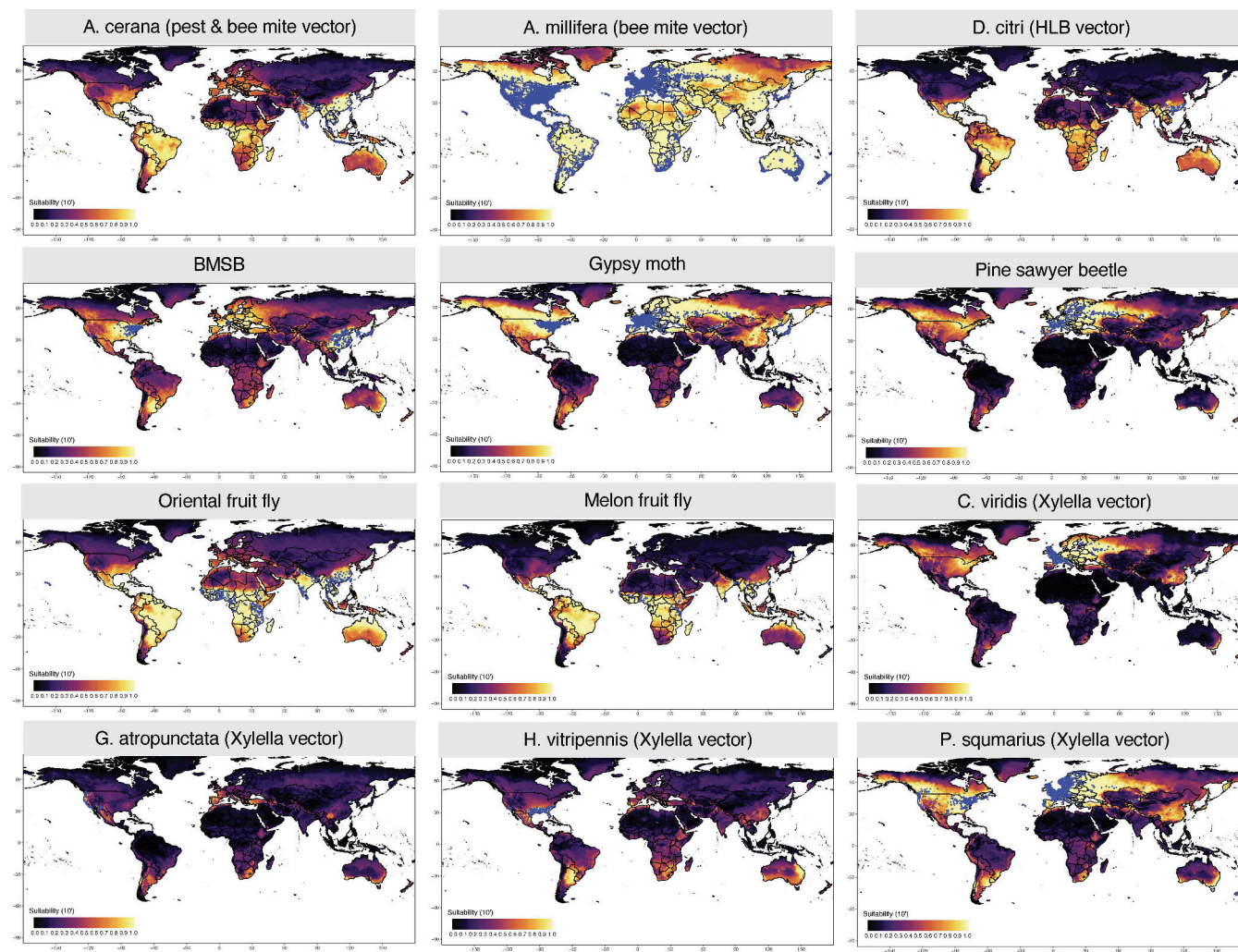
`edmaps` also removes duplicate records and thins the data to one point per 10 minutes (the resolution of the WorldClim version 2 climate data). Following this, if supplied by the user, `edmaps` will remove all occurrence records that are present in countries with no known establishment based on either a user supplied list of countries or distributional data found in a CABI datasheet<sup>2</sup>. This cleaning attempts to ensure that the remaining occurrences are most likely to be from established populations, and thus, suitable for inclusion in the range bagging analysis. In this analysis, range bagging models were fitted, and subsequently projected and resampled to a 1 km grid – our desired map resolution. The estimated climate suitability for threats are shown in Figures 3.3 & 3.4.

---

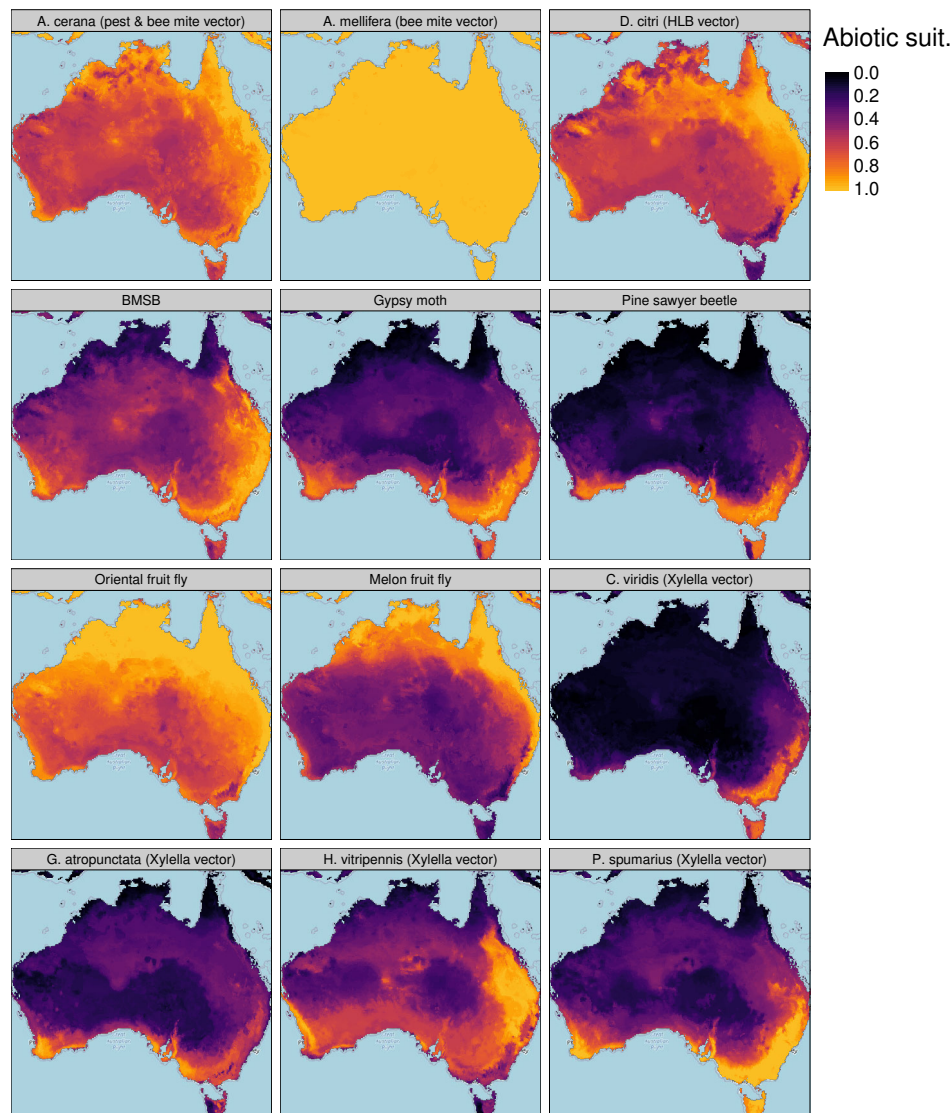
<sup>1</sup>In large databases such as GBIF, if a record does not have accurate GIS coordinates, users sometimes provide either country or capital city coordinates. While such coordinates are useful for providing information on the country or province of the record, they are likely to be too coarse when estimating the climatic suitability of a species, which involves mapping occurrence records onto fine-resolution gridded climatic data.

<sup>2</sup>Users may also wish to consult the Global Register of Introduced and Invasive Species (<http://www.griis.org/>; Pagad *et al.* (2018))





**Figure 3.3.:** Global climatic suitability for threats or threat vectors estimated from range bagging. Blue points indicate cleaned distributional data. Raster resolution: 10 minutes (approx 20 km). Suitability is the proportion of ensembled convex hulls that identify a location as climatically suitable across bootstrapped combinations of environmental variables. Lighter colours = higher climate suitability.



**Figure 3.4.:** Estimated Australian climate suitability for significant exotic pests, pathogens or pathogen vectors. Lighter colours = higher climate suitability. Note: some threats are not included due to insufficient data to estimate climate suitability. Climate suitability of pathogens is estimated indirectly via climate suitability of their vectors.



## 3.2. Using Zonation to identify priority surveillance locations across multiple biosecurity threats

Zonation contains two useful prioritisation functions – Core-Area Zonation (CAZ) and the Additive Benefit Function (ABF) – that can be used to inform multi-pest surveillance design. Each function begins by treating all cells across the landscape as candidates for surveillance. Cells are then ranked by removing, iteratively, the cell that contributes the least to the surveillance network. In the present context, this is defined as the cell whose removal leads to the minimum marginal loss in establishment potential (where marginal loss in establishment potential is given by  $\delta_i$  in Equations 3.1 & 3.2). While the two algorithms both rank cells according to their contribution to the optimal surveillance network, their definitions of optimality vary, as described below. For a comprehensive description of these functions, see Moilanen et al. (2005, 2014).

**Core-Area Zonation (CAZ)** is a prioritisation function that attempts to ensure representation of high establishment likelihood areas (i.e. *core areas*) for each threat in a multi-pest prioritisation. Raster cell removal is done by first identifying the proportional contribution of each cell to each threat’s total establishment potential, and then removing the cell for which the maximum contribution across threats (i.e. marginal loss in establishment potential upon removal,  $\delta_i$ ) is lowest. Surveillance costs and varying species weights are also accommodated. The quantity  $\delta_i$  is given by:

$$\delta_i = \max_j \frac{Q_{ij}w_j}{c_i}, \quad (3.1)$$

where  $w_j$  is the weight (or priority) of threat  $j$  and  $c_i$  is the cost of adding cell  $i$  to the surveillance network. The critical component of Equation 3.1 is  $Q_{ij}$ , which gives the establishment potential of threat  $j$  in cell  $i$ , as a proportion of the sum of establishment potential across the remaining (i.e. not yet removed) cells. That is,  $Q_{ij}$  describes the contribution of cell  $i$  to the total remaining establishment potential for threat  $j$ . When part of the distribution of a threat’s establishment likelihood is removed (via cell removal), the proportion located in each remaining cell increases. By taking the maximum across threats in Equation 3.1, Zonation ensures there is a strong preference for retaining locations of relatively high establishment potential. It is important to note that CAZ does not treat probabilities of establishment as additive; that is, having ten locations with  $p = 0.09$  is not the same as having one location with  $p = 0.9$ . However, this is strictly true only when a cost layer ( $c_i$ ) is not used in the analysis (or costs are equal across locations). When cost is used, cell removal is based on local establishment potential divided by the cost of performing surveillance in that cell. This means that a high Zonation rank for a cell can be explained by either (i) a very high establishment likelihood for some threat or (ii) low surveillance cost for the cell.

**Additive Benefit Function (ABF)**, unlike CAZ, considers all threat proportions in a given cell, rather than just the threat that has the highest proportional value. Using this algorithm, `Zonation` calculates the loss of representation for each threat due to removal of cell  $i$ , and estimates the corresponding loss of surveillance network value. Here,  $\delta_i$  is the sum over threat-specific declines in value following the loss of cell  $i$ :

$$\delta_i = \frac{\sum_j w_j \Delta V_j}{c_i}. \quad (3.2)$$

Here,  $\Delta V_j$  is the change in value for threat  $j$  due to removal of cell  $i$ , approximated internally by `Zonation` by an increasing function of representation. As before, the cost of surveillance at cell  $i$  is given by  $c_i$ , and the weight attributed to threat  $j$  is given by  $w_j$ .

Since ABF sums across all threats, the number of threats in a cell has a higher significance compared to core-area `Zonation`. For example, using ABF might lead to a situation where locations are removed even though they have a high establishment likelihood for one or two threats. This is because the  $\delta_i$  of these cells will be lower than competing cells that have high establishment likelihoods for many threats. Thus, ABF is likely to result in a surveillance design that has higher performance on average over all threats, but which captures a lower minimum proportion of establishment likelihood for the least likely (at a landscape scale) threats, compared to CAZ.

In this analysis, we examined the utility of both functions to inform multi-threat surveillance prioritisation. At the direction of DAWE, economic considerations were considered outside of scope, and as such threats were treated equally (i.e.  $w_j = 1$  for all threats) and spatially explicit surveillance costs,  $c_i$  were not included. As such, the outputs derived from this project only account for establishment potential of threats when prioritising surveillance across the landscape.

Using 1 km resolution (Australian Albers projection) `edmaps` establishment likelihood outputs for 12 individual pests/diseases and 2 aggregated pathogen groups (i.e. *Xylella*, bee mites) as `Zonation` features, we constructed multi-threat prioritisation maps using both CAZ and ABF functions at two spatial scales: 1) national; 2) metro-wide.

The national prioritisation maps allow one to prioritise multi-pest surveillance across the nation. By contrast, the metro-wide analysis prioritises multi-pest surveillance resources just within the limit of the metro-area without consideration of the establishment potential outside that area. In this analysis, metro-areas included: Sydney, Melbourne, Hobart, Adelaide, Perth, Darwin, Cairns and Brisbane, and were geographically delimited using Significant Urban Areas (SUA) boundaries ([ABS Cat: 1270.0.55.004](#)). For the national analysis, `Zonation` warp factors (i.e. how many cells are removed at a time) were set to 10,000. For the metro analysis, which contain significantly fewer raster cells, a warp factor of 1 was used. For both sets of analyses, edge removal was set to 0 (i.e. `Zonation` removes cells from anywhere and not just the edges of the remaining landscape as might be appropriate in reserve design).

For both the metro and national scales, we compared the ranking of priority locations derived from the multi-threat prioritisation maps derived from both `Zonation` algorithms to the threat-specific establishment likelihood maps derived directly from `edmaps`. We achieved this by first ranking all raster cells in each map in descending order according to their associated prioritisation scores (i.e. `Zonation` rank or establishment likelihood). We then calculated the percentage geographic overlap (i.e. agreement) between the two sets of maps as surveillance coverage increased.

`Zonation` 4.0.0 was run in a Docker container called from R via the `system()` function (see <https://github.com/wkmor1/build-zig4/blob/d35a6e6609f8c35b28e453976f3040c543e748cc/Dockerfile>).

### 3.3. Results: National Zonation outputs

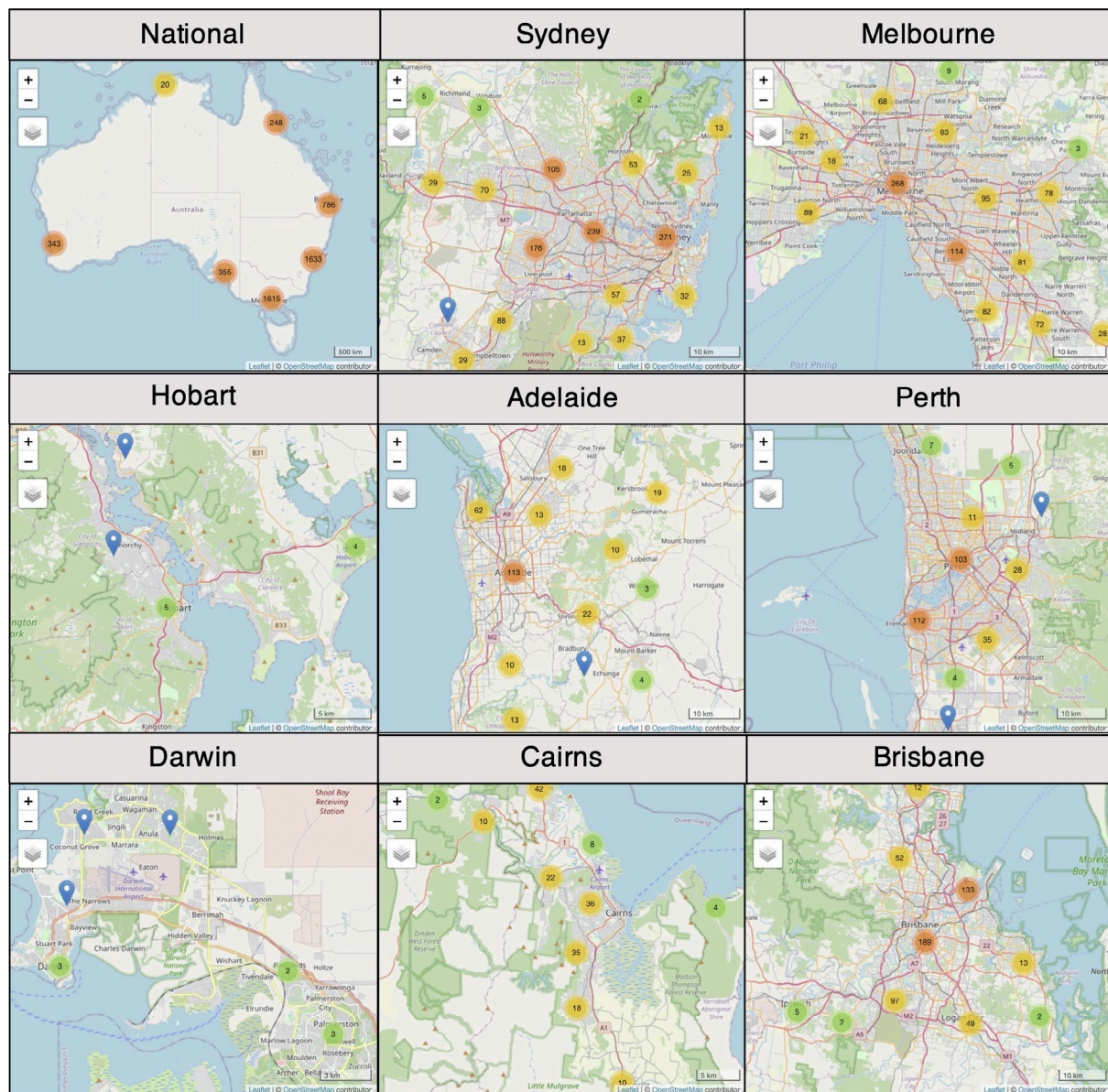
Examination of national Core-Area (CAZ) and Additive Benefit Function (ABF) Zonation outputs (Figures 3.5 & 3.6) revealed that across the 14 threats examined, Sydney, Melbourne and Brisbane contained approximately 80% (> 4000) of the top 5000 multi-threat prioritised locations, irrespective of the prioritisation function used. Closer examination of these metropolitan areas revealed that priority sites were typically situated in highly populated areas and close to either marine ports or airports. This finding is unsurprising as many of edmaps pathways are distributed post-border as a function of human population density and/or distance from key points of entry such as airports and marine ports (see Table 3.2).

