Dispersal Modelling – Quick Start Guide

Last modified: 4 November 2024

Table of Contents

Dispersal Modelling

Simulating the spread of pests and diseases plays a crucial role in informing biosecurity decisions, particularly in the realms of surveillance, impact analysis, resource allocation and the feasibility of control, containment, and eradication strategies. These simulations provide decision-makers with valuable insights into potential outbreak scenarios, allowing them to anticipate and prepare for various threats more effectively.

By simulating the potential population growth and spread patterns of pests and diseases, authorities can identify high-risk areas of subsequent outbreaks, enabling them to strategically allocate surveillance resources. This targeted approach enhances the efficiency of early detection efforts, increasing the likelihood of intercepting threats before they become widely established. Spread simulations also allow biosecurity practitioners to assess how potential economic, environmental, and social impacts of pest and disease incursions may accumulate over time, and how this may change under different management strategies -- facilitating more accurate risk assessments and cost-benefit analyses. This information is invaluable for prioritizing threats and optimizing resource allocation, ensuring that limited biosecurity resources are utilized in the most effective manner possible. The use of spread simulations empowers decision-makers to develop more robust, proactive, and data-driven biosecurity policies and response plans.

Biosecurity Commons offers users a wide range of functionality for modelling both population growth and spread/dispersal dynamics. A brief video demonstration of the workflow is available. For a detailed overview please visit the Dispersal Modelling workflow overview support article.

Dispersal Modelling linkages to other workflows

Outputs of Dispersal Modelling can be used directly as inputs in many other workflows, such as:

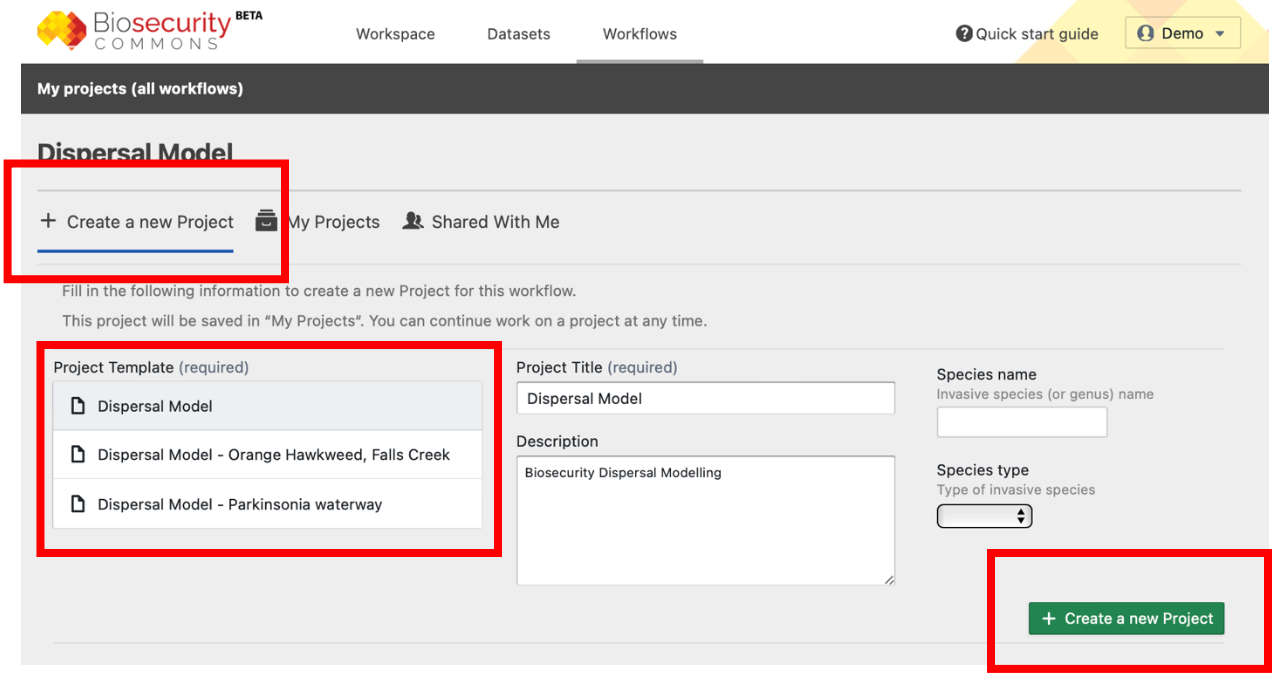
- **Surveillance Design:** For informing location specific likelihoods of presence, and thus, where best to allocate surveillance resources under different spread scenarios for the purposes of early detection, delimitation, containment, and eradication
- **Proof of Freedom:** For providing prior estimates of likelihoods of presence in Bayesian models for quantifying absence likelihoods given surveillance effort
- **Impact Assessment:** For assessing how impacts on agriculture, environment or social amenities may accumulate over time under different spread scenarios
- **Resource Allocation:** For simulating spread under different surveillance and control strategies

Creating a Dispersal Model

Step 1. Create a project

Select the Dispersal Model workflow and then select “Create a new Project” (see screenshot below).

When creating a new dispersal model project, users have the option to select from an empty template, initially titled “Dispersal Model”, which can be renamed appropriately, or users also can use one of a range of prepopulated templates that have been constructed as either examples of the workflow or based on previous case studies (e.g. “Dispersal Model – Orange Hawkweed” or “Dispersal Model – Parkinsonia waterway”).



Fill in the following information to create a new Project for this workflow.

This project will be saved in "My Projects". You can continue work on a project at any time.

Project Template (required)
- Dispersal Model
- Dispersal Model - Orange Hawkweed, Falls Creek
- Dispersal Model - Parkinsonia waterway

Project Title (required): Dispersal Model

Description: Biosecurity Dispersal Modelling

Species name
Invasive species (or genus) name

Species type
Type of invasive species

The empty template is ideal for those wishing to create a brand-new dispersal model as it contains:

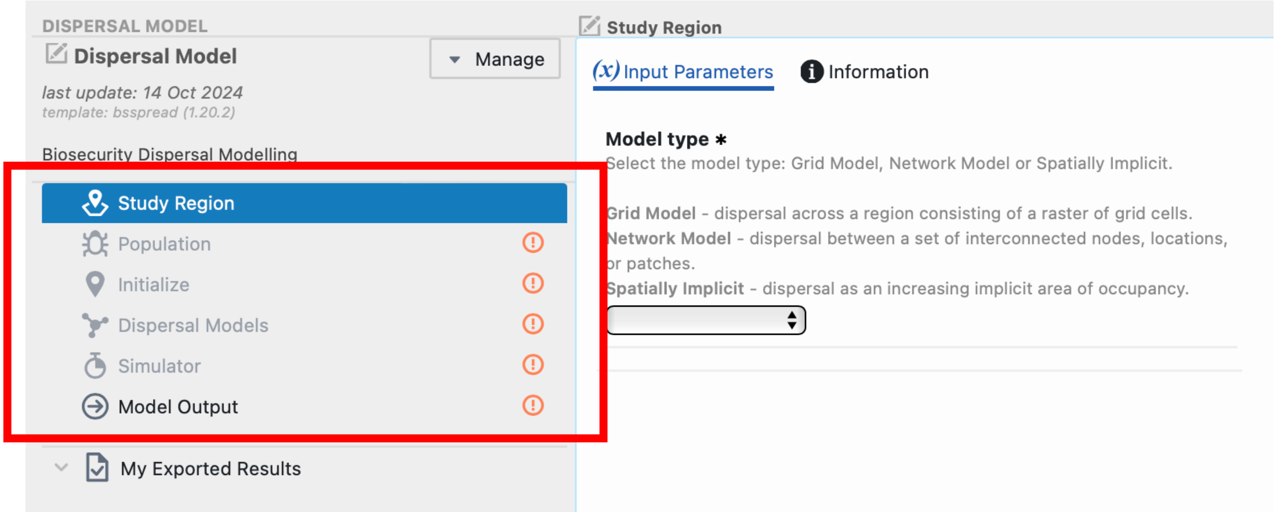
- The basic structure of the dispersal model workflow
- No preloaded datasets

By contrast, example templates provide users with the opportunity to see a completed demonstration of how dispersal models can be produced, or if based on a real-world case study, how others have attempted to create a model.

Select a template and then give your project an appropriate title. Users can also optionally provide additional descriptive details under the Description, Species name and Species type tabs fields. Project title is the only required field to be completed. These metadata are presently unused but will provide future flexibility in filtering and summarising projects.

Once details have been provided, click the green “Create a new Project” button in the bottom right-hand corner to continue.

When you start a Dispersal Modelling workflow from an empty template you will be presented with the core elements of the workflow on the left side of the screen – “Study Region”, “Population”, “Initialize”, “Dispersal Models”, “Simulator” and “Model Output”. As you progress through the project, you must address all the steps in the tree with orange exclamation points which indicate steps that require attention. These change to green ticks when complete.



last update: 14 Oct 2024
template: bsspread (1.20.2)

Model type *

Select the model type: Grid Model, Network Model or Spatially Implicit.

Grid Model - dispersal across a region consisting of a raster of grid cells.
Network Model - dispersal between a set of interconnected nodes, locations, or patches.
Spatially Implicit - dispersal as an increasing implicit area of occupancy.

Step 2. Choose your model type and study region

First, select your model type in the “Study Region” step from the following three options:

- **Grid Model:** Simulate dispersal across a region consisting of a raster of grid cells. *Useful for modelling dispersal across landscapes via functionality provided by a wide variety of rasterised data products*
- **Network Model**: Simulate dispersal between a set of interconnected nodes, locations, or patches. *Useful where the focus is modelling risk between known nodes (e.g. farms, orchards) and risk of movement between nodes are known*
- **Spatially Implicit**: Simulate dispersal as an increasing implicit area of occupancy. *Useful for some rapid risk assessments and benefit-cost analyses that do not have, or do not require, detailed information on spatial constraints*

Depending on the model type the user selects, different options will become available.