For most threats, there is a marginal loss in establishment potential captured when surveillance locations are prioritised based on a multi-threat prioritisation relative to a threat-specific prioritisation (i.e. an individual establishment likelihood map produced by edmaps, whereby cells are ranked for surveillance for that threat according to its establishment likelihood; Figure 3.7). The most obvious exception to this were the two sawyer beetles, for which establishment distributions were poorly captured by the CAZ multi-threat prioritisation and not at all captured by the ABF multi-threat prioritisation. Two reasons for this poor coverage are:

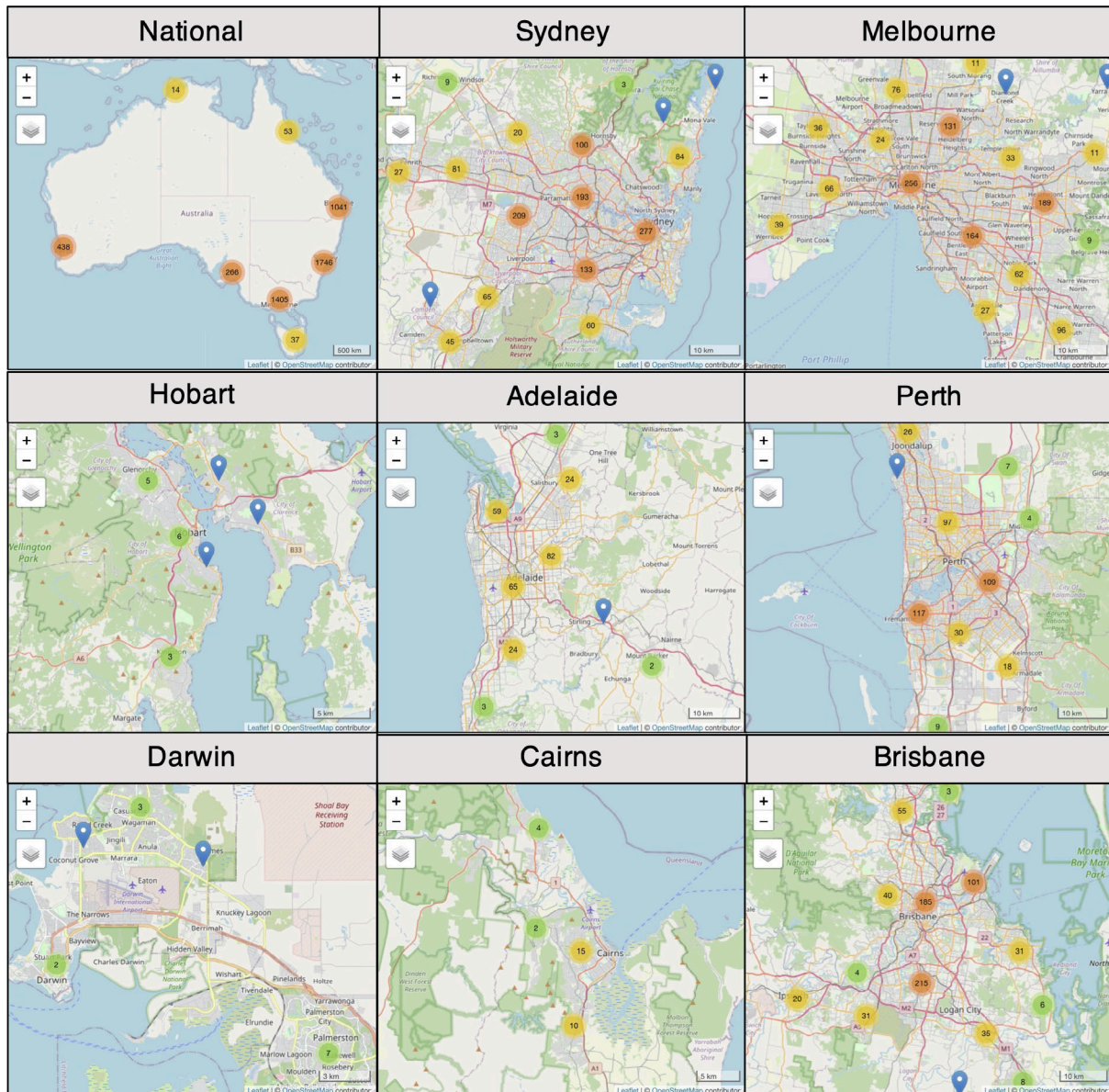
1. the sawyer beetles have a highly restricted geographic range of establishment potential (Figure 3.1) that is mostly governed by the presence of soft wood plantations outside metropolitan areas (Figure 3.2). Consequently, there is little geographic overlap with high establishment potential areas estimated for other threats, which tend to be higher within metropolitan areas due to the presence of host material (e.g. backyard fruit trees, food storage facilities, street trees or natives) coupled with high pathway-specific propagule pressure;
2. relative to other threats assessed in this study, the annual probability of a successful incursion by either sawyer beetle was estimated to be an order of magnitude or more lower than that for other threats.

The combination of these factors meant that Zonation's CAZ function – which attempts to ensure high risk establishment locations of all threats are represented – only prioritises the highest of the sawyer beetle establishment potential sites, often at the expense of higher risk locations associated with other threats. By contrast, the ABF function does not attempt to ensure threat representation, but rather maximises the average establishment potential captured across threats. Here, ABF's prioritisation is strongly influenced by threats that had a high annual establishment likelihood and had overlapping potential establishment distributions (Figures 3.7 & 3.8). As such, ABF performed well for threats such as BMSB, citrus canker, bee mites and HLB, but performed poorly for threats that had low establishment likelihoods, had high establishment potential in parts of the landscape that differed to the majority of the included threats (e.g. tropical fruit fly risk higher in northern Australia), or a combination of these factors.



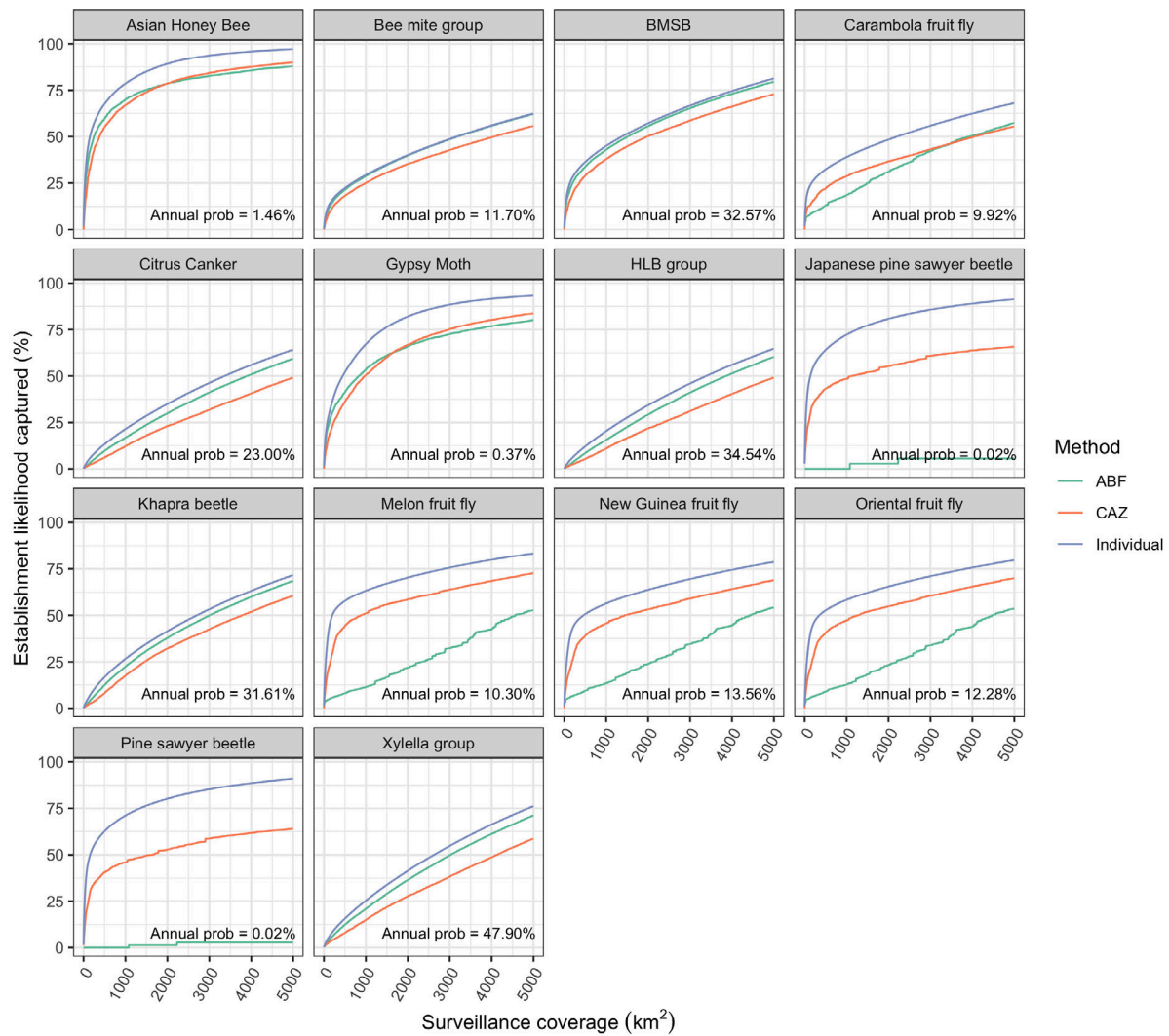


**Figure 3.5.:** The national top 5000 multi-threat priority surveillance locations as determined by Zonation’s Core-Area Zonation (CAZ) prioritisation. Top left panel shows the approximate national distribution of the top 5000 1 km<sup>2</sup> priority locations. All other panels are priority locations for major metropolitan areas. Coloured circles with numbers signify the number of priority 1 km<sup>2</sup> raster cells present within an area (warmer colours = higher number), blue markers signify single locations. Note that flagged locations are approximate – to view the exact distribution of prioritised sites, see the interactive maps associated with this report.

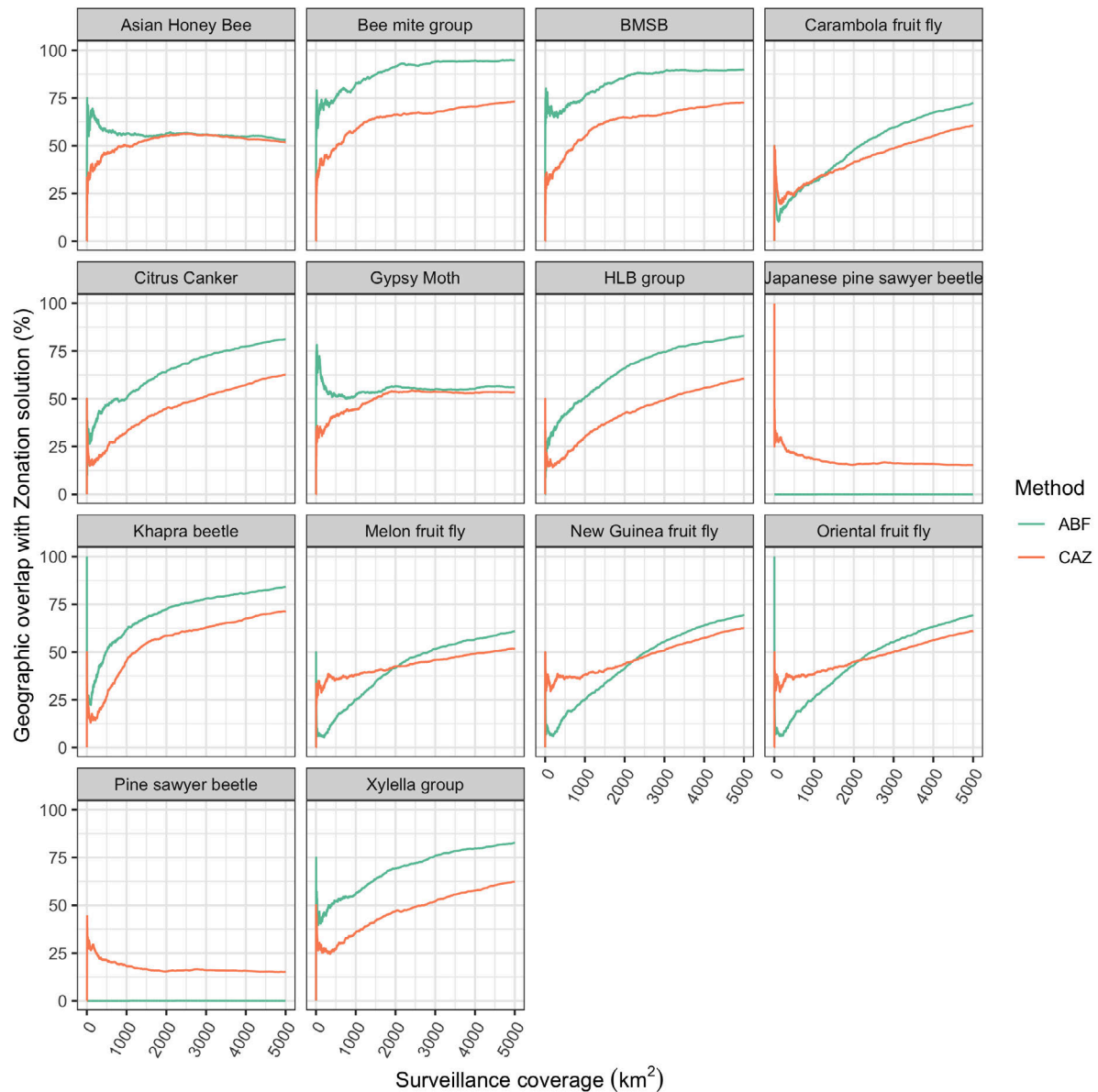


**Figure 3.6.:** The national top 5000 multi-threat priority surveillance locations as determined by Zonation's Additive Benefit Function (ABF). Top left panel shows the approximate national distribution of the top 5000 1 km<sup>2</sup> priority locations. All other panels are priority locations for major metropolitan areas. Coloured circles with numbers signify the number of priority 1 km<sup>2</sup> raster cells present within an area (warmer colours = higher number), blue markers signify single locations. Note that flagged locations are approximate.





**Figure 3.7.:** Cumulative percentage of national establishment potential captured as surveillance coverage increases, based on individual and multi-threat prioritisation methods. Blue line = prioritisation using individual threat map produced by edmaps; Orange line = Core-Area Zonation (CAZ) multi-threat prioritisation; Green line = Additive Benefit Function (ABF) multi-threat prioritisation. Annual prob is the estimated national annual probability of establishment after accounting for viable leakages across all identified pathways of entry and weighting by abiotic and biotic suitability.

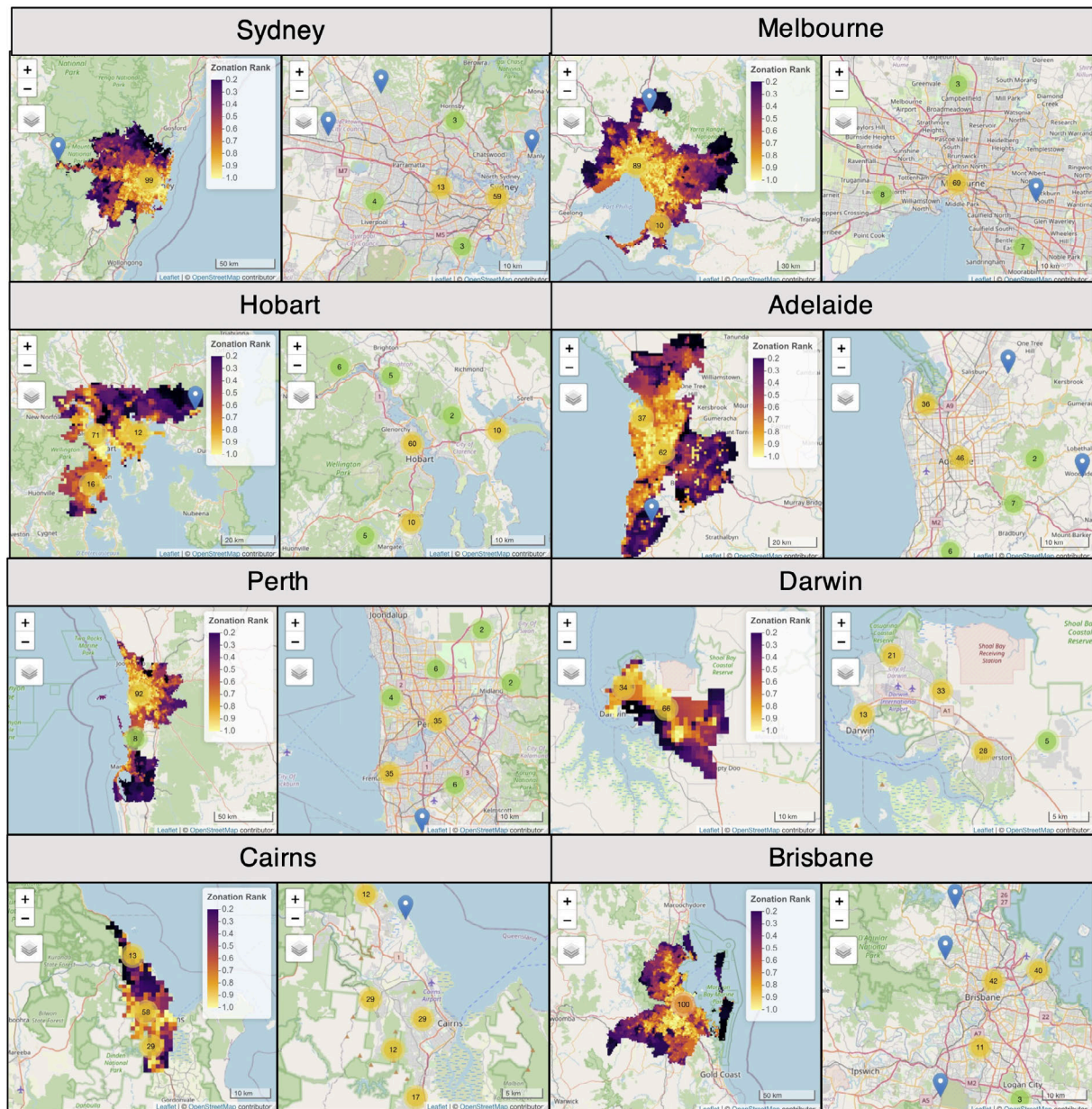


**Figure 3.8.:** Percentage of national geographic overlap between the set of cells included in the optimal surveillance network when using threat-specific edmaps output, and the set of cells included when using either Core-Area Zonation (CAZ) or Additive Benefit Function (ABF) multi-threat prioritisation, with increasing area of surveillance coverage. Orange line = CAZ; Green line = ABF.



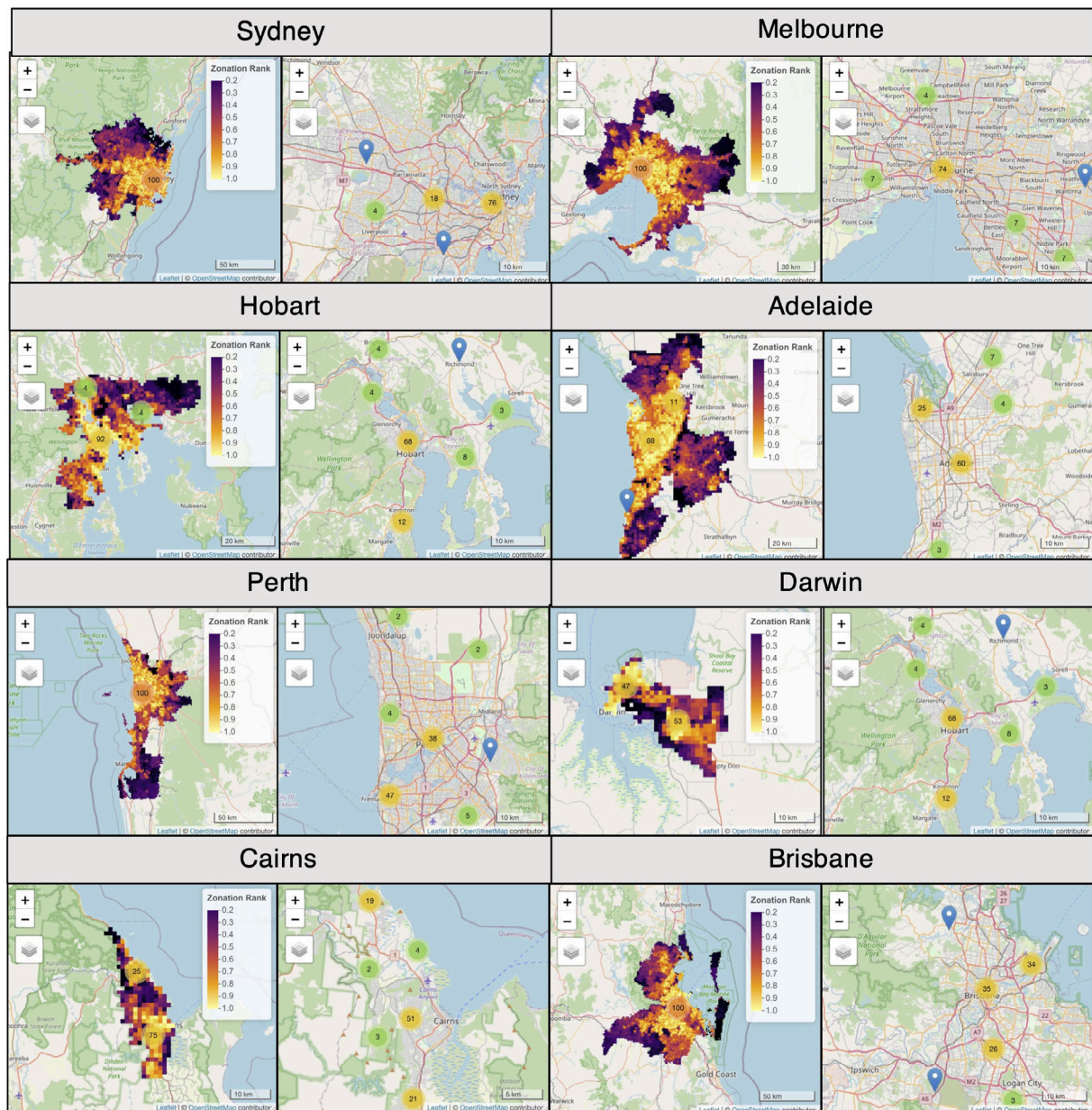
### 3.4. Results: Metropolitan Zonation outputs

Using `Zonation` to prioritise multi-threat surveillance locations independently for each major Australian metropolitan area produced broadly similar findings to the national analysis (Figures 3.9 & 3.10). This was to be expected as metropolitan areas contain the highest establishment likelihoods for most threats due to the combination of being close to points of entry, and thus, experiencing higher pathway propagule pressure, coupled with having suitable climate and host material (Figure 3.1). Minor changes in priority locations between the national and metro analyses were mostly driven by idiosyncratic differences in threat-specific geographic distributions in establishment potential. Interestingly, because establishment potential is highly consolidated in metropolitan areas, we found that ABF and CAZ prioritisation function performed similarly in terms of both the cumulative establishment likelihood captured with increasing surveillance area (Figures 3.11 & 3.13) and the percentage overlap with threat-specific maps (Figures 3.12 & 3.14). Indeed, for most threats, the multi-threat prioritisation methods performed almost as well as the threat-specific prioritisation.