1. Grid Model

Selecting “Grid Model” will prompt users to specify the following:

- **Study Region (*Required field*)**
- **Two tier (Default = “Two-tier dispersal”)**
- **Aggregation Factor (Default = 20)**
- **Inner Radius (Default = 5000)**

Study Region

Study Region is a mandatory field for “Grid Models”. Like other workflows, users can click on “Add new Input” and specify their study region in a variety of ways. For details on how to do this, please consult the risk mapping workflow).

Two-tier (and auxiliary parameters Aggregation factor and Inner Radius)

When simulating dispersal events in a Grid Model, users have two main options:

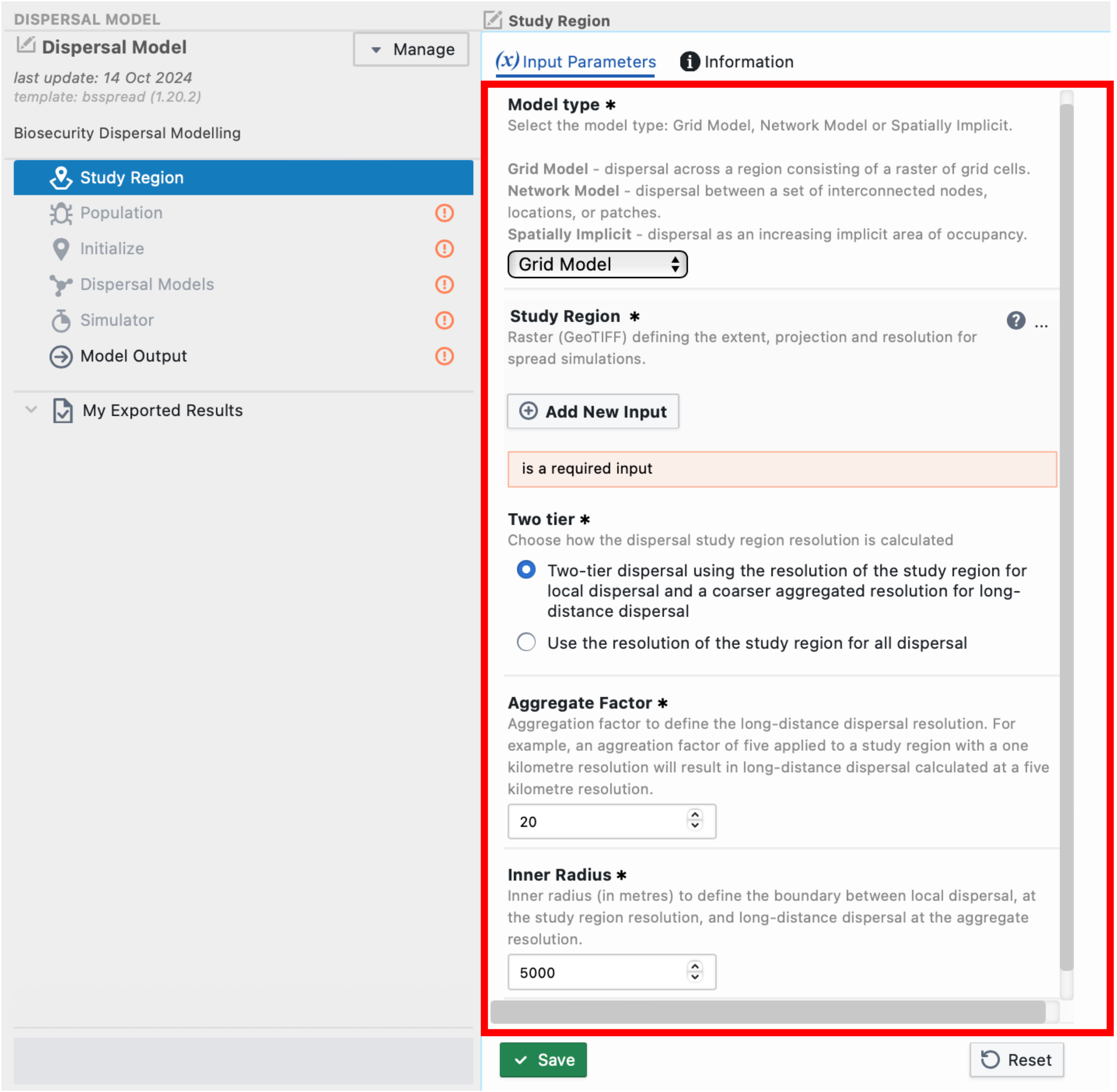
- **Non-two-tiered**: Uses the base resolution defined in the "Study Region" for all dispersal events. This method is suitable for smaller study areas or when computational resources are not a constraint
- **Two-tiered** (Default option)**:** This dispersal approach attempts to optimize computation for large-scale studies by differentiating between local and long-distance dispersal events

The Two-tiered dispersal method works by partitioning dispersal events into two types based on distance:

- **Local dispersal:** Defined by the "**Inner Radius**" parameter, measured in meters. This determines the size of the circle around each occupied cell that retains the original resolution of the study region
- **Long-distance dispersal:** Occur beyond the defined **Inner Radius** and are simulated using a coarser grid, determined by the **Aggregation Factor**

To illustrate this concept, consider a study region with a 1 km² resolution. If the Inner Radius is set to 5 km, the model uses the 1 km² cells for dispersal events that occur within the 5 km Inner Radius. For long-distance dispersal events and an **Aggregation Factor** of 20 is set, this would create 20 x 20 km (400 km²) cells beyond the Inner Radius. When a dispersal event extends beyond the Inner Radius, the model first selects one of these larger 400 km² cells, then randomly chooses a 1 km² cell within it as the specific dispersal location.

This Two-tiered dispersal approach offers several benefits. It significantly reduces computation time for large areas while preserving detailed resolution for local dispersal. It's particularly effective for modeling rare long-distance dispersal events efficiently. Moreover, it allows users to balance computational efficiency with model accuracy by adjusting the Inner Radius and Aggregation Factor. This method is especially useful for simulating dispersal over extensive areas or for species with occasional long-distance dispersal events, making it a valuable tool in large-scale ecological and biosecurity studies. By default, the platform has set Two-tiered sampling on with an **Aggregation factor** set to 20 and an **Inner Radius** set to 5000m.



Dispersal Model

last update: 14 Oct 2024

template: bsspread (1.20.2)

Biosecurity Dispersal Modelling

- Study Region
- Population ⓘ
- Initialize ⓘ
- Dispersal Models ⓘ
- Simulator ⓘ
- Model Output ⓘ
- My Exported Results

Model type *

Select the model type: Grid Model, Network Model or Spatially Implicit.

Grid Model - dispersal across a region consisting of a raster of grid cells.
Network Model - dispersal between a set of interconnected nodes, locations, or patches.
Spatially Implicit - dispersal as an increasing implicit area of occupancy.

Grid Model

Study Region *

Raster (GeoTIFF) defining the extent, projection and resolution for spread simulations.

Add New Input

is a required input

Two tier *

Choose how the dispersal study region resolution is calculated

● Two-tier dispersal using the resolution of the study region for local dispersal and a coarser aggregated resolution for long-distance dispersal

○ Use the resolution of the study region for all dispersal

Aggregate Factor *

Aggregation factor to define the long-distance dispersal resolution. For example, an aggreation factor of five applied to a study region with a one kilometre resolution will result in long-distance dispersal calculated at a five kilometre resolution.

20

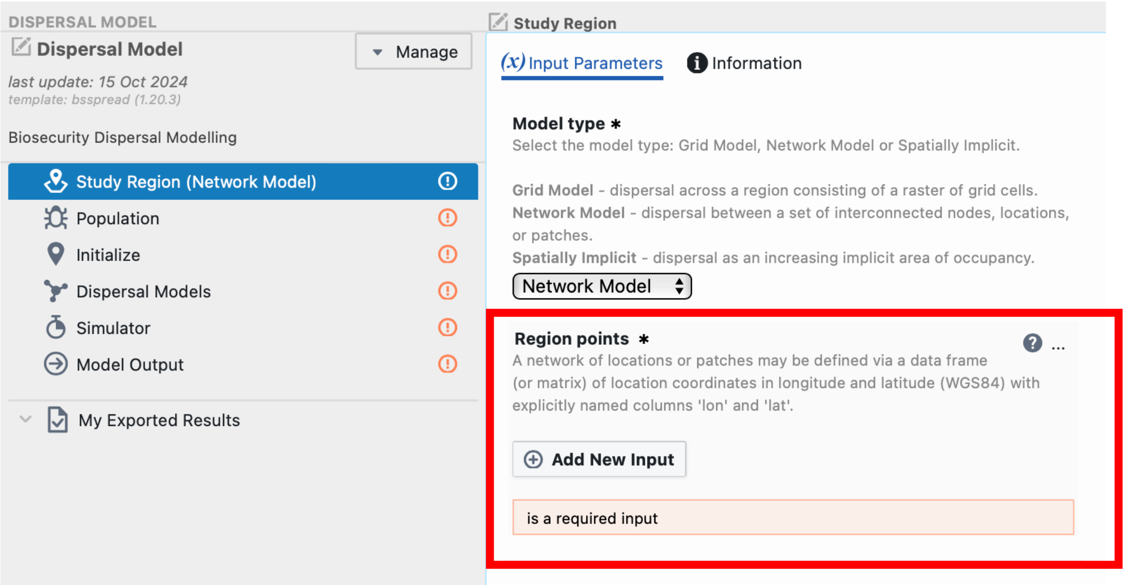
Inner Radius *

Inner radius (in metres) to define the boundary between local dispersal, at the study region resolution, and long-distance dispersal at the aggregate resolution.