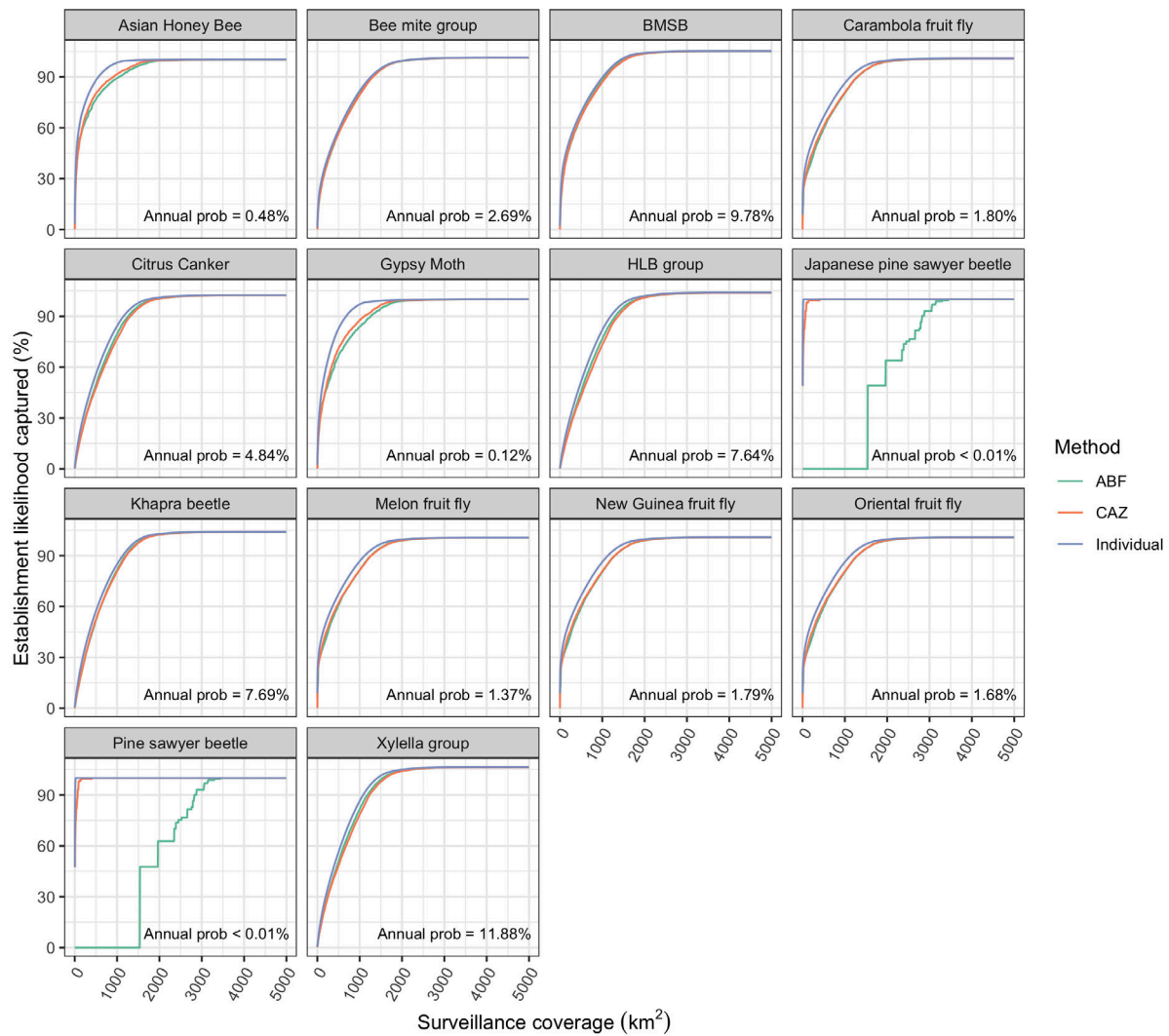


**Figure 3.9.:** Top 100 surveillance locations for major Australian metropolitan areas as determined by Core-Area Zonation (CAZ) multi-threat prioritisation. Left panels for each metropolitan area show the geographic extent of the metropolitan area and distribution of CAZ rankings, scaled to the range 0–1. Right panels show the geographic distribution of high priority surveillance locations. Warmer colours = higher priority rank (left panels) or higher number of priority sites in the vicinity (right panels), blue markers signify single locations. Note that flagged locations are approximate.

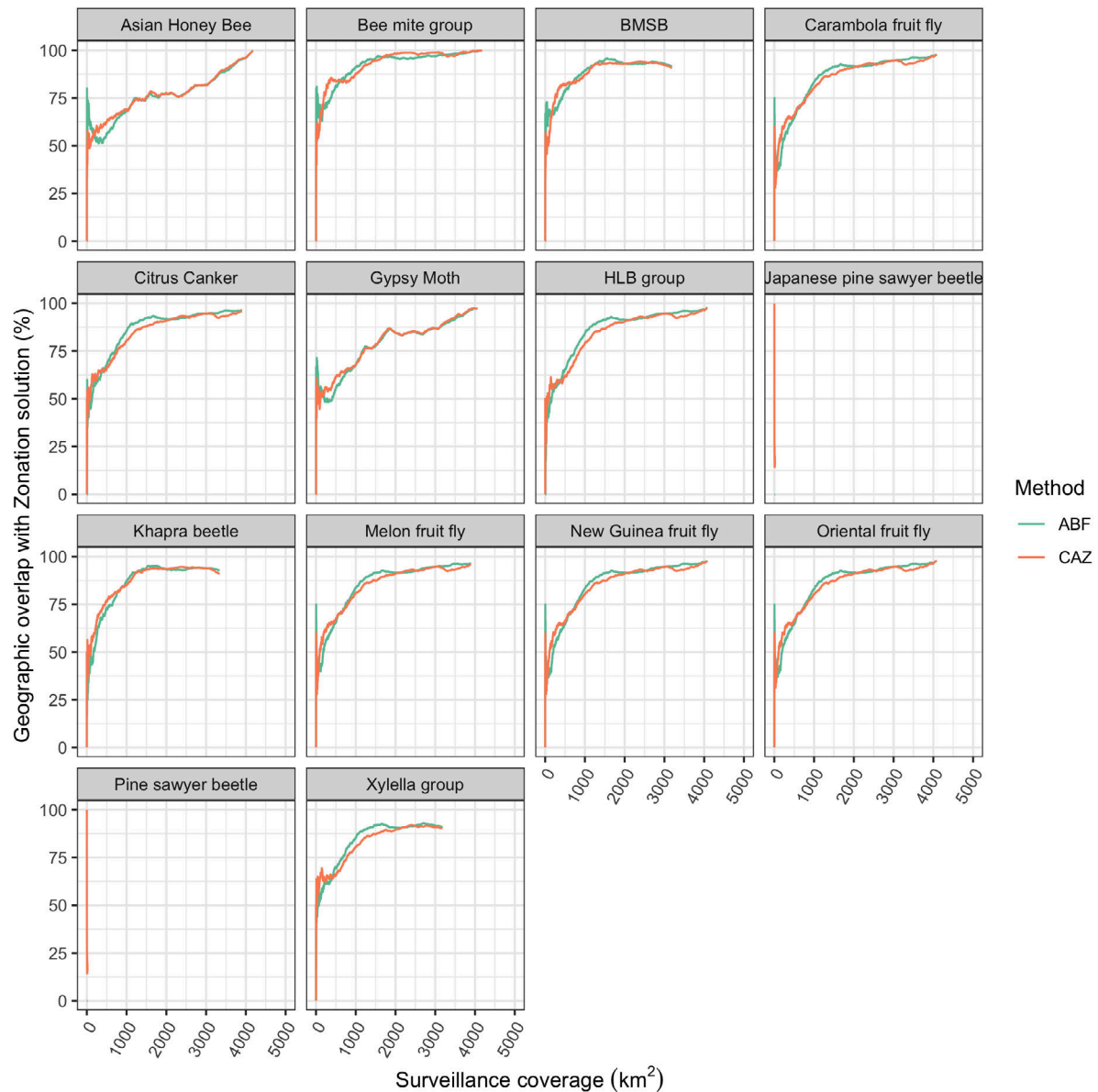




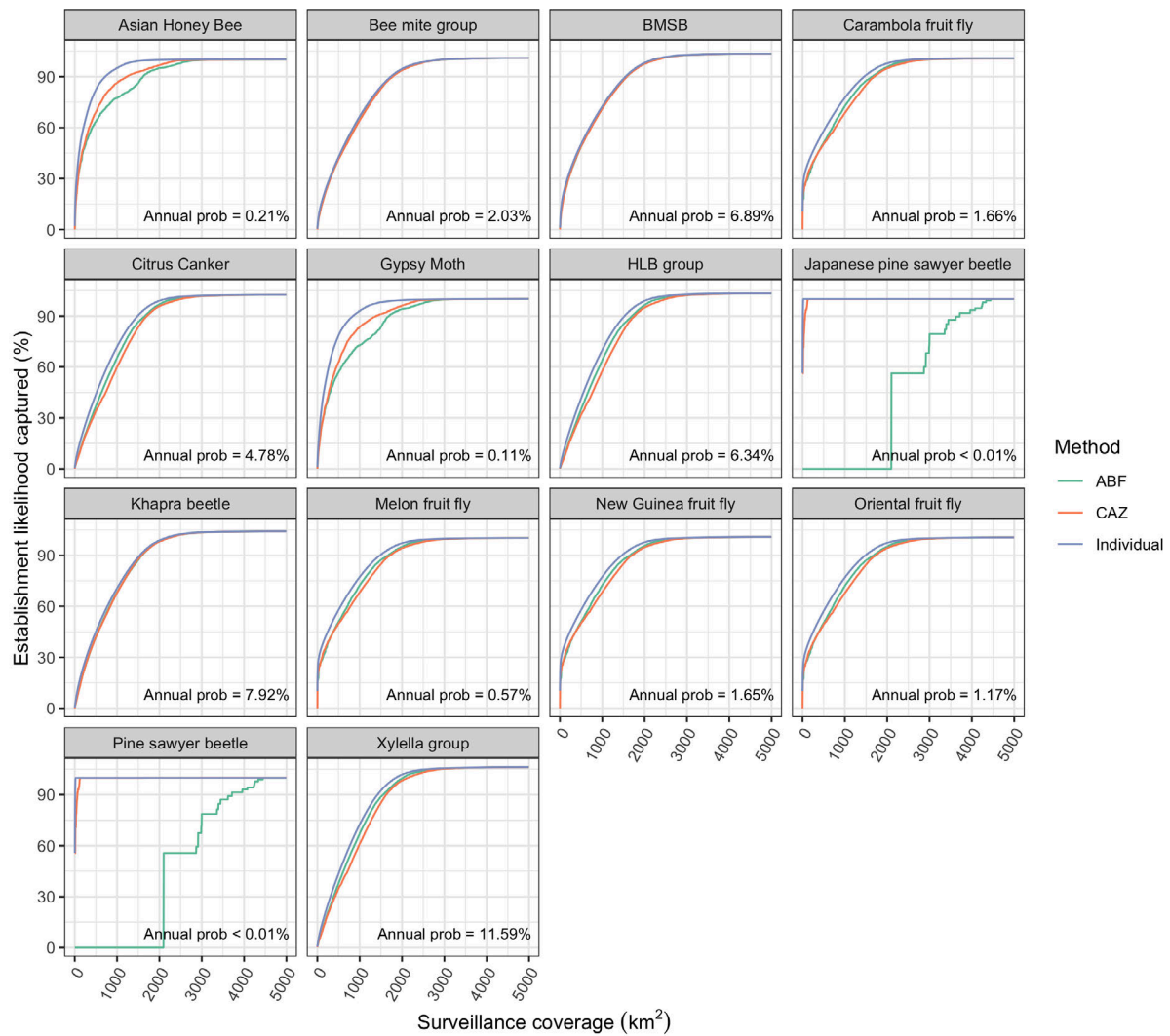
**Figure 3.10.:** Top 100 surveillance locations for major Australian metropolitan areas as determined by Additive Benefit Function (ABF) multi-threat prioritisation. Left panels for each metropolitan area show the geographic extent of the metropolitan area and distribution of ABF rankings, scaled to 0–1. Right panels show the geographic distribution of high priority surveillance locations. Warmer colours = higher priority rank (left panels) or higher number of priority sites in the vicinity (right panels), blue markers signify single locations. Note that flagged locations are approximate.



**Figure 3.11.:** Cumulative percentage of establishment potential captured in Sydney as surveillance coverage increases based on individual and multi-threat prioritisation methods. Blue line = prioritisation using individual threat map produced by edmaps; Orange line = Core-Area Zonation (CAZ) multi-threat prioritisation; Green line = Additive Benefit Function (ABF) multi-threat prioritisation. Annual prob is the estimated annual probability of establishment in Sydney after accounting for viable leakages across all identified pathways of entry and weighting by abiotic and biotic suitability.

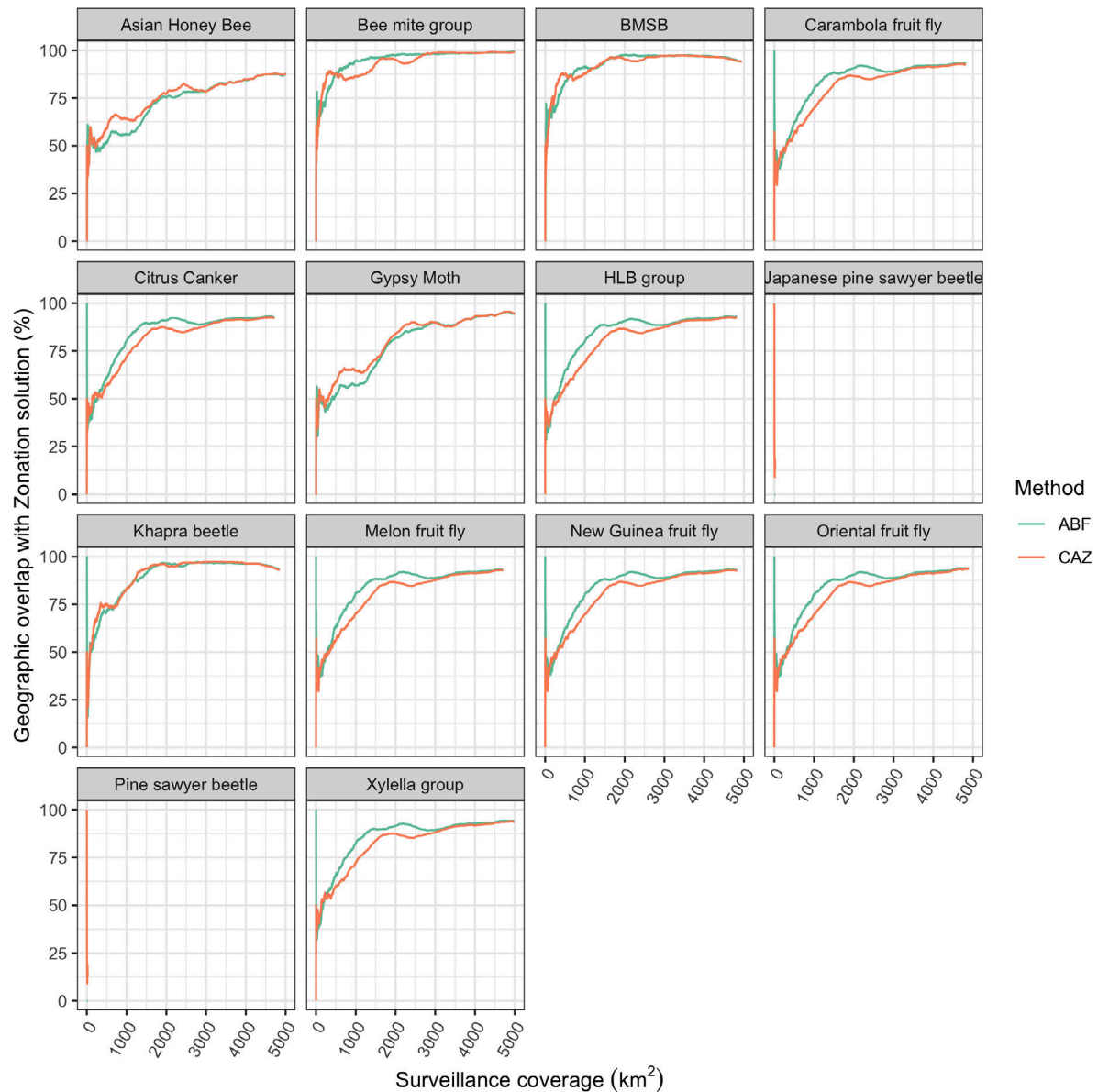


**Figure 3.12.:** Percentage of geographic overlap between the set of cells included in the optimal surveillance network for metropolitan Sydney when using threat-specific edmaps output, and the set of cells included when using either Core-Area Zonation (CAZ) or Additive Benefit Function (ABF) multi-threat prioritisation, with increasing area of surveillance coverage. Orange line = CAZ; Green line = ABF.



**Figure 3.13.:** Cumulative percentage of establishment potential captured in Melbourne as surveillance coverage increases based on individual and multi-threat prioritisation methods. Blue line = prioritisation using individual threat map produced by edmaps; Orange line = Core-Area Zonation (CAZ) multi-threat prioritisation; Green line = Additive Benefit Function (ABF) multi-threat prioritisation. Annual prob is the estimated annual probability of establishment in Melbourne after accounting for viable leakages across all identified pathways of entry and weighting by abiotic and biotic suitability.





**Figure 3.14.:** Percentage of geographic overlap between the set of cells included in the optimal surveillance network for metropolitan Melbourne when using threat-specific edmaps output, and the set of cells included when using either Core-Area Zonation (CAZ) or Additive Benefit Function (ABF) multi-threat prioritisation, with increasing area of surveillance coverage. Orange line = Core-Area Zonation (CAZ) multi-threat prioritisation; Green line = Additive Benefit Function (ABF) multi-threat prioritisation.



## 4. Conclusions, Caveats & Recommendations

### 4.1. Conclusions

Our analyses indicate that `Zonation` can be used with threat-specific establishment likelihood maps, produced by `edmaps`, to inform multi-threat surveillance prioritisation. We found that by moving from a threat-specific prioritisation (i.e. using individual threat establishment likelihood maps) to a multi-threat prioritisation, one could geographically consolidate surveillance with only a marginal reduction in the threat-specific establishment likelihood captured per unit area surveyed (Figures 3.7, 3.11 & 3.13). This was particularly effective where threats share similar geographic distributions of establishment likelihoods.

In this report we highlighted two prioritisation functions – Core-Area Zonation (CAZ) and the Additive Benefit Function (ABF). We believe both are well suited to informing where to conduct early-detection surveillance for multiple threats. Which of the two should be used depends on the objective of the end-user. Core-area Zonation is most appropriate when there is a definite set of threats, all of which must be accounted for in the surveillance design, and when tradeoffs between threats are discouraged. In this method, priority is given to core-areas (i.e. locations with the highest establishment potential for each threat). By contrast, the Additive Benefit Function (ABF) aims to maximise the *average* establishment potential captured across threats. The ABF method is more appropriate when the threats are essentially surrogates or samples from a larger threat pool, or when tradeoffs between threats are desirable, whereby greater prioritisation is weighted towards areas of shared high establishment potential, and/or threats that contain significantly higher likelihoods of establishment. This means, that threats with low establishment likelihoods that do not share the same high risk areas associated with other threats included in the analysis (e.g. sawyer beetles), will likely be severely undersampled in an ABF prioritisation.

The magnitude of differences among ABF, CAZ and individual threat prioritisation methods ultimately depends on three factors: 1) the geographic extent of the analysis; 2) the geographic similarity in establishment potential within that extent; and 3) the magnitude of difference in establishment potential among threats within the extent. In our analysis, we found that, with the exception of sawyer beetles, the majority of threats exhibited similar geographic distributions of establishment potential within major metropolitan areas such as Brisbane, Sydney (Figure 3.11) and Melbourne (Figure 3.13). This is because these areas are major hubs of human activity, and thus, are close to major points of entry (e.g. marine ports and airports) and also receive the high-

est hitch-hiking propagule pressure from human movement (e.g. air passengers) and trade pathways. Moreover, for the suite of threats included in our analysis, these areas were also estimated to have suitable biotic (Figure 3.2) and abiotic (Figure 3.4) environments. As such, when prioritisation is undertaken within individual metropolitan areas, we found little difference in performance among prioritisation methods (e.g. Figures 3.11 & 3.13). By contrast, when the prioritisation is performed at the national scale, greater geographic variability in post-border pathway movements, biotic suitability and abiotic suitability must be accounted for, which ultimately translates into greater discrepancies among methods (Figure 3.7) because each has a different objective and consequently handles this variability differently.

## 4.2. Caveats & Recommendations

The analysis presented in this report is based on a variety of model assumptions and data inputs which require further examination prior to any findings being used to implement new surveillance designs or decision-making. We briefly outline the key caveats & recommendations below.

### Pathway inputs & model outputs

At the time of writing, critical pathway inputs (i.e. pathway leakage and viability bounds) required by *edmaps* to estimate viable threat propagule pressure and consequently establishment potential, were still being estimated by the Department of Agriculture, Water and the Environment. As a consequence, this project used preliminary approximations to illustrate the *edmaps* work flow and outputs available to practitioners. The outputs in this report should therefore be treated with caution as data inputs and model assumptions have not been rigorously examined by taxon and biosecurity surveillance experts.