5000

2. Network Model

Selecting “Network Model” will prompt users to specify the **Region Points** – a .csv file containing the coordinates of each location to be specified in columns **lat** (latitude) and a **lon** (longitude). Auxiliary columns (such as name) are also accepted but not used by the model.



last update: 15 Oct 2024

template: bsspread (1.20.3)

Biosecurity Dispersal Modelling

Model type *

Select the model type: Grid Model, Network Model or Spatially Implicit.

Grid Model - dispersal across a region consisting of a raster of grid cells.

Network Model - dispersal between a set of interconnected nodes, locations, or patches.

Spatially Implicit - dispersal as an increasing implicit area of occupancy.

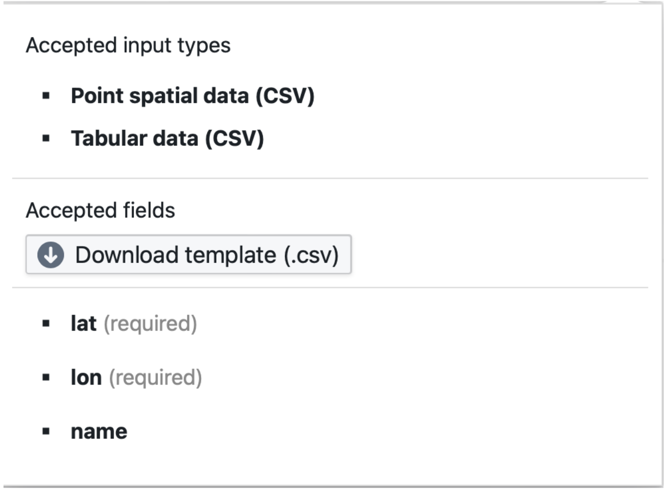
Network Model

Region points *

A network of locations or patches may be defined via a data frame (or matrix) of location coordinates in longitude and latitude (WGS84) with explicitly named columns 'lon' and 'lat'.

is a required input

This file can be added by clicking “Add New Input” and ensuring the uploaded .csv meets the required fields (see below).



Accepted input types

- **Point spatial data (CSV)**
- **Tabular data (CSV)**

Accepted fields

- **lat** (required)
- **lon** (required)
- **name**

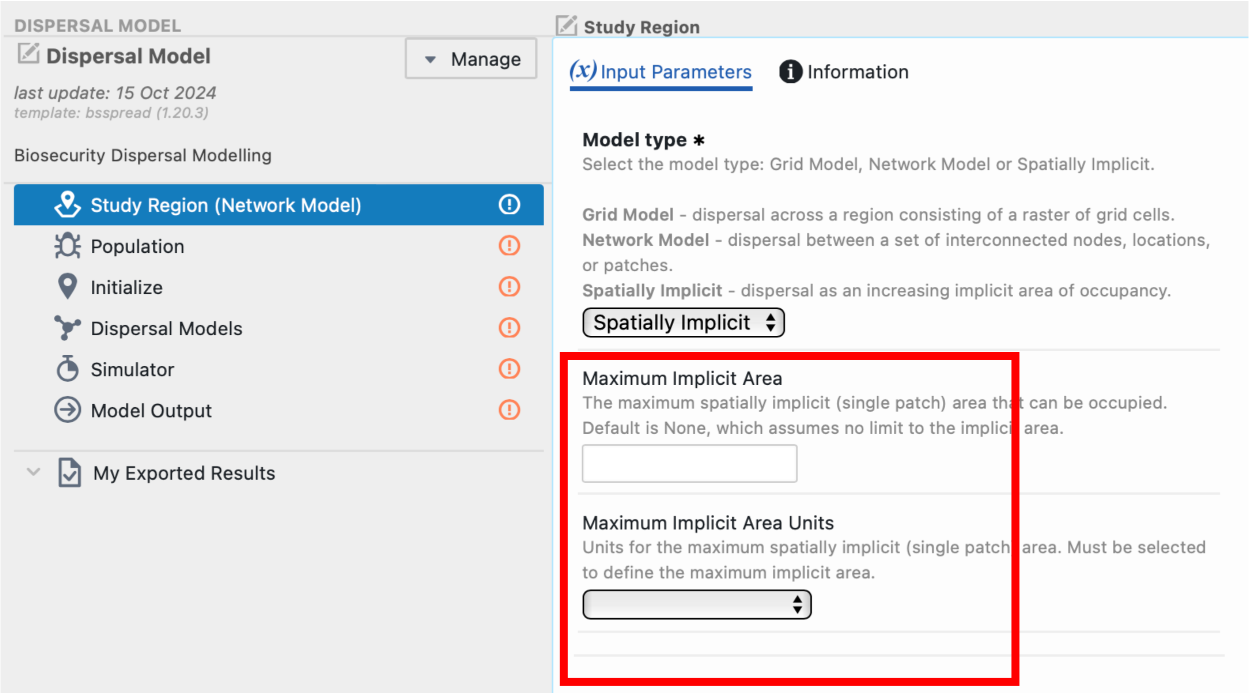
Users may also explore network data that may exist in curated datasets (Via Explore Curated Datasets) or have been previously uploaded for use in other projects via “Explore My Datasets”.

3. Spatially Implicit

Selecting “Spatially Implicit” will prompt users to specify the following:

- **Maximum Implicit Area** (*Optional*) e.g. the size of a single homogenous patch
- **Maximum Implicit Area Units** (Only required if Maximum Implicit Area is selected)

If users do not specify a **Maximum Implicit Area,** the platform will assume there is no limit to the area that can be occupied. Depending on the spread dynamics specified, this can result in unrealistic estimates of area occupied or population sizes, and in extreme cases may reach population sizes beyond integer data limits. As such, we **STRONGLY RECOMMEND** users specify a Maximum Implicit Area.



Model type *
Select the model type: Grid Model, Network Model or Spatially Implicit.

Grid Model - dispersal across a region consisting of a raster of grid cells.
Network Model - dispersal between a set of interconnected nodes, locations, or patches.
Spatially Implicit - dispersal as an increasing implicit area of occupancy.

Spatially Implicit

Maximum Implicit Area
The maximum spatially implicit (single patch) area that can be occupied. Default is None, which assumes no limit to the implicit area.

Maximum Implicit Area Units
Units for the maximum spatially implicit (single patch) area. Must be selected to define the maximum implicit area.

Irrespective of the model type you select, once you’ve provided all required fields, click the green “Save” button and move to the Population step in the tree.

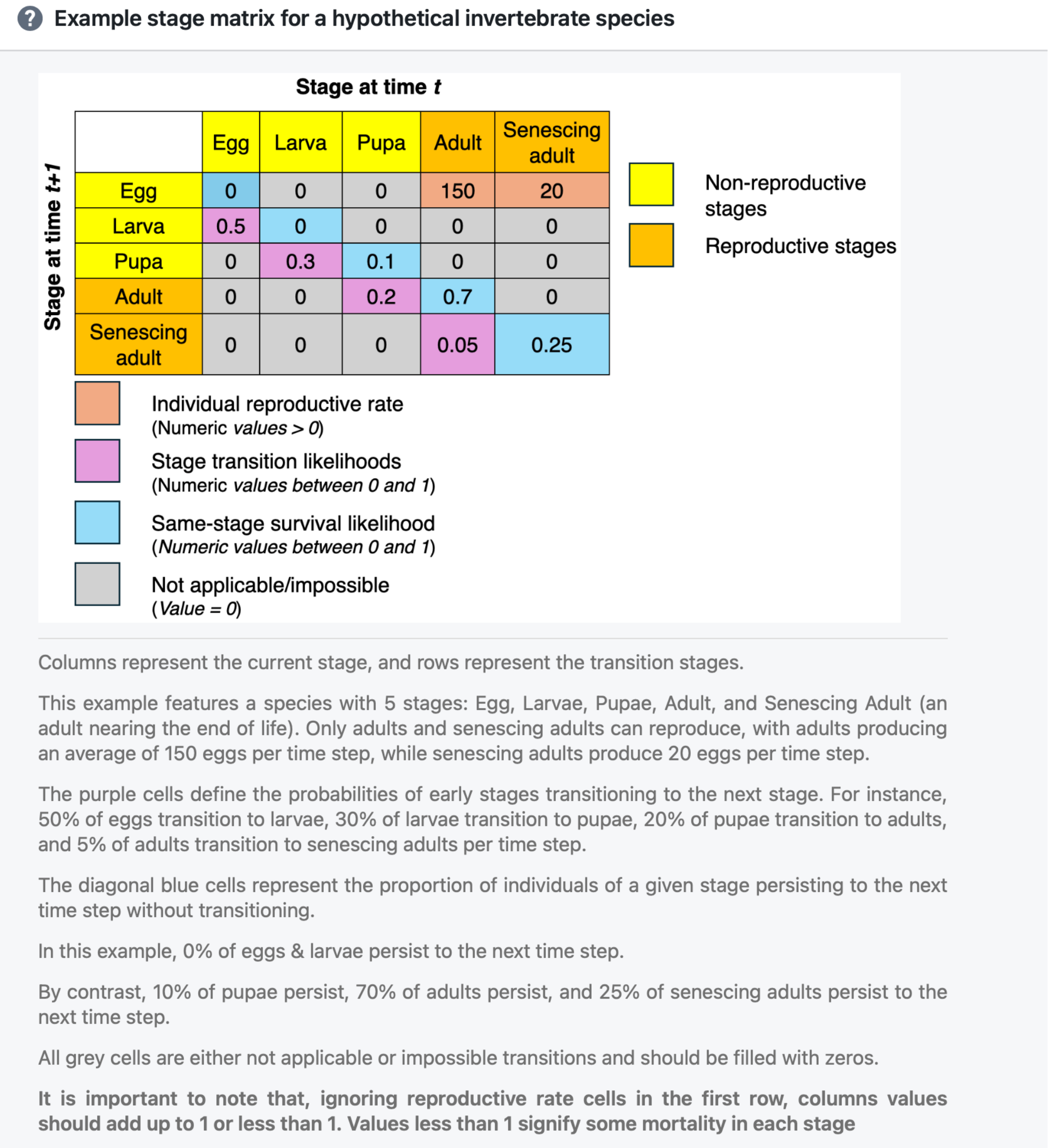
Step 3. Population Model type

The presence/absence or size of a population at a location is a critical factor in Dispersal Modelling. Biosecurity Commons provides users with the ability to use three types of population models within their spread models:

1. **Presence_only:** This is the simplest option. It ignores population structure and produces a binary output indicating only the presence and absence of a threat
2. **Unstructured:** An unstructured population focuses on modelling a population by considering the average population's intrinsic growth rate
3. **Stage_structured:** A stage-structured model requires users to specify a stage or aged-based transition matrix. This indicates how many new individuals (or offspring) will be added to the population at each time step, given species reproductive rates, and the probability of surviving, or transitioning to subsequent stages, at each simulation time step.

The parameterization, and the format in which inputs are required, depend on the model type selected in Step 2, but broadly cover the following:

- **Spread delay** (*Optional: Presence only grid or network models*): Allows users to specify the number of simulation time steps an occupied location must wait before spread can occur. This can be useful for certain threats such as weeds that must reach reproductive maturity before spread can occur
- **Growth** (*Required: Unstructured models*): Intrinsic growth rate or lambda. For example, a growth rate of 1.2 will increase the population size by 20% at each time step (or less when the population is capacity-limited – see below)
- **Growth Matrix** (*Required: Stage-structured models*): A transition matrix by age or stage that describes reproductive rates, stage transition likelihoods and same-stage survival likelihoods (see example below)



Example stage matrix for a hypothetical invertebrate species

Stage at time *t*

Stage at time *t+1*

	Egg	Larva	Pupa	Adult	Senescing adult
Egg	0	0	0	150	20
Larva	0.5	0	0	0	0
Pupa	0	0.3	0.1	0	0
Adult	0	0	0.2	0.7	0
Senescing adult	0	0	0	0.05	0.25

🟨 Non-reproductive stages

🟧 Reproductive stages

🟫 Individual reproductive rate
(Numeric *values > 0*)

🟪 Stage transition likelihoods
(Numeric *values between 0 and 1*)

🟦 Same-stage survival likelihood
(*Numeric values between 0 and 1*)

⬜ Not applicable/impossible
(*Value = 0*)

Columns represent the current stage, and rows represent the transition stages.

This example features a species with 5 stages: Egg, Larvae, Pupae, Adult, and Senescing Adult (an adult nearing the end of life). Only adults and senescing adults can reproduce, with adults producing an average of 150 eggs per time step, while senescing adults produce 20 eggs per time step.

The purple cells define the probabilities of early stages transitioning to the next stage. For instance, 50% of eggs transition to larvae, 30% of larvae transition to pupae, 20% of pupae transition to adults, and 5% of adults transition to senescing adults per time step.

The diagonal blue cells represent the proportion of individuals of a given stage persisting to the next time step without transitioning.

In this example, 0% of eggs & larvae persist to the next time step.

By contrast, 10% of pupae persist, 70% of adults persist, and 25% of senescing adults persist to the next time step.

All grey cells are either not applicable or impossible transitions and should be filled with zeros.

It is important to note that, ignoring reproductive rate cells in the first row, columns values should add up to 1 or less than 1. Values less than 1 signify some mortality in each stage

- **Threat Suitability** (*Optional: all population models for grid and network models*): A raster (Grid model) or point data (Network Model) with values between 0 and 1 that specify the suitability (or establishment likelihood) of each cell/point location. This parameter is useful to prevent or reduce the establishment when the invasive species spread to locations that are otherwise deemed unsuitable or have low suitability. Threat suitability can be based on both abiotic and biotic factors (e.g. see Risk Mapping) or a subset these. If not provided, suitability is assumed to be uniform across grid cells or point locations.
- **Carrying Capacity** (*Optional: Unstructured and Stage-structured models*): Carrying capacity defines the maximum sustainable population size at a given

location. For a spatially implicit model, this is specified as a single number. For Grid-based models it is specified as a raster describing the maximum sustainable population size. For Network models it is a point estimate per location/patch. Note: Often carrying capacity is assumed to be proportional to the suitability of a location. For example, if the carrying capacity in a location with perfect suitability (i.e. score = 1) is 5,000, then the carrying capacity in a cell with a suitability score of 0.5 = 2,500. The platform is agnostic as to how users derive their estimates of carrying capacity.

- **Capacity Stages** (*Optional: Stage***-***structured models*): Users can specify which life-stages/ages are applicable for capacity-limited growth (and survival). For example, the capacity, or limit to the number of plants that can occupy each location, usually only needs to consider the later growth stages, and is generally not affected by the number of seeds or seedlings. A checkbox will appear for each stage/age (i.e. column) specified in the growth matrix for users to select. If none are selected, all stages are assumed to be applicable.
- **Capacity Area Units** (*Required if carrying capacity is set: spatially implicit unstructured or staged structured models):* Options include:
    - *Meters squared*
    - *Kilometers squared*
    - *Maximum area (i.e. Maximum Implicit Area defined in Study region)*

When you have finished specifying a population model and its parameters, click the green “Save” button and move to the Initialize step in the tree.

Step 4. Initialize

The Initialize section allows users to specify the starting conditions in which to simulate spread. Again, how this section is specified is dependent on what was selected in the previous two steps (selecting model type and population type).

“Spatially Implicit” models require minimal initialization input:

- ***None*** (Presence Only): A population is assumed to be present, and as such, this section requires no input from the user
- **Initial Population Size** (*Required: unstructured and stage-structured*): Users will be prompted to specify the initial population size for the implicit space

If a user specifies a "Network" or "Grid" model, users will be prompted to provide information for the following:

- **Initializer type** (Required): Can be either:
    - **Initial_layer**: A raster object (Grid Model) or Point data estimate (Network model) that identifies which cells or locations contain an established population. The initial layer can be specified as 0 (absent) or 1 (present) if using a presence-only model, or the estimated total population size at each location. This is useful when dealing with an existing incursion **having** a known extent. For Grid models, users can also define their initial population by placing points on an interactive map or via drawing a polygon defining the area of incursion extent. Additional parameters for initialization via an initial layer may include:
        - **Initial Age** (Optional: Presence Only models): In some cases, it is useful for specifying the initial age of a population at each location to better account for spread delays that may be related to reproductive maturity. This can be specified as either a single value (applied uniformly across locations), as a raster (for Grid models), or as point data (for network models)
        - **Population Size** (*Required: unstructured and stage-structured*): When unstructured or stage-structured population models have been specified, and the initial layer has been specified using the draw functionality, users will also be prompted to specify the initial population size, which will be distributed across the initial area drawn/selected
        - **First Occupancy Stages** (Optional: stage-structured models): Users can specify which stages are present at the first-time step. Default is all stages are present. A checkbox will appear for each stage/age (i.e. column) specified in the growth matrix for users to select. If none are selected, all stages are assumed to be present at the first time-step
    - **Random**: A raster object (Grid Model) or point data estimate (Network model) that defines the incursion weightings/likelihood at each location. This is useful if one has a risk map (see risk mapping workflow) and is interested in simulating the risk of initial incursion and subsequent spread. Additional parameters for random initialization may include:

  - **Incursion Mean** *(Required: Unstructured and Stage-structured models)*: Numeric mean population size for initial incursion locations. The initial population size is sampled from the Poisson distribution for each incursion location
  - **Incursion Stages** *(Optional: stage-structured models only)*: What stages are capable of dispersing to new locations. A checkbox will appear for each stage/age (i.e. column) specified in the transition matrix. If none are selected, all stages are assumed to be capable of dispersal

When you have finished, click the green “Save” button and move to the “Dispersal Models” step in the tree.

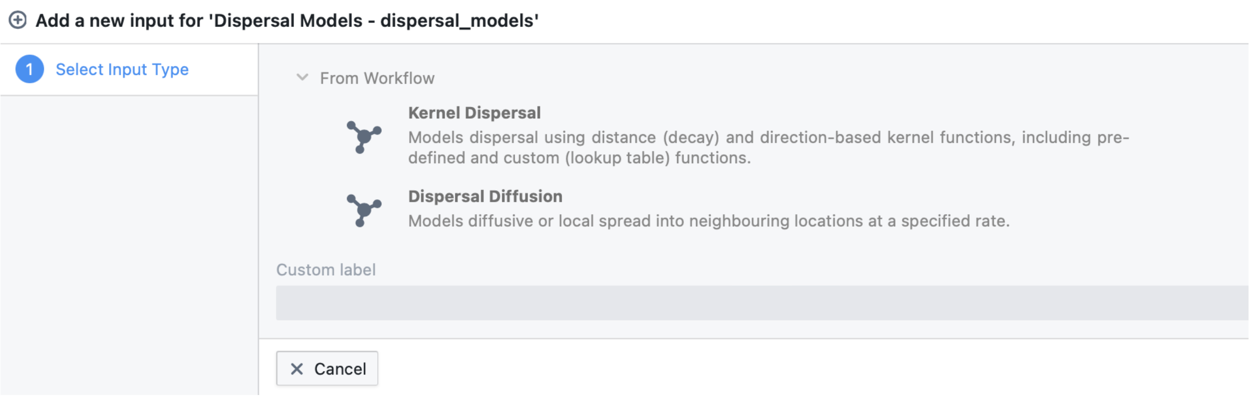
Step 5. Dispersal Models

Threat dispersal and spread can be modelled using many mathematical functions. Biosecurity Commons has attempted to consolidate these into a handful of generic functions that can incorporate a wide range of processes and constraints. These include:

- **Kernel Dispersal** (*Grid and network models, all population models)*: A highly flexible method allowing users to disperse risk as a function of distance decay, direction, spread event likelihoods, attractors, and spatial constraints
- **Dispersal Diffusion** (*Grid models, all populations models*): A simple method for modelling diffusive spread processes as a function of an average spread rate, directional weighting, attractors, and spatial constraints
- **Dispersal Gravity** (*Network models, all population models*): Dispersing risk using gravity-like diffusion between locations based on specific inter-location attraction, the strength of which decays with increasing distance between locations
- **Radial Diffusion** (*Spatially Implicit - Presence Only models*): Models spatial-implicit area as a function of radial diffusion of threat occupancy
- **Area Spread** (*Spatially Implicit - Unstructured or Stage Structured models*): Area increases proportionally to the size of the population, which may be capacity limited. When the population is capacity-limited and the maximum implicit area is specified, the time series graph of the total population size follows a typical “S-curve”

- **Reaction-Diffusion** (*Spatially Implicit - Unstructured or Stage Structured models*): A reaction-diffusion dispersal model is a mathematical framework that describes the spatial and temporal spread of a population by combining two key processes: reaction (population growth) and diffusion (spatial spread). In this model, the change in population density over time is governed by the interplay of intrinsic growth rate (r), carrying capacity (K), and diffusion rate (D). Typically, lower diffusion rates result in lower population growth, since the implicit space available to populations having capacity-limited growth does not expand as quickly as models with higher diffusion rates.

To add a Dispersal Model, simply select “Add New Input”. The functions available to the user depend on the model type and population models specified. For example, if a user has specified “Grid Model” and “Unstructured” population models, then the options available to be selected are Kernel Dispersal and Dispersal Diffusion (see below).



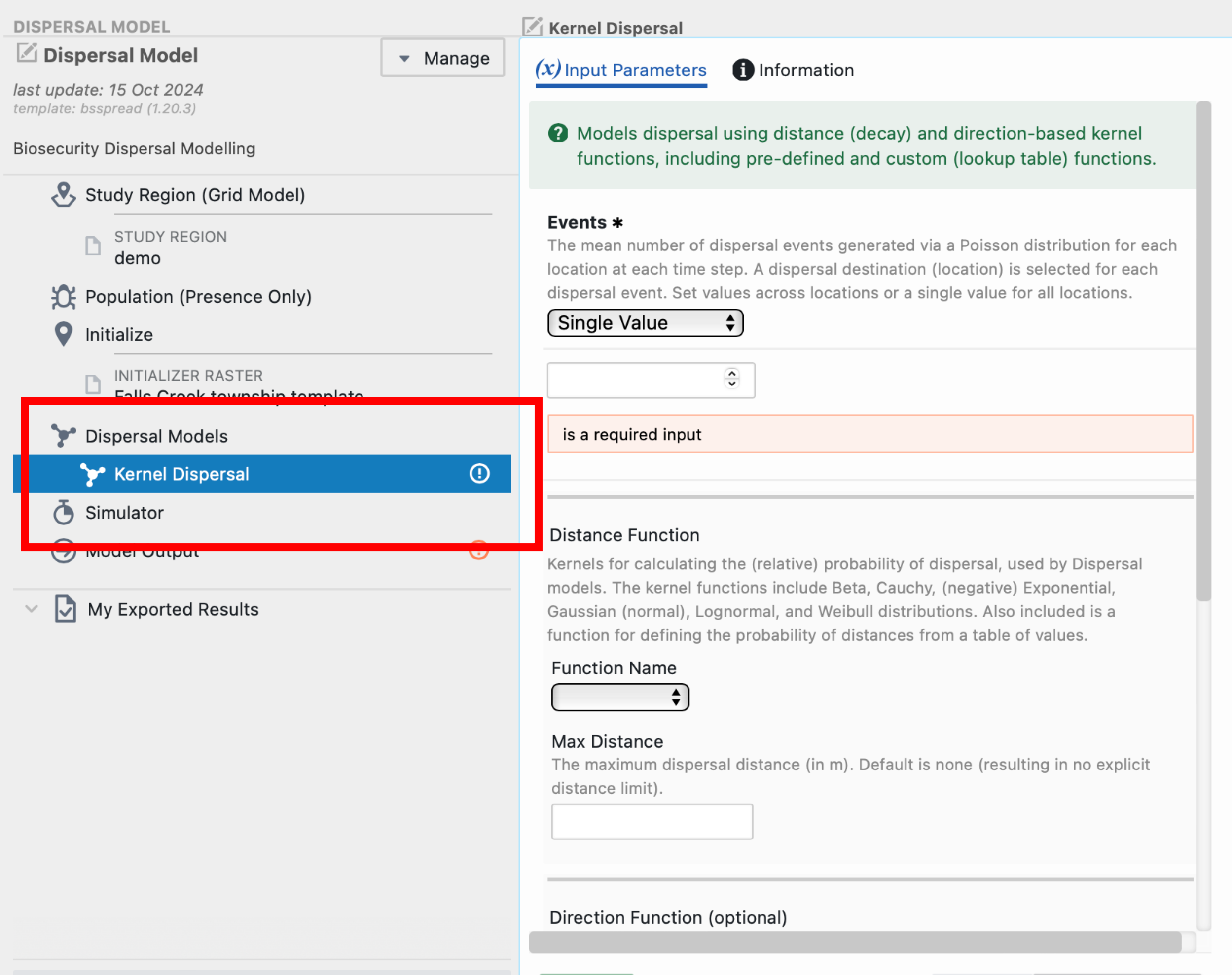
Kernel Dispersal
Models dispersal using distance (decay) and direction-based kernel functions, including pre-defined and custom (lookup table) functions.

Dispersal Diffusion
Models diffusive or local spread into neighbouring locations at a specified rate.

Multiple dispersal functions can be specified by the user, whereby each function may represent a different mode of spread. For example, a user may specify a Dispersal Diffusion function to simulate localized/natural spread and multiple Kernel Dispersal functions to simulate spread caused by different vectors (e.g. human movement, wind dispersal etc).

1. Kernel Dispersal

When users select Kernal Dispersal, a new sub step appears in the tree and multiple modification options will become available (see screenshot below). This model can only be used in spatially explicit contexts (i.e. Grid and Network models).



DISPERSAL MODEL
Dispersal Model
last update: 15 Oct 2024
template: bsspread (1.20.3)
Biosecurity Dispersal Modelling
Study Region (Grid Model)
STUDY REGION
demo
Population (Presence Only)
Initialize
INITIALIZER RASTER
Falls Creek township template
Dispersal Models
Kernel Dispersal
Simulator
Model Output
My Exported Results

Kernel Dispersal
Input Parameters | Information

Models dispersal using distance (decay) and direction-based kernel functions, including pre-defined and custom (lookup table) functions.

Events *
The mean number of dispersal events generated via a Poisson distribution for each location at each time step. A dispersal destination (location) is selected for each dispersal event. Set values across locations or a single value for all locations.
Single Value

is a required input

Distance Function
Kernels for calculating the (relative) probability of dispersal, used by Dispersal models. The kernel functions include Beta, Cauchy, (negative) Exponential, Gaussian (normal), Lognormal, and Weibull distributions. Also included is a function for defining the probability of distances from a table of values.
Function Name

Max Distance
The maximum dispersal distance (in m). Default is none (resulting in no explicit distance limit).

Direction Function (optional)

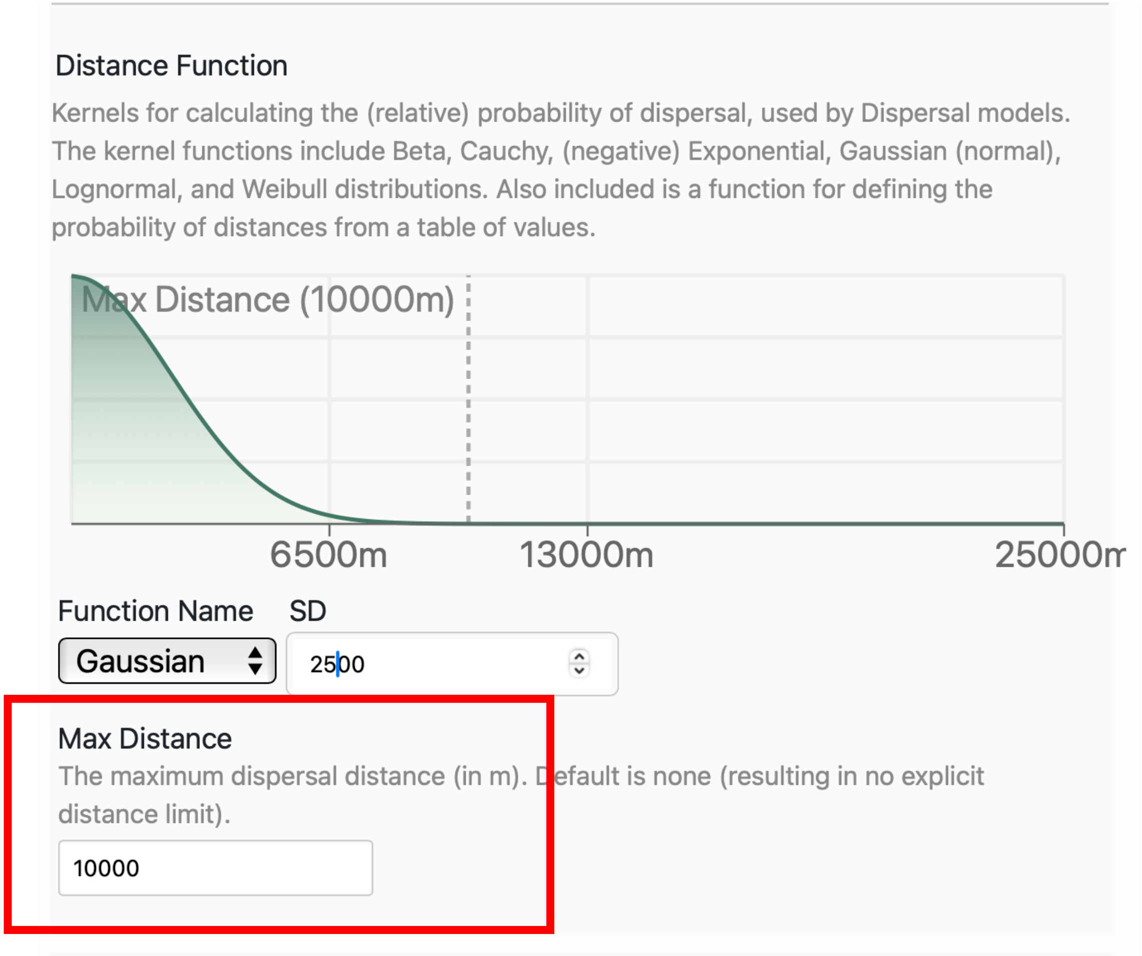
Kernel Dispersal contains a range of required and optional fields that can be filled in. These include:

- **Proportion** (*Required for unstructured or stage-structured populations*): A parameter that specifies the proportion of the population at each location that disperses at each simulation time-step
- **Events** (*Required for presence-only models, optional for unstructured or stage-structured populations*): A parameter that specifies the mean number of dispersal

events for each location at each time step. Numbers are generated using a Poisson distribution. *Users* can specify this as either:

  - **A single value**: Where the mean number is constant across locations
  - **Across locations**: Using either a raster of mean values (Grid Model) or Location/Point specific estimates (Network model)

- **Distance Function** (*Required*): A kernel function describing the relative likelihood a spread event will travel a particular distance. The platform provides users with a wide range of commonly used distribution functions that they can parameterize and visualize on the platform (see below). Distributions included:
  - **Beta** (params: Alpha, Beta, Upper distance limit)
  - **Cauchy** (params: Scale)
  - **Negative exponential** (params: Mean)
  - **Gaussian** (params: Standard Deviation)
  - **Lognormal** (params: Mean, Standard Deviation)
  - **Weibull** (params: Shape, Scale): Using these predefined functions, users can also specify an optional **Maximum Distance** parameter, which prevents dispersal beyond a given distance (measured in meters)



Distance Function

Kernels for calculating the (relative) probability of dispersal, used by Dispersal models. The kernel functions include Beta, Cauchy, (negative) Exponential, Gaussian (normal), Lognormal, and Weibull distributions. Also included is a function for defining the probability of distances from a table of values.

Function Name: Gaussian

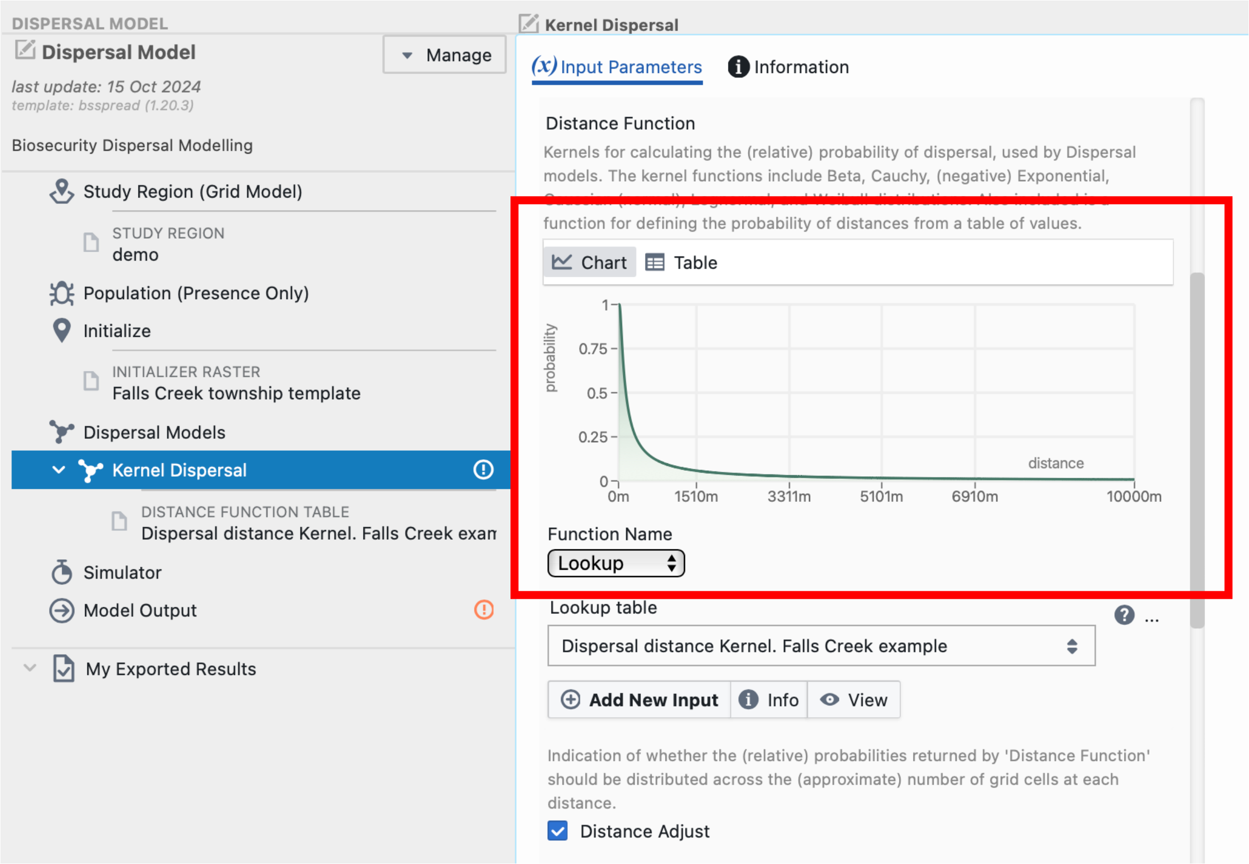
SD: 2500

Max Distance

The maximum dispersal distance (in m). Default is none (resulting in no explicit distance limit).

10000

- **Lookup**: Provides the ability for users to import their own custom kernel via a .csv file. This file requires two columns: “distance” and “probability”. Data should be ordered from shortest distance to longest distance and all distances must contain a probability. When users add the .csv file it will appear in the visualiser as a chart and a table (see screenshot below)



last update: 15 Oct 2024
template: bsspread (1.20.3)

Biosecurity Dispersal Modelling

STUDY REGION
demo

INITIALIZER RASTER
Falls Creek township template

DISTANCE FUNCTION TABLE
Dispersal distance Kernel. Falls Creek exan

Distance Function

Kernels for calculating the (relative) probability of dispersal, used by Dispersal models. The kernel functions include Beta, Cauchy, (negative) Exponential, Gaussian (normal), Lognormal, and Weibull distributions. Also included is a function for defining the probability of distances from a table of values.

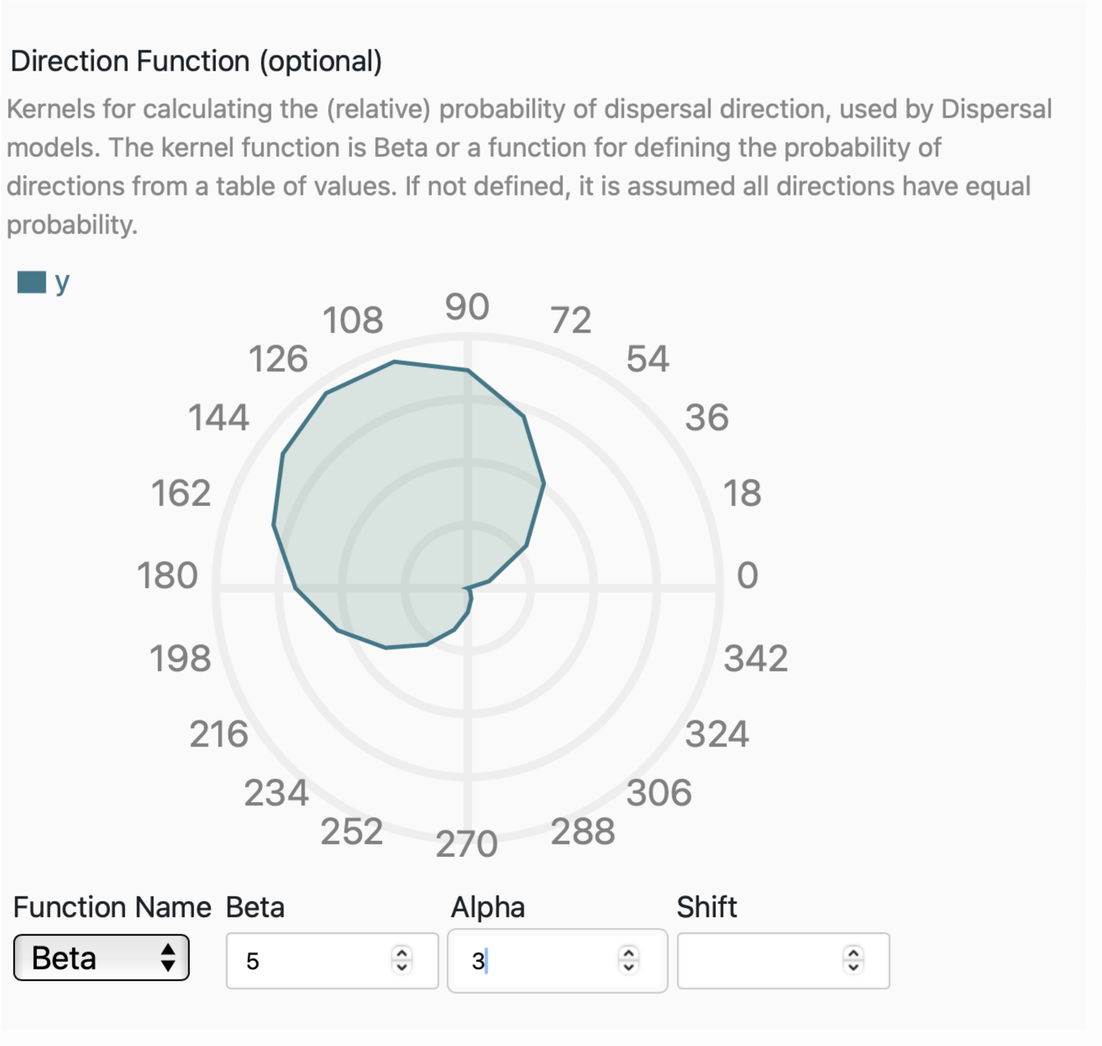
Function Name: Lookup

Lookup table: Dispersal distance Kernel. Falls Creek example

Indication of whether the (relative) probabilities returned by 'Distance Function' should be distributed across the (approximate) number of grid cells at each distance.