#### **Recommendation 1:**

**Establish expert task force for *edmaps* model interrogation, pathway parameter estimation and model implementation/use in surveillance**

To facilitate the estimation of pathway inputs and appropriate scrutiny of model outputs, we recommend the Department of Agriculture, Water and the Environment (DAWE) establish a task force consisting of experts in taxon biology, data analysis, pathway border interceptions, pre-border interventions, surveillance technologies (i.e. lure effectiveness, other logistics) and industry and biosecurity policy. Invariably, this means such a task force should have representatives from academia, industry and both commonwealth and state government agencies. These experts should have a clear understanding of the interpretation of each set of pathway inputs, how estimated inputs will be used in models such as *edmaps*, what model assumptions are made, and what data is available to estimate inputs. For threats or pathways where data are sparse, we recommend the task force approximate pathway inputs by examining data-rich

exemplars that share similar biological characteristics in combination with expert elicitation protocols such as the IDEA protocol (Hemming *et al.*, 2018). This task force will also be critical to the robust interrogation of model outputs (e.g. climate suitability, host availability and establishment potential), and thus, the identification of data/model gaps and opportunities for model improvement. To ensure model interrogation is easy, irrespective of GIS skill, edmaps by default produces intermediary (e.g. climate suitability models, host distribution models, pathway models) and final outputs (i.e. establishment potential maps) in a variety of formats including .tif for sophisticated interrogation in GIS software, .pdf for static map views, and .html for interactive map views in a web browser.

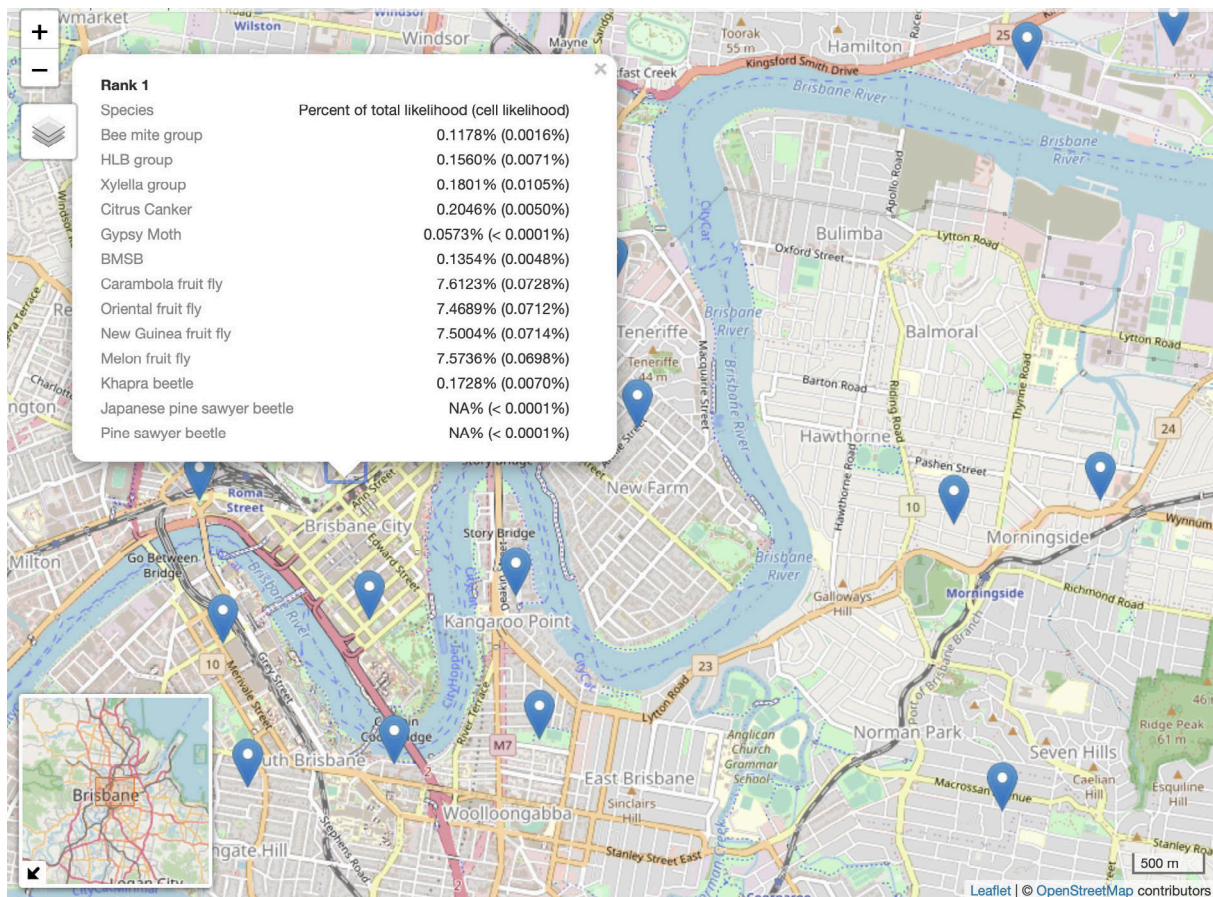
The integrated edmap and zonation workflow developed in this project also allows users to identify the top ranked priority locations at either across the nation or within a particular metropolitan area, which are then exported into a easy to use .html interactive map for further interrogation. These interactive maps not only show the geographic distribution of priority locations, but also allow the user to zoom in on a particular flagged location and obtain a threat-specific break down of the proportion of establishment potential captured if surveillance was undertaken at that location (Figure 4.1; Example interactive html appended to report). Such interactive maps further increase the capacity for biosecurity practitioners to carefully scrutinize model outputs prior to being used to inform strategic decisions about where to allocate surveillance resources.

## edmaps pathway models

Due to the limited data on post-border pathway movements, edmaps utilises a variety of simple and data-pragmatic models to approximate the dispersion of pathway-specific propagules entering the country (Camac *et al.*, 2021). While these models are grounded in common sense and use correlates that have support in the published scientific literature (e.g. human density Dodd *et al.*, 2015), additional research is required to validate these models, as well as review government and industry-held datasets that could be incorporated to better approximate pathway-specific post-border movements.

**Recommendation 2:**  
**Conduct regular reviews of pathway leakage parameters and pathway model assumptions**

We recommend DAWE regularly reviews the pathway models currently used in edmaps – ideally using the expert task force described in Recommendation 1 – to determine whether the pathway models and underlying data/assumptions are still fit-for-purpose. This review should conduct a Value of Information (VoI) analysis that identifies which data inputs contribute the most to prediction uncertainty, and thus, where governments and industry should prioritise data collection efforts. The review should also examine whether there are independent datasets or analyses (e.g. the SPEAR model; Mascaro & Woodberry 2020) available to either validate existing models and/or incorporate additional process-based complexity into pathway-specific post-



**Figure 4.1.:** Example of CAZ interactive map output highlighting priority surveillance locations within metropolitan Brisbane as well as threat-specific statistics for individual flagged locations. Users can click on any flagged location and see the proportion of the total establishment potential that occurs at that location as well as the estimated establishment likelihood (in brackets) for all threats used in multi-threat analysis. Note, flags represent the center of a 1 km raster cell.

border movements. [Camac et al. \(2021\)](#) outlined three examples of *low hanging fruit* that DAWE could capitalise on to improve pathway models in edmaps. The first is the utilisation of postcode data obtained from International Passenger Cards (IPCs) to better parametrise air passenger distance-decay functions from major international airports. The second is additional investment in post-border container movement analyses similar to the experimental analyses conducted by the Australian Bureau of Statistics ([ABS Cat. 1270.0.55.003](#)). The third is an up-to-date geo-referenced list of all Approved Arrangement sites coupled with information about the quantity of pathway-specific goods flowing to each. Lastly, we recommend additional modelling be conducted to determine how the risk associated with wind pathways is likely to disperse inland from the coast.



## Economic considerations & optimal surveillance design

The limited scope of this project meant that our analysis was unable to account for the economic considerations needed to determine the optimal early-detection design for either a fixed surveillance budget, or to minimise the net expected costs (i.e. summed cost of surveillance and potential outbreaks; [Camac \*et al.\* 2020](#)). As such, this project does not proscribe how many priority locations should be sampled, nor how much surveillance should be done at a given location. To answer questions such as how much surveillance should be done, or alternatively, how much should be invested in early-detection surveillance, one must have a sound understanding of the following five factors:

1. threat establishment likelihoods;
2. the magnitude of harm to environmental, economic and social assets;
3. surveillance efficacy (i.e. the sensitivity to detect a threat at either a tolerable population size or prevalence level per unit effort);
4. the cost of doing surveillance (i.e. cost of infrastructure, diagnostics and logistics); and
5. the likelihood a threat could spread beyond a point of establishment.

Each of these factors is likely to vary among threats and across geographic space, and recent advances in a variety of tools and data collection could be harnessed to improve prioritisation models. The `edmaps` framework provides a threat-specific estimate of how annual establishment likelihoods are likely to vary across mainland Australia. Recent work by [Dodd \*et al.\* \(2020\)](#) approximated spatially explicit estimates of environmental, agricultural and social assets and how various exotic threat groupings could impact each. [Bradhurst \*et al.\* \(2021\)](#) also recently developed a computationally efficient agent-based plant pest spread model that could be used to estimate likelihoods of subsequent spread, especially when used in combination with `edmaps` outputs. Spatially implicit costs of doing surveillance for different threats are likely known within biosecurity agencies (e.g. cost of infrastructure, logistics, diagnostic test). However, one could also account for spatially explicit costs such as travel time from different bases of operation and incorporate such information into a prioritisation algorithm.

When determining priority surveillance locations for multiple threats, `Zonation` provides an easy-to-use platform from which some of these considerations can be addressed. For example, `Zonation` allows users to weight threats differently (e.g. according to their potential magnitude of harm), which in turn, ensures a higher representation (in terms of establishment potential captured) of high impact threats relative to low impact threats. `Zonation` also allows users to input a spatially explicit cost layer, which can be used to account for logistical constraints such as travel time and/or varying costs in conducting surveillance in different landuse types. While not a full optimisation model, by incorporating these factors into `Zonation`'s prioritisation algorithms users can develop a more efficient and fit-for-purpose multi-threat surveillance prioritisation layer that not only acknowledges that threats vary in their magnitude of

harm on different assets, but also the cost constraints faced by biosecurity agencies in designing and maintaining surveillance networks.

**Recommendation 3:**

**Determine optimal surveillance effort by examining impacts of incorporating surveillance costs and potential impacts of threats**

We believe the next logical step is to examine how *Zonation* outputs change as threats are weighted according to estimated impacts, and costs such as travel time to either undertake surveillance and/or maintain infrastructure are considered within the prioritisation. Beyond this, additional work could be done to develop an optimisation model that more fully accounts for trade-offs associated with the benefits of threat early detection versus the potential costs of eradication as well as the realised costs of implementing and maintaining a comprehensive early detection surveillance program (e.g. [Nguyen \*et al.\*, 2021](#)). Such an analysis would be critical for determining the optimal number of surveillance locations that should be surveyed as well as the amount of effort required within each priority location to be confident of threat detection.

## Threat groupings

In our analysis, we found that for most threats, consolidating surveillance from a threat-specific to a multi-threat prioritisation resulted in only marginal reductions in expected establishment potential captured. The clear exception to this were the two sawyer beetles, which had locations susceptible to establishment that were poorly sampled by both CAZ and ABF prioritisation methods. As outlined earlier, the reason for this was that the two sawyer beetles had very low annual likelihoods of establishment, in areas that overlapped little with other threats. It is therefore important that when using *Zonation* to derive multi-threat prioritisation maps, one carefully considers the suite of threats that are included in the analysis, especially if the focus is to ensure all threats are adequately captured by the surveillance program.

**Recommendation 4:**

**When using *Zonation* to derive multi-threat prioritisation maps, users should always examine the proportion of establishment potential captured relative to using an individual threat map**

We recommend that when using *Zonation* to derive multi-threat prioritisation maps, users carefully examine the proportion of establishment potential captured using an individual threat map (i.e. produced by *edmaps*) vs. a multi-threat prioritisation map. Doing so will allow users to quickly identify which threats are under-represented in the surveillance design, and consequently, whether such threats should be removed from the prioritisation altogether.

# Acknowledgements

This report is a product of the Centre of Excellence for Biosecurity Risk Analysis (CE-BRA). In preparing this report, the authors acknowledge the financial and other support provided by the Australian Department of Agriculture, Water and the Environment and its predecessors, the New Zealand Ministry for Primary Industries and the University of Melbourne. We thank Prof Andrew Robinson for providing statistical guidance. We also thank Dr Mark Stanaway, Dr John Weiss, Dr Nerida Donovan and the National Citrus Surveillance Steering Committee for contributing their data and providing constructive advice throughout this project.



# Bibliography

- ABARES (2019) Catchment Scale Land Use of Australia – Update December 2018. URL <https://doi.org/10.25814/5c7728700fd2a>.
- Bradhurst R, Spring D, Stanaway M, Milner J, Kompas T (2021) A generalised and scalable framework for modelling incursions, surveillance and control of plant and environmental pests. *Environmental Modelling and Software*, **139**, 105004. doi:<https://doi.org/10.1016/j.envsoft.2021.105004>. URL <https://www.sciencedirect.com/science/article/pii/S1364815221000475>.
- Bradhurst R, Stanaway M, Milner J, Kompas T (2020) Developing models for the spread and management of national priority plant pests. Technical report prepared for the Department of Agriculture, Water and Environment. Tech. Rep. CEBRA project: 170606, Centre of Excellence for Biosecurity Risk Analysis.
- Camac JS, Baumgartner JB, Robinson A, Elith J (2021) Developing pragmatic maps of establishment likelihood for plant pests. Technical report prepared for the Department of Agriculture, Water and Environment. Tech. Rep. CEBRA project: 170607, Centre of Excellence for Biosecurity Risk Analysis.
- Camac JS, Clarke S, Niranjane A, de Majnik J (2019) Estimating probabilities of pest absence across geographic space. Technical report prepared for the Department of Agriculture, Water and Environment. Tech. Rep. CEBRA project: 1606D, Centre of Excellence for Biosecurity Risk Analysis.
- Camac JS, Dodd A, Bloomfield N, Robinson A (2020) Sampling to support claims of area freedom. Technical report prepared for the Department of Agriculture, Water and Environment. Tech. rep., Centre of Excellence for Biosecurity Risk Analysis.
- Catford JA, Jansson R, Nilsson C (2009) Reducing redundancy in invasion ecology by integrating hypotheses into a single theoretical framework. *Diversity and Distributions*, **15**, 22–40.
- Dodd A, Stoeckl N, Baumgartner J, Kompas T (2020) Key Result Summary: Valuing Australia's Biosecurity System. Technical report prepared for the Department of Agriculture, Water and Environment. . Tech. Rep. CEBRA Project: 170713, Centre of Excellence for Biosecurity Risk Analysis.
- Dodd AJ, McCarthy MA, Ainsworth N, Burgman MA (2015) Identifying hotspots of alien plant naturalisation in Australia: approaches and predictions. *Biological Invasions*, **18**, 631–645.
- Drake JM (2015) Range bagging: a new method for ecological niche modelling from presence-only data. *Journal of The Royal Society Interface*, **12**, 20150086–9.
- Fick SE, Hijmans RJ (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, **37**, 4302–4315.
- Finlay KJ, Weiss J, Blacket MJ (2018) Critical analysis of existing surveillance strategies in horticulture industries in Australia based on knowledge of natural dispersal pathways. Tech. Rep. PBCRC2153 Final Report, Plant Biosecurity CRC.
- Hemming V, Burgman MA, Hanea AM, McBride MF, Wintle BC (2018) A prac-

- tical guide to structured expert elicitation using the idea protocol. *Methods in Ecology and Evolution*, **9**, 169–180. doi:<https://doi.org/10.1111/2041-210X.12857>. URL <https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.12857>.
- Hill MP, Gallardo B, Terblanche JS (2017) A global assessment of climatic niche shifts and human influence in insect invasions. *Global Ecology and Biogeography*, **26**, 679–689.
- Kompas T, Che N (2009) A practical optimal surveillance measure: the case of papaya fruit fly in Australia.
- Kompas T, Ha P, Spring D (2017) Optimal Trapping and Review of Victoria’s Current Mediterranean Fruit Fly Surveillance Program. Report prepared for the Victorian Department of Economic Development, Jobs, Transport and Resources.
- Kriticos DJ, Kean JM, Phillips CB, Senay SD, Acosta H, Haye T (2017) The potential global distribution of the brown marmorated stink bug, *Halyomorpha halys*, a critical threat to plant biosecurity. *Journal of Pest Science*, **90**, 1–11.
- Kukkala AS, Moilanen A (2013) Core concepts of spatial prioritisation in systematic conservation planning. *Biological Reviews*, **88**, 443–464. doi:<https://doi.org/10.1111/brv.12008>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/brv.12008>.
- Mascaro S, Woodberry O (2020) Establishing a National Forest Biosecurity Surveillance Program: Forest Pest Pathway Risk Analysis Final Report.
- Moilanen A, Franco AM, Early RI, Fox R, Wintle B, Thomas CD (2005) Prioritizing multiple-use landscapes for conservation: methods for large multi-species planning problems. *Proceedings of the Royal Society B: Biological Sciences*, **272**, 1885–1891. doi:10.1098/rspb.2005.3164.
- Moilanen A, Kujala H, Leathwick J (2009) *The zonation framework and software for conservation prioritization*, pp. 196–210. Oxford University Press, United Kingdom.
- Moilanen A, Pouzol F, Meller L, Veatch V, Arponen A, Leppänen J, Kujala H (2014) *Zonation version 4 user manual*. C-BIG Conservation Biology, University of Helsinki, Finland.
- Nguyen HTM, Ha P, Kompas T (2021) Optimal surveillance against bioinvasions: The sample average approximation method applied to an agent-based spread model. *Ecological Applications*.
- Pagad S, Genovesi P, Carnevali L, Schigel D, McGeoch M (2018) Introducing the global register of introduced and invasive species. **5**, 170202.
- Pateiro-Lopez B, Rodriguez-Casal A, (2019) *alphahull: Generalization of the Convex Hull of a Sample of Points in the Plane*. URL <https://CRAN.R-project.org/package=alphahull>. R package version 2.2.
- Stanaway MA, Mengersen KL, Reeves R (2011) Hierarchical bayesian modelling of early detection surveillance for plant pest invasions. *Environmental and Ecological Statistics*, **18**, 569–591. doi:10.1007/s10651-010-0152-x.
- Wang R, Yang H, Wang M, Lu X, Huang T, Zhao J, Li Q (2019) Predicting the potential distribution of the Asian citrus psyllid, *Diaphorina citri* (Kuwayama), in China using the MaxEnt model. *PeerJ*, **7**, 1–19.
- Wilson K, Cabeza M, Klein C (2009) *Fundamental concepts of spatial conservation prioritization*, pp. 16–27. Oxford University Press, United Kingdom.
- Wintle B (2008) A review of biodiversity investment prioritization tools. A report to the biodiversity expert working group towards the development of the Investment

Framework for Environmental Resources.

- Wintle BA, Kujala H, Whitehead A, *et al.* (2019) Global synthesis of conservation studies reveals the importance of small habitat patches for biodiversity. *Proceedings of the National Academy of Sciences*, **116**, 909–914. doi:10.1073/pnas.1813051115.
- Zizka A, Antonelli A, Silvestro D (2021) sampbias, a method for quantifying geographic sampling biases in species distribution data. *Ecography*, **44**, 25–32. doi: <https://doi.org/10.1111/ecog.05102>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/ecog.05102>.
- Zizka A, Silvestro D, Andermann T, *et al.* (2019) CoordinateCleaner: Standardized cleaning of occurrence records from biological collection databases. *Methods in Ecology and Evolution*, **289**, 110–8.

## **A. Appendix A: Reference manual: edmaps R package**

# Package ‘edmaps’

May 22, 2021

**Type** Package

**Title** Estimate Likelihood of Pest Establishment

**Version** 1.9.3

**Maintainer** James Camac <james.camac@gmail.com>

**Description** This package combines spatial environmental data (e.g. distance from airports/marine ports, distribution of pest habitat/hosts) with biosecurity leakage rates (e.g. expert-elicited) to generate maps of the likelihood of pest establishment across the landscape. Final outputs include raster datasets (GTiff) indicating establishment likelihood across the area of interest, as well as interactive (html) and static (pdf) versions of these maps.