☑ Distance Adjust

- **Direction Function** (*Optional*): This specifies the probability of dispersal event moving in one direction over another. By default, dispersal events can move in any direction with equal probability. However, if users wish to specify a directional spread, simply select “Function Name” under Direction Function, and choose between:
    - **Beta** (params: Beta, Alpha, Shift): A simple beta distribution with an optional “Shift” parameter, which can be used to rotate the Beta curve (0-360 degrees). Specifying the parameters of this distribution will automatically be visualized for user interpretation

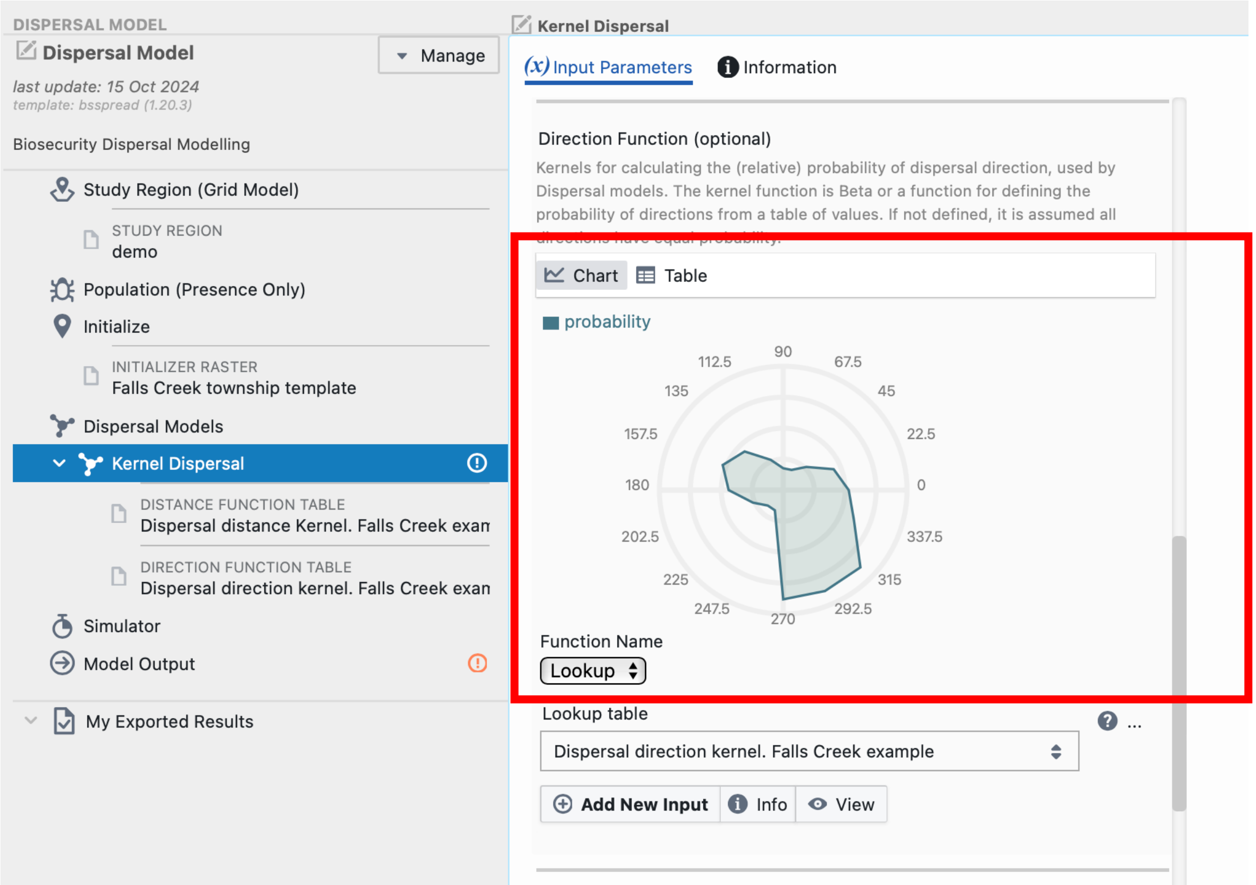


Direction Function (optional)

Kernels for calculating the (relative) probability of dispersal direction, used by Dispersal models. The kernel function is Beta or a function for defining the probability of directions from a table of values. If not defined, it is assumed all directions have equal probability.

Function Name	Beta	Alpha	Shift
Beta	5	3	

- **Lookup:** Provides the ability for users to import their own custom directional kernel via a .csv file. This file requires two columns: “direction” and “probability”. Data should be ordered from smallest to largest direction (measured in degrees) and all directions must contain a probability. When users add the .csv file it will appear in the visualiser as a chart and a table (see screenshot below)



DISPERSAL MODEL

Dispersal Model

last update: 15 Oct 2024
template: bsspread (1.20.3)

Biosecurity Dispersal Modelling

Study Region (Grid Model)

STUDY REGION
demo

Population (Presence Only)

Initialize

INITIALIZER RASTER
Falls Creek township template

Dispersal Models

Kernel Dispersal

DISTANCE FUNCTION TABLE
Dispersal distance Kernel. Falls Creek exan

DIRECTION FUNCTION TABLE
Dispersal direction kernel. Falls Creek exan

Simulator

Model Output

My Exported Results

Kernel Dispersal

Direction Function (optional)

Kernels for calculating the (relative) probability of dispersal direction, used by Dispersal models. The kernel function is Beta or a function for defining the probability of directions from a table of values. If not defined, it is assumed all directions have equal probability.

Function Name: Lookup

Lookup table: Dispersal direction kernel. Falls Creek example

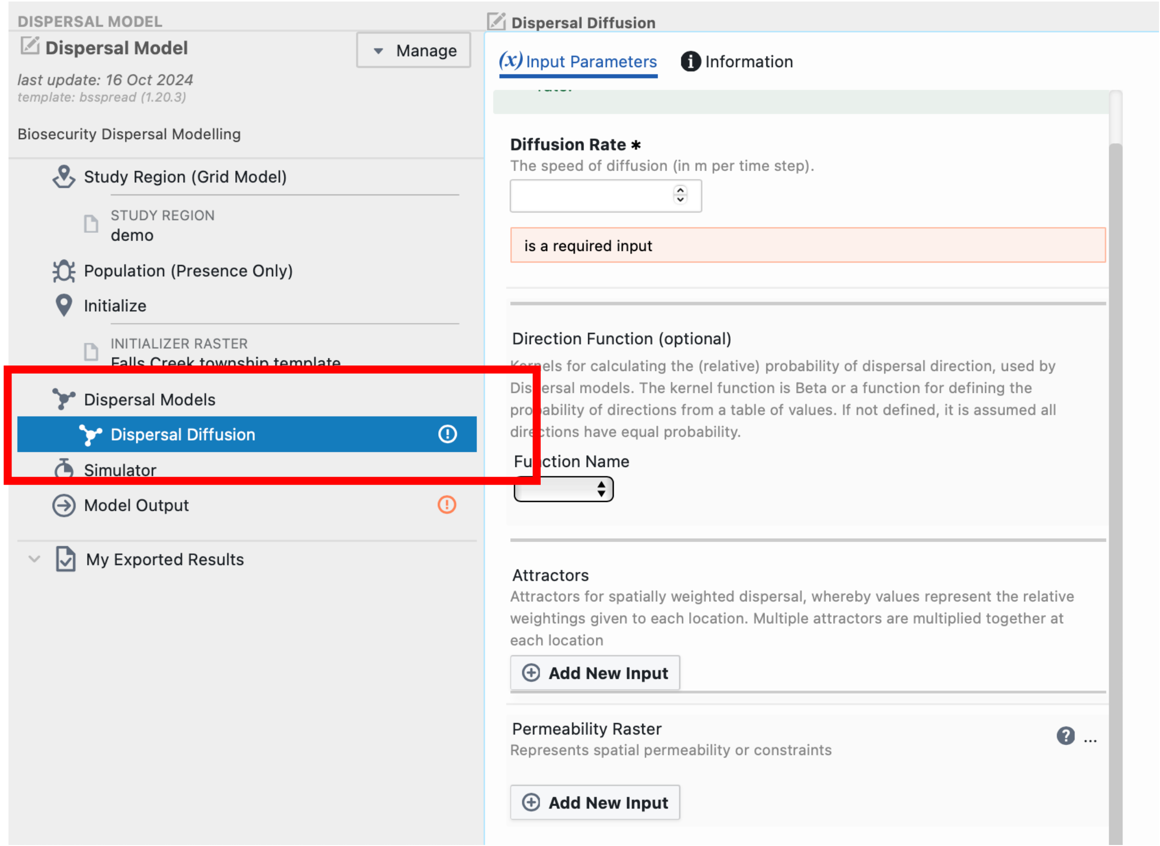
- **Attractors** (*Optional*): Users can specify one or multiple attractors that can be used to inform the relative weightings to be applied to each location during a spread event. Attractors are weightings for probability for destination locations. For “Grid Models” attractors are specified as a raster, for “Network Models” they are specified as point specific relative weightings for each location/node in tabular/csv format. Various attributes can be used to inform attractors including habitat suitability, abundance of host/food source or other landscape variables that may attract vectors of risk material movement (e.g. attractors associated with human movement). For a given kernel multiple attractors can be specified. Where multiple attractors are provided, they are multiplied together at each location to derive an attraction index
- **Network Weights** (*Optional, Network models only*): For Network models, users can specify Bi-directional weights between each pair of connected nodes. These network weights are defined via tabular data defining the path weight between locations (indices), loaded via a .csv file. In this file, each row specifies the bidirectional (relative) weight of moving between patches i and j, indexed in the order defined by in the Study Region. In effect, network weights are broadly

equivalent to permeability in grid models. By default, all nodes are connected with equal weighting