**Depends** R (>= 4.0)

**SystemRequirements** Java (>= 1.5), JRI, GNU make

**Imports** alphahull,  
countrycode,  
dplyr,  
drake,  
fasterize,  
furrr,  
gdalUtilities,  
geometry,  
ggplot2,  
glue,  
htmlwidgets,  
httr,  
leafem,  
leaflet,  
magrittr,  
mapedit,  
methods,  
purrr,  
raster,  
readr,  
readxl,  
rgbif,  
rlang,  
rnaturalearth,  
sf,

sp,  
stars,  
stplanr,  
tidyr,  
tmap,  
tmaptools

**Suggests** CoordinateCleaner,

knitr,  
rmarkdown,  
testthat,  
future.callr,  
rnaturalearthdata,  
lubridate,  
styler,  
OpenStreetMap,  
rnaturalearthhires

**VignetteBuilder** knitr

**License** MIT + file LICENSE

**Encoding** UTF-8

**LazyData** true

**RoxygenNote** 7.1.1

**ByteCompile** true

## R topics documented:

aggregate_raster . . . . .	3
arrivals_by_containers . . . . .	4
arrivals_by_fertiliser . . . . .	5
arrivals_by_food . . . . .	6
arrivals_by_goods . . . . .	7
arrivals_by_machinery . . . . .	8
arrivals_by_mail . . . . .	9
arrivals_by_nurserystock . . . . .	10
arrivals_by_residents . . . . .	11
arrivals_by_torres . . . . .	12
arrivals_by_tourists . . . . .	13
arrivals_by_vessels . . . . .	14
arrivals_by_wind . . . . .	15
binarize_and_aggregate . . . . .	16
calculate_median . . . . .	17
calc_EE . . . . .	17
calc_pathway_pr . . . . .	18
calc_proportion . . . . .	18
captured_by_ncells . . . . .	19
combine_arrivals . . . . .	19
combine_pathways . . . . .	20
container_weights . . . . .	20
download_worldclim2 . . . . .	21
establishment_likelihood . . . . .	22
excel_to_plan . . . . .	23

exp_function . . . . .	23
extract_highest_ncells . . . . .	24
extract_worldclim2 . . . . .	24
fertiliser_by_nrm . . . . .	25
fertiliser_weight . . . . .	26
fill_na . . . . .	26
gdal_reproject . . . . .	27
get_airport_dist . . . . .	29
get_gbif_records . . . . .	30
initialise_raster . . . . .	31
interactive_map . . . . .	32
max_normalize . . . . .	33
min_max_normalize . . . . .	34
na_mask . . . . .	34
ncells_4_threshold . . . . .	35
plan_globals . . . . .	35
plot_establishment_captured . . . . .	36
plot_raster . . . . .	37
port_weights . . . . .	38
pr_establishment . . . . .	39
range_bag . . . . .	39
rasterize_range . . . . .	40
rasterize_vector . . . . .	41
rasterize_wind . . . . .	42
record_flagger . . . . .	43
rle_compress . . . . .	44
species_plan . . . . .	44
static_map . . . . .	49
step_function . . . . .	50
suitability . . . . .	51
surveillance_coverage . . . . .	51
threshold_raster . . . . .	52
weight_airport_dist . . . . .	53

**Index****54**


---

aggregate_raster	<i>Downscale raster resolution</i>
------------------	------------------------------------

---

**Description**

Aggregate raster cells (and optionally layers) to coarser resolution.

**Usage**

```

aggregate_raster(
  rast,
  outfile,
  aggregate_factor,
  fun = sum,
  return_rast = FALSE
)

```



**Arguments**

rast	Raster* object or file path to a raster file.
outfile	Character. Output raster file path. If not provided, object will be returned to R. Directory will be created recursively if it does not exist.
aggregate_factor	Integer. Aggregation factor expressed as number of cells in each direction (horizontally and vertically). Or a vector of two integers (horizontal and vertical aggregation factors) or three integers (when also aggregating over layers).
fun	Function. Function used to aggregate values. Default is sum.
return_rast	Logical. Should the resulting raster be returned to R? Ignored if outfile is not provided.

**Value**

If outfile is specified, the resulting raster is saved as a geotiff to that path. If return\_rast is TRUE or outfile is not specified the resulting raster is returned, otherwise NULL is returned invisibly.

---

arrivals\_by\_containers

*Estimates pest arrivals by containers*


---

**Description**

Estimates pest arrivals by containers.

**Usage**

```
arrivals_by_containers(
  container_weights,
  port_data,
  template_raster,
  leakage_rate,
  establishment_rate,
  outfile,
  return_rast = FALSE,
  overwrite = FALSE
)
```

**Arguments**

container_weights	An <a href="#">sf</a> object or file path to a shapefile supported by OGR, as produced by <a href="#">container_weights</a> .
port_data	Character. Path to csv file containing Port Names, Port Codes, Longitude, Latitude and Container volumes. Column names must be (in this order): Name, PortCode, Longitude, Latitude, Count.
template_raster	RasterLayer or file path to a raster file. This is used to define the extent and resolution of output. Must be in CRS EPSG:3577.

leakage_rate	Numeric vector of 2 values, giving the lower and upper bounds of a 95 a random year).
establishment_rate	Numeric vector of 2 values, giving the lower and upper bounds of a 95 establishment, for leakage events).
outfile	Character. Output raster file path. If probability has length > 1, the file type must support multiband rasters (e.g. GeoTiff). If not provided, raster object will be returned to R.
return_rast	Logical. Should the raster object be returned to R? Ignored if outfile is not provided.
overwrite	Logical. If TRUE and outfile is not missing, it will be overwritten if the file specified by outfile already exists.

### Value

If outfile is specified, the resulting raster (multiband if probability has length > 1) is saved as a tiff at that path. If return\_rast is TRUE or outfile is not specified the resulting raster object is returned, otherwise NULL is returned invisibly.

### See Also

Other functions estimating arrivals: [arrivals\\_by\\_fertiliser\(\)](#), [arrivals\\_by\\_food\(\)](#), [arrivals\\_by\\_goods\(\)](#), [arrivals\\_by\\_machinery\(\)](#), [arrivals\\_by\\_mail\(\)](#), [arrivals\\_by\\_nurserystock\(\)](#), [arrivals\\_by\\_residents\(\)](#), [arrivals\\_by\\_torres\(\)](#), [arrivals\\_by\\_tourists\(\)](#), [arrivals\\_by\\_vessels\(\)](#), [arrivals\\_by\\_wind\(\)](#)

---

arrivals\_by\_fertiliser

*Estimates pest arrivals by fertiliser imports*

---

### Description

Estimates pest arrivals by fertiliser imports.

### Usage

```
arrivals_by_fertiliser(
  fertiliser_weight,
  leakage_rate,
  establishment_rate,
  outfile,
  return_rast = FALSE,
  overwrite = FALSE
)
```

### Arguments

fertiliser\_weight

A RasterLayer or file path to a raster file containing weights to distribute arrivals by.

leakage_rate	Numeric vector of 2 values, giving the lower and upper bounds of a 95 a random year).
establishment_rate	Numeric vector of 2 values, giving the lower and upper bounds of a 95 establishment, for leakage events).
outfile	Character. Output raster file path. If probability has length > 1, the file type must support multiband rasters (e.g. GeoTiff). If not provided, raster object will be returned to R.
return_rast	Logical. Should the raster object be returned to R? Ignored if outfile is not provided.
overwrite	Logical. If TRUE and outfile is not missing, it will be overwritten if the file specified by outfile already exists.

### Value

If outfile is specified, the resulting raster (multiband if probability has length > 1) is saved as a tiff at that path. If return\_rast is TRUE or outfile is not specified the resulting raster object is returned, otherwise NULL is returned invisibly.

### See Also

Other functions estimating arrivals: [arrivals\\_by\\_containers\(\)](#), [arrivals\\_by\\_food\(\)](#), [arrivals\\_by\\_goods\(\)](#), [arrivals\\_by\\_machinery\(\)](#), [arrivals\\_by\\_mail\(\)](#), [arrivals\\_by\\_nurserystock\(\)](#), [arrivals\\_by\\_residents\(\)](#), [arrivals\\_by\\_torres\(\)](#), [arrivals\\_by\\_tourists\(\)](#), [arrivals\\_by\\_vessels\(\)](#), [arrivals\\_by\\_wind\(\)](#)

---

arrivals_by_food	<i>Estimates pest arrivals through imported fresh food</i>
------------------	--

---

### Description

Estimates pest arrivals through imported fresh food as a function of population density.

### Usage

```
arrivals_by_food(
  pop_density,
  leakage_rate,
  establishment_rate,
  outfile,
  return_rast = FALSE,
  overwrite = FALSE
)
```

### Arguments

pop_density	A RasterLayer or file path to a raster file containing population density.
leakage_rate	Numeric vector of 2 values, giving the lower and upper bounds of a 95 a random year).
establishment_rate	Numeric vector of 2 values, giving the lower and upper bounds of a 95 establishment, for leakage events).

outfile	Character. Output raster file path. If probability has length > 1, the file type must support multiband rasters (e.g. GeoTiff). If not provided, raster object will be returned to R.
return_rast	Logical. Should the raster object be returned to R? Ignored if outfile is not provided.
overwrite	Logical. If TRUE and outfile is not missing, it will be overwritten if the file specified by outfile already exists.

### Value

If outfile is specified, the resulting raster (multiband if probability has length > 1) is saved as a tiff at that path. If return\_rast is TRUE or outfile is not specified the resulting raster object is returned, otherwise NULL is returned invisibly.

### See Also

Other functions estimating arrivals: [arrivals\\_by\\_containers\(\)](#), [arrivals\\_by\\_fertiliser\(\)](#), [arrivals\\_by\\_goods\(\)](#), [arrivals\\_by\\_machinery\(\)](#), [arrivals\\_by\\_mail\(\)](#), [arrivals\\_by\\_nurserystock\(\)](#), [arrivals\\_by\\_residents\(\)](#), [arrivals\\_by\\_torres\(\)](#), [arrivals\\_by\\_tourists\(\)](#), [arrivals\\_by\\_vessels\(\)](#), [arrivals\\_by\\_wind\(\)](#)

---

arrivals_by_goods	<i>Estimates pest arrivals through imported goods</i>
-------------------	---

---

### Description

Estimates pest arrivals through imported goods as a function of population density.

### Usage

```
arrivals_by_goods(
  pop_density,
  leakage_rate,
  establishment_rate,
  outfile,
  return_rast = FALSE,
  overwrite = FALSE
)
```

### Arguments

pop_density	A RasterLayer or file path to a raster file containing population density.
leakage_rate	Numeric vector of 2 values, giving the lower and upper bounds of a 95 a random year).
establishment_rate	Numeric vector of 2 values, giving the lower and upper bounds of a 95 establishment, for leakage events).
outfile	Character. Output raster file path. If probability has length > 1, the file type must support multiband rasters (e.g. GeoTiff). If not provided, raster object will be returned to R.

return_rast	Logical. Should the raster object be returned to R? Ignored if outfile is not provided.
overwrite	Logical. If TRUE and outfile is not missing, it will be overwritten if the file specified by outfile already exists.

### Value

If outfile is specified, the resulting raster (multiband if probability has length > 1) is saved as a tiff at that path. If return\_rast is TRUE or outfile is not specified the resulting raster object is returned, otherwise NULL is returned invisibly.

### See Also

Other functions estimating arrivals: [arrivals\\_by\\_containers\(\)](#), [arrivals\\_by\\_fertiliser\(\)](#), [arrivals\\_by\\_food\(\)](#), [arrivals\\_by\\_machinery\(\)](#), [arrivals\\_by\\_mail\(\)](#), [arrivals\\_by\\_nurserystock\(\)](#), [arrivals\\_by\\_residents\(\)](#), [arrivals\\_by\\_torres\(\)](#), [arrivals\\_by\\_tourists\(\)](#), [arrivals\\_by\\_vessels\(\)](#), [arrivals\\_by\\_wind\(\)](#)

---

arrivals\_by\_machinery *Estimates pest arrivals through imported machinery*

---

### Description

Estimates pest arrivals through imported machinery as a function of population density.

### Usage

```
arrivals_by_machinery(
  pop_density,
  leakage_rate,
  establishment_rate,
  outfile,
  return_rast = FALSE,
  overwrite = FALSE
)
```

### Arguments

pop_density	A RasterLayer or file path to a raster file containing population density.
leakage_rate	Numeric vector of 2 values, giving the lower and upper bounds of a 95 a random year).
establishment_rate	Numeric vector of 2 values, giving the lower and upper bounds of a 95 establishment, for leakage events).
outfile	Character. Output raster file path. If probability has length > 1, the file type must support multiband rasters (e.g. GeoTiff). If not provided, raster object will be returned to R.
return_rast	Logical. Should the raster object be returned to R? Ignored if outfile is not provided.
overwrite	Logical. If TRUE and outfile is not missing, it will be overwritten if the file specified by outfile already exists.

## Details

This pathway is used under the assumption that the vast majority of imported machinery comprises new and used motor vehicles.

## Value

If outfile is specified, the resulting raster (multiband if probability has length > 1) is saved as a tiff at that path. If return\_rast is TRUE or outfile is not specified the resulting raster object is returned, otherwise NULL is returned invisibly.

## See Also

Other functions estimating arrivals: [arrivals\\_by\\_containers\(\)](#), [arrivals\\_by\\_fertiliser\(\)](#), [arrivals\\_by\\_food\(\)](#), [arrivals\\_by\\_goods\(\)](#), [arrivals\\_by\\_mail\(\)](#), [arrivals\\_by\\_nurserystock\(\)](#), [arrivals\\_by\\_residents\(\)](#), [arrivals\\_by\\_torres\(\)](#), [arrivals\\_by\\_tourists\(\)](#), [arrivals\\_by\\_vessels\(\)](#), [arrivals\\_by\\_wind\(\)](#)

---

arrivals_by_mail	<i>Estimates pest arrivals through mail</i>
------------------	---

---

## Description

Estimates arrival rate attributable to mail as a function of population density.

## Usage

```
arrivals_by_mail(
  pop_density,
  leakage_rate,
  establishment_rate,
  outfile,
  return_rast = FALSE,
  overwrite = FALSE
)
```

## Arguments

pop_density	A RasterLayer or file path to a raster file containing population density.
leakage_rate	Numeric vector of 2 values, giving the lower and upper bounds of a 95 a random year).
establishment_rate	Numeric vector of 2 values, giving the lower and upper bounds of a 95 establishment, for leakage events).
outfile	Character. Output raster file path. If probability has length > 1, the file type must support multiband rasters (e.g. GeoTiff). If not provided, raster object will be returned to R.
return_rast	Logical. Should the raster object be returned to R? Ignored if outfile is not provided.
overwrite	Logical. If TRUE and outfile is not missing, it will be overwritten if the file specified by outfile already exists.



**Value**

If outfile is specified, the resulting raster (multiband if probability has length > 1) is saved as a tiff at that path. If return\_rast is TRUE or outfile is not specified the resulting raster object is returned, otherwise NULL is returned invisibly.

**See Also**

Other functions estimating arrivals: [arrivals\\_by\\_containers\(\)](#), [arrivals\\_by\\_fertiliser\(\)](#), [arrivals\\_by\\_food\(\)](#), [arrivals\\_by\\_goods\(\)](#), [arrivals\\_by\\_machinery\(\)](#), [arrivals\\_by\\_nurserystock\(\)](#), [arrivals\\_by\\_residents\(\)](#), [arrivals\\_by\\_torres\(\)](#), [arrivals\\_by\\_tourists\(\)](#), [arrivals\\_by\\_vessels\(\)](#), [arrivals\\_by\\_wind\(\)](#)

---

arrivals\_by\_nurserystock

*Estimates pest arrivals through imported nursery stock*


---

**Description**

Estimates pest arrivals through imported nursery stock as a function of population density.

**Usage**

```
arrivals_by_nurserystock(
  pop_density,
  leakage_rate,
  establishment_rate,
  outfile,
  return_rast = FALSE,
  overwrite = FALSE
)
```

**Arguments**

pop_density	A RasterLayer or file path to a raster file containing population density.
leakage_rate	Numeric vector of 2 values, giving the lower and upper bounds of a 95 a random year).
establishment_rate	Numeric vector of 2 values, giving the lower and upper bounds of a 95 establishment, for leakage events).
outfile	Character. Output raster file path. If probability has length > 1, the file type must support multiband rasters (e.g. GeoTiff). If not provided, raster object will be returned to R.
return_rast	Logical. Should the raster object be returned to R? Ignored if outfile is not provided.
overwrite	Logical. If TRUE and outfile is not missing, it will be overwritten if the file specified by outfile already exists.

**Value**

If outfile is specified, the resulting raster (multiband if probability has length > 1) is saved as a tiff at that path. If return\_rast is TRUE or outfile is not specified the resulting raster object is returned, otherwise NULL is returned invisibly.

**See Also**

Other functions estimating arrivals: [arrivals\\_by\\_containers\(\)](#), [arrivals\\_by\\_fertiliser\(\)](#), [arrivals\\_by\\_food\(\)](#), [arrivals\\_by\\_goods\(\)](#), [arrivals\\_by\\_machinery\(\)](#), [arrivals\\_by\\_mail\(\)](#), [arrivals\\_by\\_residents\(\)](#), [arrivals\\_by\\_torres\(\)](#), [arrivals\\_by\\_tourists\(\)](#), [arrivals\\_by\\_vessels\(\)](#), [arrivals\\_by\\_wind\(\)](#)

---

arrivals\_by\_residents *Estimates pest arrivals by returning residents*

---

**Description**

Estimates arrival rate attributable to returning residents as a function of population density.

**Usage**

```
arrivals_by_residents(
  pop_density,
  leakage_rate,
  establishment_rate,
  outfile,
  return_rast = FALSE,
  overwrite = FALSE
)
```

**Arguments**

pop_density	A RasterLayer or file path to a raster file containing population density.
leakage_rate	Numeric vector of 2 values, giving the lower and upper bounds of a 95 a random year).
establishment_rate	Numeric vector of 2 values, giving the lower and upper bounds of a 95 establishment, for leakage events).
outfile	Character. Output raster file path. If probability has length > 1, the file type must support multiband rasters (e.g. GeoTiff). If not provided, raster object will be returned to R.
return_rast	Logical. Should the raster object be returned to R? Ignored if outfile is not provided.
overwrite	Logical. If TRUE and outfile is not missing, it will be overwritten if the file specified by outfile already exists.

**Value**

If outfile is specified, the resulting raster (multiband if probability has length > 1) is saved as a tiff at that path. If return\_rast is TRUE or outfile is not specified the resulting raster object is returned, otherwise NULL is returned invisibly.

**See Also**

Other functions estimating arrivals: [arrivals\\_by\\_containers\(\)](#), [arrivals\\_by\\_fertiliser\(\)](#), [arrivals\\_by\\_food\(\)](#), [arrivals\\_by\\_goods\(\)](#), [arrivals\\_by\\_machinery\(\)](#), [arrivals\\_by\\_mail\(\)](#), [arrivals\\_by\\_nurserystock\(\)](#), [arrivals\\_by\\_torres\(\)](#), [arrivals\\_by\\_tourists\(\)](#), [arrivals\\_by\\_vessels\(\)](#), [arrivals\\_by\\_wind\(\)](#)

---

arrivals_by_torres	<i>Estimates pest arrivals due to Torres Strait air passengers</i>
--------------------	--

---

**Description**

Estimates pest arrivals due to Torres Strait air passengers coming into Cairns.

**Usage**

```
arrivals_by_torres(
  pop_density,
  airport_weight,
  leakage_rate,
  establishment_rate,
  outfile,
  return_rast = FALSE,
  overwrite = FALSE
)
```

**Arguments**

pop_density	A RasterLayer or file path to a raster file containing population density.
airport_weight	A RasterLayer or file path to a raster file containing distance from Cairns airport weights.
leakage_rate	Numeric vector of 2 values, giving the lower and upper bounds of a 95 a random year).
establishment_rate	Numeric vector of 2 values, giving the lower and upper bounds of a 95 establishment, for leakage events).
outfile	Character. Output raster file path. If probability has length > 1, the file type must support multiband rasters (e.g. GeoTiff). If not provided, raster object will be returned to R.
return_rast	Logical. Should the raster object be returned to R? Ignored if outfile is not provided.
overwrite	Logical. If TRUE and outfile is not missing, it will be overwritten if the file specified by outfile already exists.

**Value**

If outfile is specified, the resulting raster (multiband if probability has length > 1) is saved as a tiff at that path. If return\_rast is TRUE or outfile is not specified the resulting raster object is returned, otherwise NULL is returned invisibly.

**See Also**

Other functions estimating arrivals: [arrivals\\_by\\_containers\(\)](#), [arrivals\\_by\\_fertiliser\(\)](#), [arrivals\\_by\\_food\(\)](#), [arrivals\\_by\\_goods\(\)](#), [arrivals\\_by\\_machinery\(\)](#), [arrivals\\_by\\_mail\(\)](#), [arrivals\\_by\\_nurserystock\(\)](#), [arrivals\\_by\\_residents\(\)](#), [arrivals\\_by\\_tourists\(\)](#), [arrivals\\_by\\_vessels\(\)](#), [arrivals\\_by\\_wind\(\)](#)

---

arrivals\_by\_tourists    *Estimates pest arrival due to tourists*

---

**Description**

Estimates pest arrival due to tourists as a function of distance from airport and tourist accommodation.

**Usage**

```
arrivals_by_tourists(
  tourist_beds,
  airport_weights,
  leakage_rate,
  establishment_rate,
  outfile,
  return_rast = FALSE,
  overwrite = FALSE
)
```

**Arguments**

tourist_beds	A RasterLayer or file path to a raster file containing tourist bed numbers.
airport_weights	A RasterLayer or file path to a raster file containing airport distance weights.
leakage_rate	Numeric vector of 2 values, giving the lower and upper bounds of a 95 a random year).
establishment_rate	Numeric vector of 2 values, giving the lower and upper bounds of a 95 establishment, for leakage events).
outfile	Character. Output raster file path. If probability has length > 1, the file type must support multiband rasters (e.g. GeoTiff). If not provided, raster object will be returned to R.
return_rast	Logical. Should the raster object be returned to R? Ignored if outfile is not provided.
overwrite	Logical. If TRUE and outfile is not missing, it will be overwritten if the file specified by outfile already exists.

**Value**

If outfile is specified, the resulting raster (multiband if probability has length > 1) is saved as a tiff at that path. If return\_rast is TRUE or outfile is not specified the resulting raster object is returned, otherwise NULL is returned invisibly.

**See Also**

Other functions estimating arrivals: [arrivals\\_by\\_containers\(\)](#), [arrivals\\_by\\_fertiliser\(\)](#), [arrivals\\_by\\_food\(\)](#), [arrivals\\_by\\_goods\(\)](#), [arrivals\\_by\\_machinery\(\)](#), [arrivals\\_by\\_mail\(\)](#), [arrivals\\_by\\_nurserystock\(\)](#), [arrivals\\_by\\_residents\(\)](#), [arrivals\\_by\\_torres\(\)](#), [arrivals\\_by\\_vessels\(\)](#), [arrivals\\_by\\_wind\(\)](#)

---

arrivals_by_vessels	<i>Estimates pest arrivals by vessels</i>
---------------------	---

---

**Description**

Estimates pest arrivals by vessels.

**Usage**

```
arrivals_by_vessels(
  port_weight,
  leakage_rate,
  establishment_rate,
  outfile,
  return_rast = FALSE,
  overwrite = FALSE
)
```

**Arguments**

port_weight	A RasterLayer or file path to a raster file containing weights to be used to distribute arrivals.
leakage_rate	Numeric vector of 2 values, giving the lower and upper bounds of a 95 a random year).
establishment_rate	Numeric vector of 2 values, giving the lower and upper bounds of a 95 establishment, for leakage events).
outfile	Character. Output raster file path. If probability has length > 1, the file type must support multiband rasters (e.g. GeoTiff). If not provided, raster object will be returned to R.
return_rast	Logical. Should the raster object be returned to R? Ignored if outfile is not provided.
overwrite	Logical. If TRUE and outfile is not missing, it will be overwritten if the file specified by outfile already exists.

**Value**

If outfile is specified, the resulting raster (multiband if probability has length > 1) is saved as a tiff at that path. If return\_rast is TRUE or outfile is not specified the resulting raster object is returned, otherwise NULL is returned invisibly.

**See Also**

Other functions estimating arrivals: [arrivals\\_by\\_containers\(\)](#), [arrivals\\_by\\_fertiliser\(\)](#), [arrivals\\_by\\_food\(\)](#), [arrivals\\_by\\_goods\(\)](#), [arrivals\\_by\\_machinery\(\)](#), [arrivals\\_by\\_mail\(\)](#), [arrivals\\_by\\_nurserystock\(\)](#), [arrivals\\_by\\_residents\(\)](#), [arrivals\\_by\\_torres\(\)](#), [arrivals\\_by\\_tourists\(\)](#), [arrivals\\_by\\_wind\(\)](#)

---

arrivals_by_wind	<i>Estimates pest arrivals via wind</i>
------------------	---

---

**Description**

Estimates pest arrivals via coastal onshore winds.

**Usage**

```
arrivals_by_wind(
  wind_speed,
  leakage_rate,
  establishment_rate,
  outfile,
  return_rast = FALSE,
  overwrite = FALSE
)
```

**Arguments**

wind_speed	A RasterLayer or file path to a raster file containing wind speed.
leakage_rate	Numeric vector of 2 values, giving the lower and upper bounds of a 95 a random year).
establishment_rate	Numeric vector of 2 values, giving the lower and upper bounds of a 95 establishment, for leakage events).
outfile	Character. Output raster file path. If not provided, raster object will be returned to R.
return_rast	Logical. Should the raster object be returned to R? Ignored if outfile is not provided.
overwrite	Logical. If TRUE and outfile is not missing, it will be overwritten if the file specified by outfile already exists.

**Value**

If outfile is specified, the resulting raster (multiband if probability has length > 1) is saved as a tiff at that path. If return\_rast is TRUE or outfile is not specified the resulting raster object is returned, otherwise NULL is returned invisibly.

**See Also**

Other functions estimating arrivals: [arrivals\\_by\\_containers\(\)](#), [arrivals\\_by\\_fertiliser\(\)](#), [arrivals\\_by\\_food\(\)](#), [arrivals\\_by\\_goods\(\)](#), [arrivals\\_by\\_machinery\(\)](#), [arrivals\\_by\\_mail\(\)](#), [arrivals\\_by\\_nurserystock\(\)](#), [arrivals\\_by\\_residents\(\)](#), [arrivals\\_by\\_torres\(\)](#), [arrivals\\_by\\_tourists\(\)](#), [arrivals\\_by\\_vessels\(\)](#)



---

binarize\_and\_aggregate

*Binarize a categorical raster and optionally aggregate*


---

## Description

Binarize a categorical raster and optionally aggregate and/or resample.

## Usage

```
binarize_and_aggregate(
  infile,
  rle,
  outfile,
  extent,
  res,
  categories,
  overwrite = FALSE,
  return_rast = FALSE,
  quiet = FALSE
)
```

## Arguments

infile	Character. File path to a categorical raster. If rle is provided, infile must not be provided.
rle	A raster_rle object generated by <a href="#">rle_compress</a> or a file path to such an object saved as .rds.
outfile	Character. The target file path for the binarized raster.
extent	Output extent as an Extent object or an object from which an Extent object can be extracted. If not provided, extent will be taken from input. If extent differs from that of infile, new cells will be assigned value 1 if any original cells belonging to categories have their centroids within the new cell.
res	output resolution as a numeric vector of 1 or 2 elements, or a Raster* object from which resolution can be extracted. If not provided, resolution will be taken from input. If res differs from that of infile, new cells will be assigned value 1 if any original cells belonging to categories have their centroids within the new cell.
categories	Integer or numeric vector of class values to be labelled as 1 in the target raster.
overwrite	Logical. Should outfile be overwritten if it exists?
return_rast	Logical. Should the target raster be returned to R as a Raster layer (TRUE) or not returned (FALSE)?
quiet	Logical. Should progress messages be suppressed?

## Value

A binarized raster layer is written to outfile, and if return\_rast is TRUE, the raster is additionally returned to R as a Raster layer.

---

calculate_median	<i>Calculate median of raster objects</i>
------------------	---

---

**Description**

Calculate the cellwise median of two or more Raster\* objects, by layer.

**Usage**

```
calculate_median(files, outfile)
```

**Arguments**

files	A vector of file paths to (optionally multiband) raster files.
outfile	Character. Output raster file path. If not provided, object will be returned to R. Directory will be created recursively if it does not exist.

**Value**

If outfile is specified, the resulting raster is saved as a geotiff to that path and NULL is returned. If outfile is not specified the resulting raster is returned.

---

calc_EE	<i>Calculate establishment probability</i>
---------	--

---

**Description**

Calculate the probability of one or more establishment events occurring as a function of estimated leakage numbers and the probability a leakage event could result in establishment.

**Usage**

```
calc_EE(n_events, p_establish, nsims = 1e+05)
```

**Arguments**

n_events	Numeric vector containing lower and upper bounds (95 the number of leakage events).
p_establish	Numeric vector containing the lower and upper bounds (95 CI) for the probability a leakage event could result in an establishment.
nsims	Integer. Number of samples to be taken from event and establishment distributions. Default = 100000.

**Value**

A data.frame containing the possible number of incursions that may occur, and their corresponding probabilities.

**Author(s)**

James Camac (<james.camac@gmail.com>)

---

calc_pathway_pr	<i>Calculate the pathway-specific probability of pest arrival</i>
-----------------	---

---

**Description**

Calculate the pathway-specific probability of pest arrival for each raster cell.

**Usage**

```
calc_pathway_pr(EE, rast, outfile, return_rast = TRUE)
```

**Arguments**

EE	data.frame object obtained from <a href="#">calc_EE</a> .
rast	Raster object or character path to raster file containing dispersal weights.
outfile	Character. Output raster file path. If not provided, the RasterLayer will be returned to R.
return_rast	Logical. Should the RasterLayer be returned to R? Ignored if outfile is not provided.

**Value**

If outfile is specified, the resulting RasterLayer is saved to outfile. If return\_rast is TRUE or outfile is not specified, the resulting RasterLayer is returned, otherwise NULL is returned invisibly.

---

calc_proportion	<i>Calculate raster cell values as proportions</i>
-----------------	--

---

**Description**

Calculates raster cell values as a proportion of the sum of all cells' values.

**Usage**

```
calc_proportion(rast)
```

**Arguments**

rast	Raster object.
------	----------------

**Value**

A RasterLayer.

---

captured_by_ncells	<i>Calculate establishment likelihood captured in top n cells</i>
--------------------	---

---

**Description**

Calculate the proportion of establishment likelihood captured in top n cells.

**Usage**

```
captured_by_ncells(infiles, names, n_cells, all = TRUE)
```

**Arguments**

infiles	Character vector. File path(s) to one or more raster files.
names	Character vector. Names corresponding to infiles.
n_cells	Integer. The number of cells to consider.
all	Logical. If TRUE, return the proportion captured from 1 to <i>n</i> cells. If FALSE, return the cumulative proportion.

**Value**

Proportion of establishment likelihood captured, or a vector of cumulative proportions.

---

combine_arrivals	<i>Sum arrivals across entry pathways</i>
------------------	---

---

**Description**

Sums estimated arrivals rates across all entry pathways.

**Usage**

```
combine_arrivals(
  x,
  outfile,
  summarise_uncertainty = FALSE,
  return_rast = FALSE
)
```

**Arguments**

x	Character vector giving file path(s) to rasters to be included in the summation. All rasters must have the same extent and resolution.
outfile	Character. Output raster file path. If not provided, the RasterLayer will be returned to R.
summarise_uncertainty	Logical. If TRUE, and if probability has length > 1, the arrival frequency surfaces for the values of probability will be summarised to their median, min, and max cell-wise values (in that order).
return_rast	Logical. Should the RasterLayer be returned to R? Ignored if outfile is not provided.

**Value**

If outfile is specified, the resulting RasterLayer is saved to outfile. If return\_rast is TRUE or outfile is not specified, the resulting RasterLayer is returned, otherwise NULL is returned invisibly. If summarise\_uncertainty is TRUE and rasters passed to 'x' are multiband, then the resulting stack will have three layers equal to the cell-wise median, minimum, and maximum of the layers, respectively.

---

combine_pathways	<i>Calculate the probability of arrival across pathways</i>
------------------	---

---

**Description**

Calculate the probability of arrival across pathways.

**Usage**

```
combine_pathways(x, outfile, return_rast = FALSE)
```

**Arguments**

x	Either a RasterStack or a character vector giving file path(s) to rasters. Each raster should describe a pathway to be included in the summation, and cell values give the probability of arrival for the corresponding pathway. All rasters must have the same extent and resolution.
outfile	Character. Output raster file path. If not provided, the RasterLayer will be returned to R.
return_rast	Logical. Should the RasterLayer be returned to R? Ignored if outfile is not provided.

**Value**

If outfile is specified, the resulting RasterLayer is saved to outfile. If return\_rast is TRUE or outfile is not specified, the resulting RasterLayer is returned, otherwise NULL is returned invisibly.

---

container_weights	<i>Creates a weight by postcode for each major port</i>
-------------------	---

---

**Description**

Creates a weight by postcode for each major port.

**Usage**

```

container_weights(
  path,
  sheet_nums,
  range = "A7:M2217",
  postcode_shp,
  na = c("", "-", "np"),
  outfile,
  return_sf = FALSE
)

```

**Arguments**

path	Character. File path to xls file containing containers by postcode for each port.
sheet_nums	Integer. Vector of integers signifying the sheet numbers to read in.
range	A cell range to read from, as described in cell-specification. Includes typical <i>Excel</i> ranges such as "B3:D87", possibly including the sheet name like "Budget!B2:G14", and more. Interpreted strictly, even if the range forces the inclusion of leading or trailing empty rows or columns.
postcode_shp	Character. File path to postcode shape file.
na	Character vector of strings to interpret as missing values. By default, readxl treats blank cells as missing data.
outfile	Character. Name of shapefile where output will be saved. If not provided, sf object will be returned to R.
return_sf	Logical. Should the sf object be returned to R? Ignored if outfile is not provided.

**Details**

For the purposes of this analysis missing data (i.e. NAs) will be treated as zeroes.

**Value**

An sf object or shapefile export.

---

download\_worldclim2      *Download climate layers from WorldClim 2.0*

---

**Description**

Download climate layers from WorldClim 2.0 (current climate only).

**Usage**

```
download_worldclim2(outfile, variable, resolution)
```



**Arguments**

outfile	Character. Target file for downloaded .zip archive.
variable	Character. Can be one of "bio" (19 standard bioclim variables), "tmin", "tmax", "tavg", "srad", "wind" or "vapr".
resolution	Character. Can be one of "10m", "5m", "2.5m", or "30s".

**Value**

A zipfile is downloaded and NULL is returned invisibly.

**See Also**

[extract\\_worldclim2](#)

**Examples**

```
## Not run:
download_worldclim2('bioclim_10m', 'bio', '10m')

## End(Not run)
```

---

establishment\_likelihood

*Estimate establishment likelihood*

---

**Description**

Estimate the likelihood of pest establishment based on total arrivals and environmental suitability.

**Usage**

```
establishment_likelihood(
  total_arrivals,
  suitability,
  outfile,
  return_rast = FALSE
)
```

**Arguments**

total_arrivals	A Raster* object or path to raster file containing total arrival estimates.
suitability	A RasterLayer or path to raster file containing suitability scores.
outfile	Character. Output raster file path. Must support multiband raster if total_arrivals has multiple layers. If not provided, the Raster* object will be returned to R.
return_rast	Logical. Should the RasterLayer be returned to R? Ignored if outfile is not provided.

**Value**

If outfile is specified, the resulting Raster\* object is saved as to that path. If return\_rast is TRUE or outfile is not specified the resulting RasterLayer is returned, otherwise NULL is returned invisibly.

---

 excel\_to\_plan

*Create a drake plan using tabular species data*


---

### Description

Import parameters from an Excel xlsx file and create a drake plan to estimate establishment likelihoods.

### Usage

```
excel_to_plan(file)
```

### Arguments

**file** Character. File path to an xlsx file containing required parameters for creating a species plan. See Details and [species\\_plan](#) for further information.

### Details

To simplify reproducibility, edmaps provides an *Excel* interface for specifying species parameters relevant to estimating establishment likelihood. An example spreadsheet is bundled with the package, available at the path given by `system.file('extdata/parameters.xlsx', package='edmaps')`. The spreadsheet has two sheets, the first specifying "global" parameters that will apply to all species (e.g. file paths to rasters that will), be used regardless of species identity and the second specifying parameters that can vary by species. In the second sheet, each row corresponds to a separate species. Tooltips and data validation guide the user with respect to expected/allowable data.

### Value

A drake plan with a combined workflow for all species.

---

 exp\_function

*Exponentiate raster values*


---

### Description

Apply exponential function to raster values.

### Usage

```
exp_function(rast, beta)
```

### Arguments

**rast** Raster object.  
**beta** Numeric. The beta coefficient of a standard exponential function.

### Value

A Raster object.

---

```
extract_highest_ncells
```

*Returns a raster with values for the  $n$  cells with highest establishment likelihood*

---

### Description

Returns a raster with values only for the  $n$  cells with highest establishment likelihood.

### Usage

```
extract_highest_ncells(infile, n_cells, outfile, return_rast = FALSE)
```

### Arguments

<code>infile</code>	File path to a raster containing estimated likelihoods of establishment or arrival.
<code>n_cells</code>	Integer. The number of cells to return.
<code>outfile</code>	Character. Output raster file path. If not provided, the RasterLayer will be returned to R.
<code>return_rast</code>	Logical. Should the RasterLayer be returned to R? Ignored if outfile is not provided.

### Details

This function returns a raster containing values for only the top  $n$  cells.

### Value

If outfile is specified, the resulting RasterLayer is saved as to that path. If return\_rast is TRUE or outfile is not specified, the resulting RasterLayer is returned, otherwise NULL is returned invisibly.

### Warning

This function uses the quantile approach for determining the top  $n$  cells. As such, if risk is highly aggregated this function may run into issues when  $n\_cells$  is high.

---

```
extract_worldclim2
```

*Extract WorldClim 2.0 data*

---

### Description

Extract WorldClim 2.0 data

### Usage

```
extract_worldclim2(path_2_zip, outdir)
```

**Arguments**

path_2_zip	Character. Path to .zip archive downloaded by <a href="#">download_worldclim2</a> .
outdir	Character. File path to which contained files should be extracted. Will be created (recursively) if necessary.

**Value**

Raster data are extracted to outdir and NULL is returned invisibly.

**See Also**

[download\\_worldclim2](#)

**Examples**

```
## Not run:
download_worldclim2('bioclim_10m.zip', 'bio', '10m')
extract_worldclim2('bioclim_10m.zip', outdir='bioclim')

## End(Not run)
```

---

fertiliser_by_nrm	<i>Create fertiliser by nrm sf object</i>
-------------------	---

---

**Description**

Create fertiliser by nrm sf object.

**Usage**

```
fertiliser_by_nrm(abs_data, nrm_shapefile, outfile, return_sf = FALSE)
```

**Arguments**

abs_data	Character. File path to ABS .csv file.
nrm_shapefile	Character. File path to NRM shapefile.
outfile	Character. Name of shapefile (or other vector data format supported by OGR) where output will be saved. If not provided, sf object will be returned to R.
return_sf	Logical. Should the sf object be returned to R? Ignored if outfile is not provided.

**Value**

An sf object or vector data export.

---

fertiliser_weight	<i>Create a fertiliser weight raster</i>
-------------------	--

---

### Description

Create a fertiliser weight raster as a function of estimated nrm fertiliser tonnes and landuses.

### Usage

```
fertiliser_weight(fert_nrm, fert_landuses, outfile, return_rast = FALSE)
```

### Arguments

fert_nrm	A file path to vector data or an sf object.
fert_landuses	A file path to raster file or a RasterLayer object.
outfile	Character. Output raster file path. If not provided, RasterLayer will be returned to R.
return_rast	Logical. Should the RasterLayer be returned to R? Ignored if outfile is not provided.

### Value

If outfile is specified, the resulting RasterLayer is saved as to that path. If return\_rast is TRUE or outfile is not specified, the resulting RasterLayer is returned, otherwise NULL is returned invisibly.

---

fill_na	<i>Modify NAs within neighbourhood of data cells</i>
---------	--

---

### Description

Apply a function to NA cells within a moving window.

### Usage

```
fill_na(x, fun, w, outfile, return_rast = FALSE, overwrite = FALSE)
```

### Arguments

x	A Raster* object.
fun	The function (name or symbol) to apply to the moving windows. First argument should represent the vector of cells contained in the focal window. Only non-NA cell values will be passed to the function.
w	A weights matrix (see <a href="#">focalWeight</a> and <a href="#">focal</a> defining the focal window to which fun will be applied. Note that all contributing cells will be given equal weight (i.e. varying weights are not respected - non-zero & non-NA weights will be replaced by 1).
outfile	Character. File path to an output raster file. If missing, a temporary file will be used.
return_rast	Logical. Should the RasterLayer be returned to R?
overwrite	Logical. Should outfile be replaced if it already exists?

**Details**

NA values within focal blocks will be ignored. Raster edges will be NA-padded to allow focal computations at the edge (see [focal](#)).

**Value**

Returns the resulting RasterLayer if return\_rast is TRUE. Returns the output file path otherwise.

---

gdal_reproject	<i>Reproject and resample a raster</i>
----------------	--

---

**Description**

This function can be used to change resolution, projection and extent of a raster.

**Usage**

```
gdal_reproject(
  infile,
  outfile,
  src_proj,
  tgt_proj,
  res,
  resampling_method = "near",
  tgt_extent,
  buffer,
  src_nodata,
  datatype = "Float32",
  return_rast = FALSE,
  overwrite = TRUE
)
```

**Arguments**

infile	Character. File path to a raster file.
outfile	Character. Output raster file path.
src_proj	Character. The source file coordinate system. Only needs to be set if you wish to reproject object and it is not specified in the infile. The coordinate systems that can be passed includes EPSG PCS and GCSes (i.e. EPSG:4296), PROJ.4 declarations (as above), or the name of a .prj file containing well known text. Starting with GDAL 2.2, if the SRS has an explicit vertical datum that points to a PROJ.4 geoidgrids, and the input dataset is a single band dataset, a vertical correction will be applied to the values of the dataset.
tgt_proj	Character. The target coordinate system. See src_proj for details. If not specified, outfile will inherit the coordinate system of infile.
res	Numeric vector containing x and y resolution. e.g. c(xres,yres). Must be specified in the units of tgt_proj.



resampling_method	Character. One of "near", "bilinear", "cubic", "cubicspline", "lanczos", "average", "mode", "max", "min", "med", "q1", "q3". See Details. Default is "near".
tgt_extent	Numeric vector containing corner coordinates specified as c(xmin, ymin, xmax, ymax). Must be specified in the units of tgt_proj.
buffer	Numeric. Add buffer (specified in units of tgt_proj). Useful for adding buffers around coasts, and for filling in small gaps within raster specified by infile.
src_nodata	Integer. The nodata value for an input file.
datatype	A character string indicating the output data type. See the <a href="#">gdalwarp docs</a> for more information.
return_rast	Logical. Return RasterLayer to R?
overwrite	Logical. Should outfile be overwritten if it already exists?

## Details

The resampling methods available are as follows:

- near: nearest neighbour resampling (default, fastest algorithm, worst interpolation quality).
- bilinear: bilinear resampling.
- cubic: cubic resampling.
- cubicspline: cubic spline resampling.
- lanczos: Lanczos windowed sinc resampling.
- average: average resampling, computes the average of all non-NODATA contributing pixels. (GDAL >= 1.10.0).
- mode: mode resampling, selects the value which appears most often of all the sampled points. (GDAL >= 1.10.0).
- max: maximum resampling, selects the maximum value from all non-NODATA contributing pixels. (GDAL >= 2.0.0).
- min: minimum resampling, selects the minimum value from all non-NODATA contributing pixels. (GDAL >= 2.0.0).
- med: median resampling, selects the median value of all non-NODATA contributing pixels. (GDAL >= 2.0.0).
- q1: first quartile resampling, selects the first quartile value of all non-NODATA contributing pixels. (GDAL >= 2.0.0).
- q3: third quartile resampling, selects the third quartile value of all non-NODATA contributing pixels. (GDAL >= 2.0.0).

## Value

`_rast` A raster file is produced on disk. Additionally, if `return_rast` is TRUE a RasterLayer object is returned to R.

---

get_airport_dist	<i>Estimate distance to nearest Australian international airport</i>
------------------	--

---

## Description

Generates a raster proximity map indicating the distance from the center of each pixel to the centre of the nearest pixel containing a major airport. Airport coordinates are in the source raster.

## Usage

```
get_airport_dist(
  vector_data,
  outfile,
  template_raster,
  extent,
  res,
  crs,
  airport_codes,
  return_rast = FALSE,
  overwrite = FALSE
)
```

## Arguments

vector_data	Character. Path to airport data (available <a href="#">here</a> ).
outfile	Character. Output raster file path. Directory will be created recursively if it does not exist.
template_raster	Optional. Raster* object or a file path to template raster. If this is provided, extent, res, and crs will be taken from this raster unless they are also passed to this function. If template_raster is not provided, then extent and res must be provided. The template raster will also be used to mask the output raster so that NA cell values will be propagated.
extent	Either a character path to a raster file, an <a href="#">Extent</a> object (or an object from which such an extent can be extracted), or a numeric vector with four elements giving xmin, xmax, ymin, ymax.
res	Numeric or integer vector giving the horizontal and vertical spatial resolution, in units of crs. If a single value is given, it will be used for both horizontal and vertical resolution.
crs	Target coordinate reference system as a PROJ string (character) or an object of class CRS.
airport_codes	Numeric. Airport codes to be used as targets in order to estimate cell proximity.
return_rast	Logical. Return RasterLayer to R?
overwrite	Logical. Should outfile be removed if it already exists?

**Value**

A proximity raster is written to outfile. If return\_rast is TRUE, the raster object is also returned to R, otherwise NULL is returned invisibly. This function assumes that crs is either an unprojected coordinate system, or that the units of crs are metres. Resulting distances are expressed in kilometres.

---

get_gbif_records	<i>Downloads GBIF records of species</i>
------------------	--

---

**Description**

Downloads GBIF records of species, removing auxiliary columns and records recorded prior to a specified year, or that have coordinate uncertainty above a specified amount.

**Usage**

```
get_gbif_records(
  taxon,
  min_year,
  coord_uncertainty,
  method = c("search", "download"),
  username,
  pwd,
  email,
  retries = 10,
  cleanup = TRUE
)
```

**Arguments**

taxon	Character. Species taxonomic name.
min_year	Integer. The minimum year for which records should be collated. Default is 'NULL', i.e. no minimum.
coord_uncertainty	Integer. The maximum allowable documented coordinate uncertainty (in metres). If specified, this argument will result in removal of any records that have a documented uncertainty above this limit. Records with no uncertainty documented will be returned regardless of the value of 'coord_uncertainty'.
method	Either 'search' (uses the GBIF '/occurrence/search' API endpoint) or 'download' (uses the GBIF '/occurrence/download' API endpoint). The former makes paginated queries to the API, while the latter performs an asynchronous query (but waits for the resulting dataset to be ready for download). The 'search' method is limited to 100,000 records; for large datasets, consider using 'download'. When using 'method='download'', the arguments 'username', 'pwd', and 'email' must be provided.
username	GBIF username, required when method is 'download'.
pwd	GBIF password, required when method is 'download'.
email	Email address, required when 'method = 'download''. This _may_ be used to notify user when download is ready.

retries	If 'method='download'' and file download fails, how many additional attempts should be made to download the file?
cleanup	Logical. Should temporary files associated with 'method='download'' be deleted? Default is 'TRUE'.

### Details

This function is a wrapper of 'rgbif' such that it can be readily used with the 'CoordinateCleaner' package.

### Value

A 'data.frame' of species occurrence records.

---

initialise_raster	<i>Create an empty raster</i>
-------------------	-------------------------------

---

### Description

Create an empty or constant raster with specified attributes.

### Usage

```
initialise_raster(
  x,
  outfile,
  extent,
  res,
  crs,
  init = NA,
  datatype = "FLT4S",
  overwrite = TRUE,
  return_rast = FALSE
)
```

### Arguments

x	Optional. Raster* object or a file path to template raster. If this is provided, extent, res, and crs will be taken from this raster unless they are also passed to this function. If x is not provided, then extent and res must be provided.
outfile	Target raster file path. Directory will be created (recursively) if it doesn't exist.
extent	Either a character path to a raster file, an <a href="#">Extent</a> object (or an object from which such an extent can be extracted), or a numeric vector with four elements giving xmin, xmax, ymin, ymax.
res	Numeric or integer vector giving the horizontal and vertical spatial resolution of the target raster, in units of crs. If a single value is given, it will be used for both horizontal and vertical resolution.
crs	Target coordinate reference system as a PROJ string (character) or an object of class CRS. If missing and x is supplied, the crs of x will be used.
init	Numeric. A value assigned to all cells of the created raster.

<code>datatype</code>	Character. Data type for the created raster. See <a href="#">dataType</a> .
<code>overwrite</code>	Logical. Should outfile be overwritten if it exists?
<code>return_rast</code>	Logical. Should the resulting raster be returned?

### Value

An empty raster is created at outfile, and the corresponding RasterLayer is returned if `return_rast` is TRUE.

---

<code>interactive_map</code>	<i>Produce an interactive html map</i>
------------------------------	--

---

### Description

Produce an interactive html map.

### Usage

```
interactive_map(
  ras,
  layer_name = NULL,
  palette = "inferno",
  transparency = 0.8,
  legend = TRUE,
  set_value_range = NULL,
  discrete = FALSE,
  scale_type = "none",
  outfile = NULL,
  surveillance_locs = NULL,
  pt_col = "red",
  cleanup = FALSE
)
```

### Arguments

<code>ras</code>	A RasterLayer or file path to a raster file.
<code>layer_name</code>	Character. An optional name to assign to <code>ras</code> .
<code>palette</code>	Either a vector of 2 or more colours (e.g. as hex codes or colour names) or the name of a palette function supported by <code>tmap</code> (see <a href="#">palette_explorer</a> and <a href="#">tm_raster</a> ).
<code>transparency</code>	Numeric. Value between 0 and 1 defining the opacity of the plotted raster data (1 = fully opaque; 0 = fully transparent).
<code>legend</code>	Logical. Should a legend be plotted?
<code>set_value_range</code>	A numeric vector giving upper and lower limits for raster values. Values outside this range (including the limits) will be set to NA.
<code>discrete</code>	Logical. Are the values of <code>ras</code> discrete (categorical)?

scale_type	Character. Can be one of: "none" (raw data, no rescaling), "log10", "max normalize" (proportional to maximum value), "minmax normalize" (rescale values to be between 0 and 1 based on min and max) or "logit". Note that if "log10" or "logit" is used, 0 or 1 values must be masked (using set_value_range) or rescaled before passing to this function. scale_type is ignored if discrete is TRUE.
outfile	Character. If NULL, map will be returned to R and not saved. Otherwise, map will be exported as a html file. Full path address must be used. If pandoc is available, a standalone html file is created (see details).
surveillance_locs	A spatial object or a path to a .csv file containing columns named "Latitude" and "Longitude".
pt_col	Character. The plotting colour for surveillance points.
cleanup	Logical. If a standalone html file is created, should accessory files be removed after the standalone file is generated? This will be a folder created within tempdir().

### Details

To create a standalone html file, the **pandoc** software must be installed and available to R. If pandoc is unavailable, the html file will be accompanied by a folder of accessory files.

### Value

A html map.

---

max_normalize	<i>Raster max normalization</i>
---------------	---------------------------------

---

### Description

Rescales raster values such that maximum is 1.

### Usage

```
max_normalize(rast, outfile)
```

### Arguments

rast	Raster object.
outfile	Character. Output raster file path. If not specified, the resulting RasterLayer will be returned to R. Directory will be created recursively if required.

### Value

A normalised RasterLayer will be written to outfile if specified, and returned to R otherwise.

---

min_max_normalize	<i>Raster min-max normalization</i>
-------------------	-------------------------------------

---

**Description**

Performs a min-max normalization on a raster object.

**Usage**

```
min_max_normalize(rast, outfile)
```

**Arguments**

rast	Raster object.
outfile	Character. Output raster file path. If not specified, the resulting RasterLayer will be returned to R. Directory will be created recursively if required.

**Value**

A normalised RasterLayer will be written to outfile if specified, and returned to R otherwise.

---

na_mask	<i>Create a mask raster</i>
---------	-----------------------------

---

**Description**

Create a mask raster, retaining NA and setting all non-NA cells to 1. Input raster will be resampled if necessary, using "max" resampling (i.e. if any contributing cells are not NA, the new cell will be set to 1).

**Usage**

```
na_mask(infile, outfile, res, extent)
```

**Arguments**

infile	Character. File path to input raster file.
outfile	Character. Output raster file path. Directory will be created (recursively) if it doesn't exist.
res	Numeric. A vector of one or two numbers giving the desired output resolution. If missing, the input resolution will be used. If a single number is provided, it will be used for both horizontal and vertical resolution.
extent	One of of: a numeric vector giving the desired extent of the output raster as c(xmin, xmax, ymin, ymax); an <a href="#">Extent</a> object; or an object from which such an Extent object can be extracted.

**Value**

A raster is written to outfile, with NA cells transferred from infile and all other cells set to 1.



---

ncells_4_threshold	<i>Extract number of cells required to meet threshold proportion of risk</i>
--------------------	--

---

**Description**

Extract number of cells required to meet threshold proportion of risk.

**Usage**

```
ncells_4_threshold(
  risk_rasters,
  names,
  proportion_captured = c(0.6, 0.8, 0.9, 0.95)
)
```

**Arguments**

`risk_rasters`     Character. File path(s) to rasters to be loaded.  
`names`             Character. Names corresponding to infiles.  
`proportion_captured`     Numeric vector. Proportion(s) of risk to be captured.

**Value**

A data.frame containing the number of cells to be trapped to capture given proportion(s) of total risk across multiple input files.

---

plan_globals	<i>Create a drake plan describing targets relating to global parameters</i>
--------------	---

---

**Description**

Create a drake plan that describes targets relating to global parameters to be used by individual species plans.

**Usage**

```
plan_globals(
  clum_path,
  nvis_path,
  ndvi_path,
  fertiliser_data_path,
  nrm_path,
  containers_data_path,
  postcode_path,
  airport_beta,
  airport_tsi_beta,
  basemap_mode = c("osm", "boundaries")
)
```

**Arguments**

clum_path	Character. File path to <b>Catchment Scale Land Use of Australia (ACLUM)</b> raster.
nvis_path	Character. File path to the <b>National Vegetation Information System (NVIS)</b> raster dataset.
ndvi_path	Character. File path to the <b>Normalised Difference Vegetation Index (NDVI)</b> raster dataset.
fertiliser_data_path	Character. File path to a csv containing data describing fertiliser use, available <a href="#">here</a> .
nrm_path	Character. File path to shapefile of <b>Natural Resource Management Regions</b> .
containers_data_path	Character. File path to xlsx file containing data about shipping container movements, available <a href="#">here</a> .
postcode_path	Character. File path to postal areas (i.e. post codes) shapefile, available <a href="#">here</a> .
airport_beta	Numeric. Parameter controlling the distribution of tourists (international air passengers) around Australian international airports. Distance to nearest airport is multiplied by this value and exponentiated to give the relative density of tourists at a location. To generate a distribution that ensures proportion $p$ of tourists within distance $d$ of nearest airport, specify $\text{airport\_beta} = \log(p)/d$ (e.g. to have 50 of an airport, use $\log(0.5)/200$ ).
airport_tsi_beta	Numeric. Interpretation is as for <code>airport_beta</code> , but applies to air passengers arriving at Cairns International Airport (CNS) from the Torres Strait Islands.
basemap_mode	Type of basemap for static maps. Either 'osm' (default), or 'boundaries' (polygons delineating borders of states/territories).

**Value**

A drake plan containing targets that generate objects used across species.

---

plot\_establishment\_captured

*Plots the cumulative proportion of establishment likelihood captured*

---

**Description**

Plots the cumulative proportion of establishment likelihood captured as a greater number of top ranking risk cells is considered.

**Usage**

```
plot_establishment_captured(
  data,
  xlab = "Number of cells",
  ylab = "Proportion",
  legend_title = NULL,
  legend = "right",
  prop_line,
```

```

    y_limit = NULL,
    width = NA,
    height = NA,
    units = c("in", "cm", "mm"),
    outfile
  )

```

### Arguments

data	A data.frame as derived from <a href="#">captured_by_ncells</a> .
xlab	Character. x axis label.
ylab	Character. y axis label.
legend_title	Character. Optional legend title.
legend	Legend position. Either a position name (one of "none", "top", "bottom", "left", or "right", or a vector of two normalised coordinates ranging from 0 to 1, e.g. c(1,0) (for bottom-right).
prop_line	Numeric. Proportion in which to add a vertical line. Default is NULL (no line).
y_limit	Numeric vector giving the minimum and maximum y-axis limits. If omitted this will be determined based on the data.
width	Width of plot. If not defined will use size of current graphic device.
height	Height of plot. If not defined will use size of current graphic device.
units	Character. Units corresponding to height and width. Can be "in", "cm", or "mm". Default is inches ("in").
outfile	Character. Output image file path. Containing directory will be created recursively if it does not already exist.

### Value

An image is written to outfile if provided, and otherwise a ggplot object is returned.

---

plot_raster	<i>Plot a raster</i>
-------------	----------------------

---

### Description

Plot a raster.

### Usage

```

plot_raster(
  object,
  legend_title,
  occurrence_data = NULL,
  pt_col = "red",
  height,
  compass = FALSE,
  outfile
)

```

**Arguments**

object	A RasterLayer or file path to raster file.
legend_title	Character. If missing, the name of the raster layer will be used.
occurrence_data	A data.frame, sf object, SpatialPointsDataFrame object, or path to a .csv file containing columns named "Latitude" and "Longitude". If NULL, no points will be plotted.
pt_col	Character. Colour of points (if plotted).
height	Height of plot in inches (will be rendered at 300 dpi). Required if outfile is provided.
compass	Logical. Should a North arrow be shown?
outfile	Character. Path to save output.

**Value**

A tmap object. If outfile is provided, a map will also be written to that file.

---

port_weights	<i>Create raster giving the establishment likelihoods relating to ports.</i>
--------------	--

---

**Description**

Creates spatial port establishment likelihood raster as a function of port container volumes and distance from port.

**Usage**

```
port_weights(template_raster, port_data, beta, outfile, return_rast = FALSE)
```

**Arguments**

template_raster	A RasterLayer or path to supported raster file.
port_data	Character. Path to csv file containing port container volumes and named Latitude and Longitude (in WGS84) for each port of interest.
beta	Numeric. The beta coefficient exponential function of how risk changes with distance from port. beta should be in units per 1km.
outfile	Character. Output raster file path. Directory will be created if it does not exist. If not provided, object will be returned to R.
return_rast	Logical. Should the resulting raster be returned to R? Ignored if outfile is not provided.

**Value**

If outfile is specified, the resulting raster is saved as a geotiff to that path. If return\_rast is TRUE or outfile is not specified the resulting raster is returned, otherwise NULL is returned invisibly.

---

pr_establishment	<i>Calculate probability of establishment</i>
------------------	---

---

**Description**

Calculate probability of establishment for each raster cell as a function of probabilities of arrival and climate and/or biotic suitability.

**Usage**

```
pr_establishment(rast, suitability, outfile, return_rast)
```

**Arguments**

rast	Raster object or path to file containing pathway probability of pest arrival.
suitability	Raster object or path to file containing a suitability raster. Cell values must be between 0 (not suitable) and 1 (ideal suitability).
outfile	Character. Output raster file path. If not provided, the RasterLayer will be returned to R.
return_rast	Logical. Should the RasterLayer be returned to R? Ignored if outfile is not provided.

**Value**

If outfile is specified, the resulting RasterLayer is saved to outfile. If return\_rast is TRUE or outfile is not specified, the resulting RasterLayer is returned, otherwise NULL is returned invisibly.

---

range_bag	<i>Fit and project range bag model</i>
-----------	--

---

**Description**

Fit and project range bag model.

**Usage**

```
range_bag(
  occurrence_data,
  bioclim_dir,
  n_dims = 2,
  n_models = 100,
  p = 0.5,
  exclude_vars = NULL,
  outfile
)
```

## Arguments

occurrence_data	sf object, data.frame or character path to a csv file containing occurrence coordinates (must contain columns named "Latitude" and "Longitude").
bioclim_dir	Path. Path to directory containing WorldClim raster data.
n_dims	Integer. The number of dimensions ranges to bag.
n_models	Integer. The number of bootstrapped model ensembles to run.
p	Numeric between 0 and 1. The proportion of occurrence records to include in bootstrapping .
exclude_vars	Character vector. A vector of bioclim variables to exclude from analysis. Default is NULL.
outfile	Character. Output raster file path. Parent directory will be created recursively if required. If NULL, the RasterLayer will be returned in R.

## Value

A RasterLayer of model predictions is written to outfile if provided, and returned to R otherwise. The raster's extent, resolution and CRS are taken from the raster data in bioclim\_dir. Cell values give the fraction of bootstrapped models for which the cell's environment fell within the species' modelled climate envelope.

## References

This function is a modified version of the rb function provided in Drake, J.M. & Richards, R.L. (2019) [Data from: Estimating environmental suitability](#). Dryad, Dataset, doi:10.5061/dryad.g5p7d1c.

See also: Drake, J.M. (2015) [Range bagging: a new method for ecological niche modelling from presence-only data](#). *Journal of the Royal Society Interface*, 12(107), 20150086. doi:<https://doi.org/10.1098/rsif.2015.0086>

---

rasterize_range	<i>Rasterize species range</i>
-----------------	--------------------------------

---

## Description

Rasterize a species' range, either by burning points directly into raster, or by calculating an alpha hull around points and burning the hull into raster.

## Usage

```
rasterize_range(
  xy,
  method,
  alpha,
  point_buffer = 0,
  template,
  outfile,
  xy_crs,
  plot = TRUE
)
```

**Arguments**

xy	Coordinates to define host/species' range. This must be one of: a matrix with two columns giving longitude and latitude (in that order); a file path to a csv file that contains columns "Longitude" and "Latitude"; a SpatialPoints object; or a sf multipoints object. Points will be reduced to the subset that falls within the template extent.
method	Either points to burn xy points into raster, or alphahull to calculate the alpha hull of xy, and burn those polygons into raster.
alpha	Alpha parameter for alpha hull calculation. Ignored if method is 'points'.
point_buffer	When method = "points", the width of a buffer to generate around xy points before burning into the raster. Expected to be in the CRS of template. Ignored if method is 'alphahull'.
template	RasterLayer or file path to raster file. The host raster resulting from this function will use the extent and resolution of this template.
outfile	Optional file path to write out resulting host raster.
xy_crs	Coordinate reference system of xy, passed as numeric EPSG code, or any other format accepted by <code>sf::st_set_crs</code> . If missing and xy is an sp or sf object, the CRS is defined by the object. If the latter is undefined, or if xy_crs is missing and xy is a matrix or path to a csv file, xy will inherit the CRS of template, and if that is undefined, EPSG:4326 (WGS84) will be assumed. If xy_crs is provided and xy is an sf or sp object with a defined CRS, xy_crs will be ignored.
plot	Logical. Plot the resulting raster? Default = TRUE.

**Value**

A RasterLayer with the resulting range burnt into it. Additionally, if outfile is not missing, the raster is written to that file.

---

rasterize_vector	<i>Rasterize spatial vector dataset</i>
------------------	---

---

**Description**

Converts vector object to a raster with specified extent and resolution.

**Usage**

```
rasterize_vector(
  vector_data,
  outfile,
  template_raster,
  extent,
  res,
  crs,
  field,
  burn_value,
  datatype = "Float32",
  overwrite = FALSE,
  return_rast = FALSE
)
```



**Arguments**

vector_data	Character. File path to vector spatial data.
outfile	Raster output file path. Parent directory will be created recursively if required.
template_raster	Optional. Raster* object or a file path to template raster. If this is provided, extent, res, and crs will be taken from this raster unless they are also passed to this function. If x is not provided, then extent and res must be provided.
extent	Either a character path to a raster file, an <a href="#">Extent</a> object (or an object from which such an extent can be extracted), or a numeric vector with four elements giving xmin, xmax, ymin, ymax.
res	Numeric or integer vector giving the horizontal and vertical spatial resolution of the target raster, in units of crs. If a single value is given, it will be used for both horizontal and vertical resolution.
crs	Target coordinate reference system as a PROJ string (character) an object of class CRS. If missing and template_raster is supplied, the crs of template_raster will be used. If template_raster is also not supplied, the CRS of vector_data will be used.
field	Character. Name of attribute to be burned into raster.
burn_value	Numeric. A constant value to burn into raster.
datatype	Character. Output data type (see <a href="#">gdal_rasterize</a> documentation).
overwrite	Logical. Should outfile be overwritten if it exists?
return_rast	Logical. Return object to R?

**Value**

A binarized raster is written to outfile, and returned to R as a RasterLayer if return\_rast is TRUE.

---

rasterize_wind	<i>Rasterize wind data</i>
----------------	----------------------------

---

**Description**

Process and rasterize wind data, creating raster datasets that describe wind speed within a specified distance of the coastline.

**Usage**

```
rasterize_wind(data, wind_column, template, width, outfile)
```

**Arguments**

data	Character. A vector polygon dataset (e.g. Shapefile or GeoPackage) describing onshore winds, or a file path to such a file.
wind_column	Character. The column name for the column of the object defined by data that contains wind speed data.

template	A RasterLayer or stars object, or a character file path to a file that can be read by GDAL, defining the extent and resolution of analysis, and defining the coastline (edge of non-NA values). If a stars object, the first attribute will be used. Must have a valid coordinate system. CRS is assumed to be Australian Albers (3577).
width	Numeric. The width of the coastal buffer (in metres), defining how far inland the wind has an effect. E.g. if the pest is thought to be carried up to 50km inland by the wind, set this value to 50000.
outfile	Character. The target file path for the wind raster.

---

record_flagger	<i>Flags records that occur in countries that are not reported in either infected_countries or CABI datasheet</i>
----------------	---

---

## Description

‘record\_flagger’ cross references occurrence records against either a vector of country names with established populations specified in ‘infected\_countries’ or those presented in a CABI distribution file. It will then either flag or automatically remove occurrence records that occur in countries outside one of these lists.

## Usage

```
record_flagger(
  occurrence_records,
  infected_countries,
  cabi_ref,
  manual_check = FALSE,
  return_df = FALSE
)
```

## Arguments

occurrence_records	A ‘data.frame’, ‘sf’ object or ‘SpatialPoints*’ object, or a path to a .csv file containing the locations of species of interest. (If a ‘data.frame’ is provided, it must contain the named columns "Latitude" and "Longitude").
infected_countries	Character vector of countries with known established populations.
cabi_ref	Character. The path to a .csv file downloaded from a CABI datasheet containing the country-scale distributional data for the pest. Note that if infected_countries the CABI data will not be used.
manual_check	Logical. Allows interactive selection of which unverified points to retain. Interactive map will only be produced if unverified records are present and if this argument is set to ‘TRUE’.
return_df	Logical. Return the object as a ‘data.frame’ or as an ‘sf’ object (i.e. ‘TRUE’, the default).

**Value**

A 'data.frame' or an 'sf' points object flagging records found in countries in which do not occur in infected\_countries OR is not reported as being present in CABI.

---

rle_compress	<i>Compress raster data using run length encoding</i>
--------------	---

---

**Description**

Compress categorical raster data using run length encoding.

**Usage**

```
rle_compress(x, outfile, quiet = FALSE)
```

**Arguments**

x	File path to the categorical raster to be compressed, or a Raster* object.
outfile	Character (optional). Path to target .rds file that will store RLE results. Directory will be created recursively if it doesn't exist.
quiet	Logical. Should progress messages be suppressed?

**Value**

A list with five elements:

**starts** Cell numbers corresponding to run starts

**lengths** Run lengths

**values** Run values

**extent** Raster extent

**res** Raster resolution

This object is additionally saved in rds format to outfile, if provided.

---

species_plan	<i>Generate a drake plan for a species</i>
--------------	--

---

**Description**

Generate a drake plan that facilitates reproducible generation of species outputs.

**Usage**

```
species_plan(  
  species,  
  clum_classes,  
  nvis_classes,  
  host_path,  
  pathways,  
  include_abiotic_weight = TRUE,  
  climate_suitability_path,  
  exclude_bioclim_vars = NULL,  
  include_ndvi = TRUE,  
  aggregated_res = c(5000, 5000),  
  make_interactive_maps = TRUE,  
  clum_path,  
  nvis_path,  
  ndvi_path,  
  airport_beta = log(0.5)/200,  
  airport_tsi_beta = log(0.5)/10,  
  port_data_path,  
  port_weight_beta,  
  fertiliser_data_path,  
  nrm_path,  
  containers_data_path,  
  postcode_path,  
  occurrence_path,  
  infected_countries,  
  cabi_path,  
  use_gbif = FALSE,  
  gbif_species,  
  gbif_min_year = 1970,  
  gbif_max_uncertainty = 20000,  
  gbif_username,  
  gbif_password,  
  basemap_mode = c("osm", "boundaries"),  
  minimum_probability_for_maps = 1e-05,  
  manual_check_flagged_records = FALSE,  
  wind_effect_width,  
  leakage_tourists,  
  establishment_tourists,  
  leakage_returning,  
  establishment_returning,  
  leakage_torres,  
  establishment_torres,  
  leakage_mail,  
  establishment_mail,  
  leakage_vessels,  
  establishment_vessels,  
  leakage_fertiliser,  
  establishment_fertiliser,  
  leakage_machinery,  
  establishment_machinery,  
  leakage_containers,
```

```

establishment_containers,
leakage_nurserystock,
establishment_nurserystock,
leakage_food,
establishment_food,
leakage_goods,
establishment_goods,
leakage_northwind,
establishment_northwind,
leakage_pacificwind,
establishment_pacificwind,
leakage_nzwind,
establishment_nzwind,
overwrite = TRUE
)

```

### Arguments

species	The name of the species. This will be used for naming output files and folders.
clum_classes	An integer vector indicating which ACLUM classes are considered host plants for species. Either clum_classes or nvis_classes (or both) must be provided. If user_host_path is also provided, the union of the two datasets will be used to define host distribution.
nvis_classes	An integer vector indicating which NVIS classes are considered host plants for species. Either clum_classes or nvis_classes (or both) must be provided.
host_path	Character. Optional file path to a raster dataset that describes the distribution of host material for the species. The file must be readable by GDAL. All cell values other than zero or NA will be considered to be host material. Dataset will be projected and resampled as necessary to match the spatial extent and resolution of analysis. If clum_classes is also provided, the union of the two datasets will be used to define host distribution. A valid coordinate reference system must be associated with the spatial dataset.
pathways	A character vector of invasion pathways that should be included. Can be one or more of: 'containers', 'fertiliser', 'food', 'goods', 'machinery', 'mail', 'nurserystock', 'residents', 'torres', 'tourists', 'vessels', 'northwind', 'pacificwind', 'nzwind'.
include_abiotic_weight	Logical. Should suitability be dependent on climate? Considered TRUE if climate_suitability_path is provided, or if use_gbif is TRUE.
climate_suitability_path	Optional file path to a raster describing climatic suitability across the landscape. If provided, the raster must have the Australian Albers coordinate system (EPSG:3577), spatial resolution of 1000 m, and must have xmin = -1888000, xmax = 2122000, ymin = -4847000, ymax = -1010000. If not provided and include_abiotic_weight is TRUE, a range bag model will be fit to estimate climatic suitability.
exclude_bioclim_vars	Character vector of bioclim variables that should not be used when fitting a range bag model (see <a href="#">range_bag</a> ) of climatic suitability. Variables should be specified as, e.g., c("bio01", "bio12"). Ignored if climate_suitability_path is provided.

include_ndvi	Logical. Should biotic suitability be dependent on NDVI?
aggregated_res	A numeric vector of 2 elements, indicating the desired resolution of aggregated establishment likelihood rasters, in metres.
make_interactive_maps	Logical. Should interactive html maps be generated?
clum_path	Path to the ACLUM raster.
nvis_path	Path to the NVIS raster.
ndvi_path	Path to the NDVI raster.
airport_beta	Numeric. Parameter controlling the distribution of international tourists passengers around international airport. Default is $\log(0.5)/200$ (i.e. 50% of passengers within 200km of airport).
airport_tsi_beta	Numeric. Parameter controlling the distribution of Torres Strait passengers around Cairns airport. Default is $\log(0.5)/10$ (i.e., 50% of passengers within 10km of Cairns airport).
port_data_path	File path to the marine ports .csv file.
port_weight_beta	Numeric. Defines the decay rate of an exponential model. In the context of pests entering via the vessel pathway, this reflects the decrease in the relative likelihood of pest arrival at locations distant from marine ports. For example, $\text{prob\_weight\_beta} = \log(0.5)/10$ would lead to distance-decay that leads to 50% (i.e. 0.5) of establishment likelihood (prior to considering other relevant pathways) within a distance of 10 map units (i.e., 10 kilometres when res is 1000).
fertiliser_data_path	File path to a csv file containing information about fertiliser usage by NRM.
nrm_path	File path to a polygon shapefile of NRMs (natural resource management areas).
containers_data_path	File path to the dataset giving the distribution of containers by postcode.
postcode_path	File path to postal areas shapefile.
occurrence_path	Path to a .csv file containing occurrence data. Must include columns Longitude and Latitude. Coordinates are expected to be in decimal degrees (WGS84).
infected_countries	A character vector of countries within which the species occurs. Ignored if <code>climate_suitability_path</code> is provided. Only one of <code>infected_countries</code> or <code>cabi_path</code> should be provided.
cabi_path	Path to a .csv file downloaded from CABI, indicating the countries within which the species occurs. Download links to these files can be found at the bottom of CABI species datasheet webpages, e.g. <a href="https://www.cabi.org/isc/datasheet/17685">https://www.cabi.org/isc/datasheet/17685</a> . Ignored if <code>climate_suitability_path</code> is provided. Only one of <code>infected_countries</code> or <code>cabi_path</code> should be provided.
use_gbif	Logical. Should species occurrence records be sourced from GBIF? Ignored if <code>climate_suitability_path</code> is provided.
gbif_species	Character vector. Taxon names to use when querying GBIF. Ignored if <code>climate_suitability_path</code> is provided.
gbif_min_year	Integer. The minimum year (yyyy) to be included when downloading GBIF data. Ignored if <code>climate_suitability_path</code> is provided.

gbif_max_uncertainty	Numeric. The maximum permissible coordinate uncertainty for GBIF records. Ignored if <code>climate_suitability_path</code> is provided.
gbif_username	GBIF username to use for querying GBIF's occurrence download API endpoint. If missing, the less efficient "search" endpoint is used. Register at <a href="http://gbif.org">http://gbif.org</a> .
gbif_password	GBIF password to use for querying GBIF's occurrence download API endpoint. If missing, the less efficient "search" endpoint is used. Register at <a href="http://gbif.org">http://gbif.org</a> .
basemap_mode	Type of basemap for static maps. Either 'osm' (default), or 'boundaries' (polygons delineating borders of states/territories).
minimum_probability_for_maps	Numeric. A value between 0 and 1, defining the minimum establishment probability to be displayed on plotted maps. Values below this threshold will be excluded. Default is 1E-5.
manual_check_flagged_records	Logical. Should an interactive map be used for manually checking flagged occurrence records? If TRUE, the user will have the opportunity to select dubious points (i.e. occurrences in countries for which CABI has no record of the species' establishment), to be retained. If FALSE (the default), all such dubious points will be excluded. Ignored if <code>climate_suitability_path</code> is provided. Note that manual checking is not possible when using <a href="#">excel_to_plan</a> since the required interactivity will interrupt plan processing.
wind_effect_width	Numeric. For wind pathways, the distance (in metres) inland from the coastline, over which wind-based arrival applies. E.g. if pests are expected to be carried by wind up to 50 km inland from the coast, use 50000. Ignored if pathways does not include one or more of northwind, pacificwind, or codenzwind.
leakage_tourists, leakage_returning, leakage_torres, leakage_mail, leakage_vessels, leakage_fertil	Numeric vector with length = 2, giving the range (bounds of 95 number of leakage events per year for the pathway.
establishment_tourists, establishment_returning, establishment_torres, establishment_mail, establish	Numeric vector with length = 2, giving the bounds of the 95 of survival & establishment to end of pathway, for leakage events on the pathway.
overwrite	Logical. Should the code executed by the resulting plan be allowed to overwrite existing raster files? Default is TRUE.

## Details

To simplify reproducibility, `edmaps` provides an *Excel* interface for specifying species parameters relevant to estimating establishment likelihood. An example spreadsheet is bundled with the package, available at the path given by `system.file('extdata/parameters.xlsx', package='edmaps')`. The spreadsheet has two sheets, the first specifying "global" parameters that will apply to all species (e.g. file paths to rasters that will), be used regardless of species identity and the second specifying parameters that can vary by species. In the second sheet, each row corresponds to a separate species. Tooltips and data validation guide the user with respect to expected/allowable data.

## See Also

[excel\\_to\\_plan](#)



static\_map

*Produce a static map of establishment likelihood***Description**

Produce a static map of establishment likelihood, with OpenStreetMap base layer.

**Usage**

```
static_map(
  ras,
  xlim,
  ylim,
  layer,
  layer_names,
  legend_title,
  set_value_range,
  scale_type = "none",
  basemap_mode = c("osm", "boundaries"),
  transparency = 0.7,
  colramp_entire_range = TRUE,
  surveillance_locs,
  pt_col = "red",
  aggregate_raster,
  nrow,
  height,
  outfile
)
```

**Arguments**

ras	A Raster* object or file path to a (potentially multiband) raster file.
xlim	Numeric vector. The longitudinal extent of the area to plot.
ylim	Numeric vector. The latitudinal extent of the area to plot.
layer	Character. A layer name to be plotted. Only relevant if loading a RasterStack or RasterBrick.
layer_names	Optional panel titles names for multipanel maps. If not provided, panels will not be given titles.
legend_title	Character. Legend title.
set_value_range	A numeric vector containing an upper and lower bound value (in units of raster). Values outside this range (including values falling on the boundaries) will be masked.
scale_type	Character. Can be: "none" (raw data, no rescaling), "log10", "max normalize" (proportional to maximum value), "minmax normalize" (rescale values to be between zero and 1 based on min and max), or "logit". Note that if "log10" or "logit" is used 0 or 1 values. must be masked (using set_value_range) or rescaled outside of this function.

basemap_mode	Either ‘‘boundaries’’ or ‘‘osm’’ (i.e., OpenStreetMap), defining whether OpenStreetMap imagery should be used for static map basemaps, or simple administrative boundaries. Default is ‘‘osm’’.
transparency	Numeric. Transparency of raster, between 0-1.
colramp_entire_range	Logical. Whether to set colour ramp limits based on national risk range (TRUE) or by risk range present in region specified by xlim and ylim.
surveillance_locs	A spatial object or a path to a .csv file containing columns named "Latitude" and "Longitude".
pt_col	Character. The plotting colour for surveillance points.
aggregate_raster	NULL or a list containing the aggregation factor (i.e. number of raster cells to aggregate) and the aggregation function e.g. <code>list(10, sum)</code> .
nrow	For multipanel plots, an optional numeric value defining the number of rows of panels.
height	height of plot in inches (will be rendered at 300 dpi). If not defined will use size of current graphic device. Width will be determined automatically, based on the aspect ratio given by the plotting extent.
outfile	Character. File path to save map.

### Details

This function relies on the OpenStreetMap package to obtain base layer tiles. This in turn requires Java to be installed, and linking R with Java can sometimes prove challenging. On macOS, Java errors can often be resolved by entering `sudo R CMD javareconf` in a terminal (which updates R's Java configuration variables). On Windows, ensure that the Java architecture (32-bit/64-bit) matches that of R. Additionally, some Java errors arise when using RStudio but not when using R.

### Value

If outfile is provided, a map is saved to that file. Otherwise, a tmap object is returned to R.

---

step_function	<i>Apply a step function to raster values</i>
---------------	---

---

### Description

Reclassify raster values either side of a defined threshold value.

### Usage

```
step_function(rast, threshold, lower_value, upper_value)
```

### Arguments

rast	Raster object
threshold	Numeric. The threshold value.
lower_value	Numeric. The value to set below the threshold.
upper_value	Numeric. The value to set above the threshold.

**Value**

A raster object

---

suitability	<i>Combine abiotic and biotic layers to create suitability raster</i>
-------------	---

---

**Description**

Combine biotic (and abiotic) layers into a single suitability raster to be used to scale arrival rates.

**Usage**

```
suitability(x, outfile, return_rast = FALSE)
```

**Arguments**

x	One of: a RasterStack; a RasterBrick; a list of RasterLayer objects, or a vector of file paths to raster files.
outfile	Character. Name of geotiff where output will be saved. If not provided, raster object will be returned to R.
return_rast	Logical. Should the raster object be returned to R? Ignored if outfile is not provided.

**Value**

If outfile is specified, the resulting raster is saved as a geotiff to that path. If return\_rast is TRUE or outfile is not specified the resulting raster is returned, otherwise NULL is returned invisibly.

---

surveillance_coverage	<i>Calculate the proportion of total establishment likelihood captured by current surveillance</i>
-----------------------	--

---

**Description**

Calculate the proportion of total establishment likelihood captured by current surveillance.

**Usage**

```
surveillance_coverage(  
  establishment_rasters,  
  layer_names,  
  surveillance_locations  
)
```

**Arguments**

establishment_rasters	Character vector. A vector of one or more file path to rasters describing establishment likelihood.
layer_names	Character. Name to be assigned to each establishment raster. If not specified, names of the raster layers will be used.
surveillance_locations	Character. Path to a csv file containing surveillance locations, including columns titled "Longitude" and "Latitude" with coordinates given in decimal degrees (GDA94).

**Value**

The proportion of establishment likelihood captured by current surveillance.

---

threshold_raster	<i>Threshold raster</i>
------------------	-------------------------

---

**Description**

Threshold raster.

**Usage**

```
threshold_raster(rast, threshold, value = 0, outfile)
```

**Arguments**

rast	A file path to a raster file.
threshold	A named list giving the minimum and/or maximum values defining the range of values to retain. Values outside this range will be replaced with value. Can be <code>list(min = Y, max = Z)</code> or <code>list(min = Y)</code> or <code>list(max = Z)</code> .
value	Numeric. The value supplied to cells beyond the threshold(s).
outfile	Character. Output raster file path. Parent directory will be created recursively if required. If <code>outfile</code> is not provided, the resulting RasterLayer will be returned to R.

**Value**

A RasterLayer will be written to `outfile` if provided, and returned to R otherwise.

---

weight_airport_dist	<i>Transform airport distances</i>
---------------------	------------------------------------

---

**Description**

Transform airport distances according to a negative exponential function.

**Usage**

```
weight_airport_dist(  
  airport_dist,  
  beta = log(0.5)/200,  
  outfile,  
  overwrite = FALSE,  
  return_rast = FALSE  
)
```

**Arguments**

airport_dist	Character. File path to a raster file containing proximity to airports. Map units are expected to be kilometres.
beta	Numeric. Parameter passed to the exponential function. Distance to nearest airport is multiplied by this value and exponentiated to give the relative density of tourists at a location. To generate a distribution that ensures proportion $p$ of tourists within distance $d$ of nearest airport, specify <code>airport_beta=log(p)/d</code> (e.g. to have 50 tourists within 200 km of an airport, use <code>log(0.5)/200</code> ).
outfile	Character. Output raster file path. If missing, object will be returned to R.
overwrite	Logical. Overwrite the target raster if it already exists?
return_rast	Logical. Should the raster object be returned to R?

**Value**

If `return_rast` is TRUE, or if `outfile` is missing, the resulting RasterLayer object will be returned. Otherwise NULL is returned invisibly.

# Index

## \* functions estimating arrivals

- arrivals\_by\_containers, 4
- arrivals\_by\_fertiliser, 5
- arrivals\_by\_food, 6
- arrivals\_by\_goods, 7
- arrivals\_by\_machinery, 8
- arrivals\_by\_mail, 9
- arrivals\_by\_nurserystock, 10
- arrivals\_by\_residents, 11
- arrivals\_by\_torres, 12
- arrivals\_by\_tourists, 13
- arrivals\_by\_vessels, 14
- arrivals\_by\_wind, 15

aggregate\_raster, 3

arrivals\_by\_containers, 4, 6–15

arrivals\_by\_fertiliser, 5, 5, 7–15

arrivals\_by\_food, 5, 6, 6, 8–15

arrivals\_by\_goods, 5–7, 7, 9–15

arrivals\_by\_machinery, 5–8, 8, 10–15

arrivals\_by\_mail, 5–9, 9, 11–15

arrivals\_by\_nurserystock, 5–10, 10, 12–15

arrivals\_by\_residents, 5–11, 11, 13–15

arrivals\_by\_torres, 5–12, 12, 14, 15

arrivals\_by\_tourists, 5–13, 13, 15

arrivals\_by\_vessels, 5–14, 14, 15

arrivals\_by\_wind, 5–15, 15

binarize\_and\_aggregate, 16

calc\_EE, 17, 18

calc\_pathway\_pr, 18

calc\_proportion, 18

calculate\_median, 17

captured\_by\_ncells, 19, 37

combine\_arrivals, 19

combine\_pathways, 20

container\_weights, 4, 20

dataType, 32

download\_worldclim2, 21, 25

establishment\_likelihood, 22

excel\_to\_plan, 23, 48

exp\_function, 23

Extent, 29, 31, 34, 42

extract\_highest\_ncells, 24

extract\_worldclim2, 22, 24

fertiliser\_by\_nrm, 25

fertiliser\_weight, 26

fill\_na, 26

focal, 26, 27

focalWeight, 26

gdal\_reproject, 27

get\_airport\_dist, 29

get\_gbif\_records, 30

initialise\_raster, 31

interactive\_map, 32

max\_normalize, 33

min\_max\_normalize, 34

na\_mask, 34

ncells\_4\_threshold, 35

palette\_explorer, 32

plan\_globals, 35

plot\_establishment\_captured, 36

plot\_raster, 37

port\_weights, 38

pr\_establishment, 39

range\_bag, 39, 46

rasterize\_range, 40

rasterize\_vector, 41

rasterize\_wind, 42

record\_flagger, 43

rle\_compress, 16, 44

sf, 4

sf::st\_set\_crs, 41

species\_plan, 23, 44

static\_map, 49

step\_function, 50

suitability, 51

surveillance\_coverage, 51

threshold\_raster, [52](#)

tm\_raster, [32](#)

weight\_airport\_dist, [53](#)