- **Permeability** (*Optional, Grid models only*): A raster of permeability values for each cell. In the context of spread modeling, permeability refers to how easily a species (such as a pest or disease) can move through different types of environments or landscapes. Think of it as a measure of how "permeable" or "passable" the landscape is to the spreading organism. Permeability should be between 0 (spread is completely constrained) and 1 (there is no restriction to spread). Factors such as traffic speed, vegetation density, human density, topographical features, water bodies and human infrastructure can be used to inform spread permeability

2. Dispersal Diffusion

When users select Dispersal Diffusion, a new sub step appears in the tree prompting users to specify required and optional fields (see screenshot below).



last update: 16 Oct 2024
template: bsspread (1.20.3)

The speed of diffusion (in m per time step).

is a required input

Kernels for calculating the (relative) probability of dispersal direction, used by Dis ersal models. The kernel function is Beta or a function for defining the pro ability of directions from a table of values. If not defined, it is assumed all dire tions have equal probability.

Attractors for spatially weighted dispersal, whereby values represent the relative weightings given to each location. Multiple attractors are multiplied together at each location

Represents spatial permeability or constraints

- **Diffusion Rate** (*Required*): The average rate of spread (in meters per time step)

- **Direction Function** (*Optional*): See Kernel Dispersal for description
- **Attractors** (*Optional*): See Kernel Dispersal for description
- **Permeability** (*Optional, Grid models only*): See Kernel Dispersal for description

3. Dispersal Gravity (*Network models*)

Network dispersal models commonly utilize a combination of spatially weighted attractors in combination with distances between nodes to determine the relative likelihood of movement between nodes. When selecting Dispersal Gravity a new sub step appears in the tree prompting users to specify required and optional fields.

- **Proportion** (*Required for unstructured or stage-structured populations*): See Kernel Dispersal for description
- **Events** (*Required for presence-only models, optional for unstructured or stage-structured populations*): See Kernel Dispersal for description
- **Direction Function** (*Optional*): See Kernel Dispersal for description
- **Attractors** (*Required*): See Kernel Dispersal for description
- **Network Weights** (*Optional*): See Dispersal Diffusion for description
- **Beta** (*Required*): A numeric constant for shaping the effect of distance within gravity dispersal, where dispersal = f(attractors)/distance^beta
- **Distance Unit** (*Required*): Unit for distances used within gravity dispersal, i.e. dispersal = f(attractors)/distance^beta. The units selected should be the same for which beta is defined. Default unit is kilometers

4. Radial Diffusion (*Spatially Implicit – Presence Only Models*)

Radial Diffusion is the most simplistic of spread functions that can be applied in a spatially implicit model. Radial diffusion spreads at a constant rate across an implicit homogeneous landscape with no constraints. When selecting Radial Diffusion a new sub step appears in the tree prompting users to specify a single parameter:

- **Diffusion Rate** (*Required*): The average speed at which diffusions occurs (in metres per time step)

5. Area Spread (*Spatially Implicit – unstructured or stage-structured models*)

Area spread is another simple spatially implicit dispersal function for unstructured or stage-based models. It contains no parameters, but rather makes the explicit assumption that area of infestation increases proportionally with population size, which may be capacity limited.

6. Reaction-Diffusion (*Spatially Implicit – unstructured or stage-structured models*)

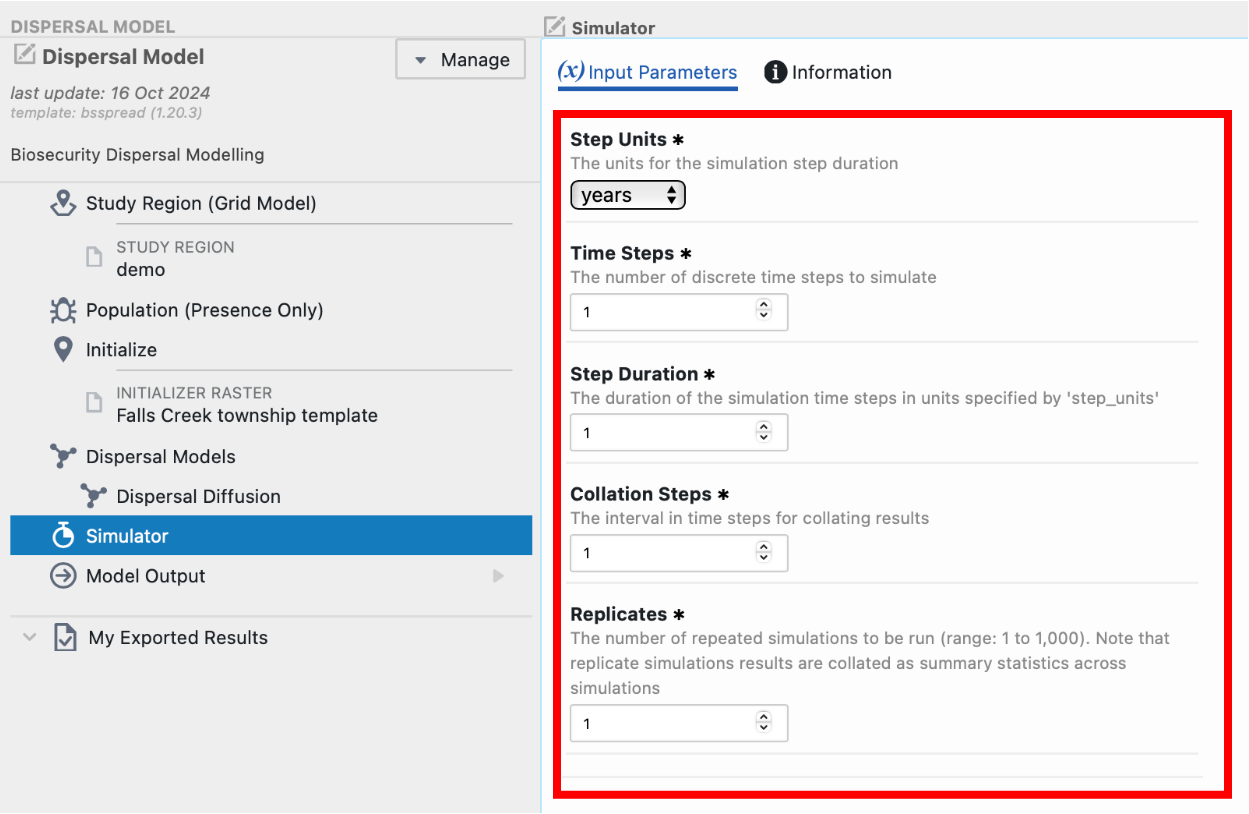
A reaction-diffusion dispersal model is a mathematical framework that describes the spatial and temporal spread of a population by combining two key processes: reaction (population growth) and diffusion (spatial spread). Within this function a single parameter must be specified (with other parameters specified in the population and study area components).

- **Diffusion Rate** (*Required*): The average rate of spread (in metres per time step)

When you have finished, click the green “Save” button and you can move to the “Simulator” step in the tree.

Step 4. Simulator

The Simulator allows users to set the overall parameters for the simulation including the number of time steps, their unit and the number of simulations to be run (see screenshot below).



Step Units *
The units for the simulation step duration
years

Time Steps *
The number of discrete time steps to simulate
1

Step Duration *
The duration of the simulation time steps in units specified by 'step_units'
1

Collation Steps *
The interval in time steps for collating results
1

Replicates *
The number of repeated simulations to be run (range: 1 to 1,000). Note that replicate simulations results are collated as summary statistics across simulations
1

The simulator requires users to specify a range of parameters, including:

- **Step Units** (*Required*)**:** The unit of time in which to simulate dispersal/spread. Options included are:
  - years
  - months
  - weeks
  - days
- **Time Steps** *(Required):* Integer value (Default = 1). How many discrete time steps to simulate spread?
- **Step Duration:** Integer value (Default = 1). This parameter allows users to specify the duration of a step that aligns with the biology of a species (e.g.

generation times). For example, say a species has a generation every six weeks. Users can specify Step Units as weeks, and then specify the Step Duration as six, meaning the step duration simulated is six weeks – a generation time

- **Collation Steps:** Integer value (default = 1, no collation). How many steps to collate. Can be useful when running simulations containing many steps
- **Replicates:** The number of simulation replicates to run. It is important to note that most of the decisions on where organisms disperse to at each time step are chosen based on probabilities, so results will change from one simulation to the next. To get an overview of patterns (means and standard deviations) of possible dispersal you will almost always run simulations multiple times (1,000+ times if you have probabilities with wide-tailed or long-tailed distributions)

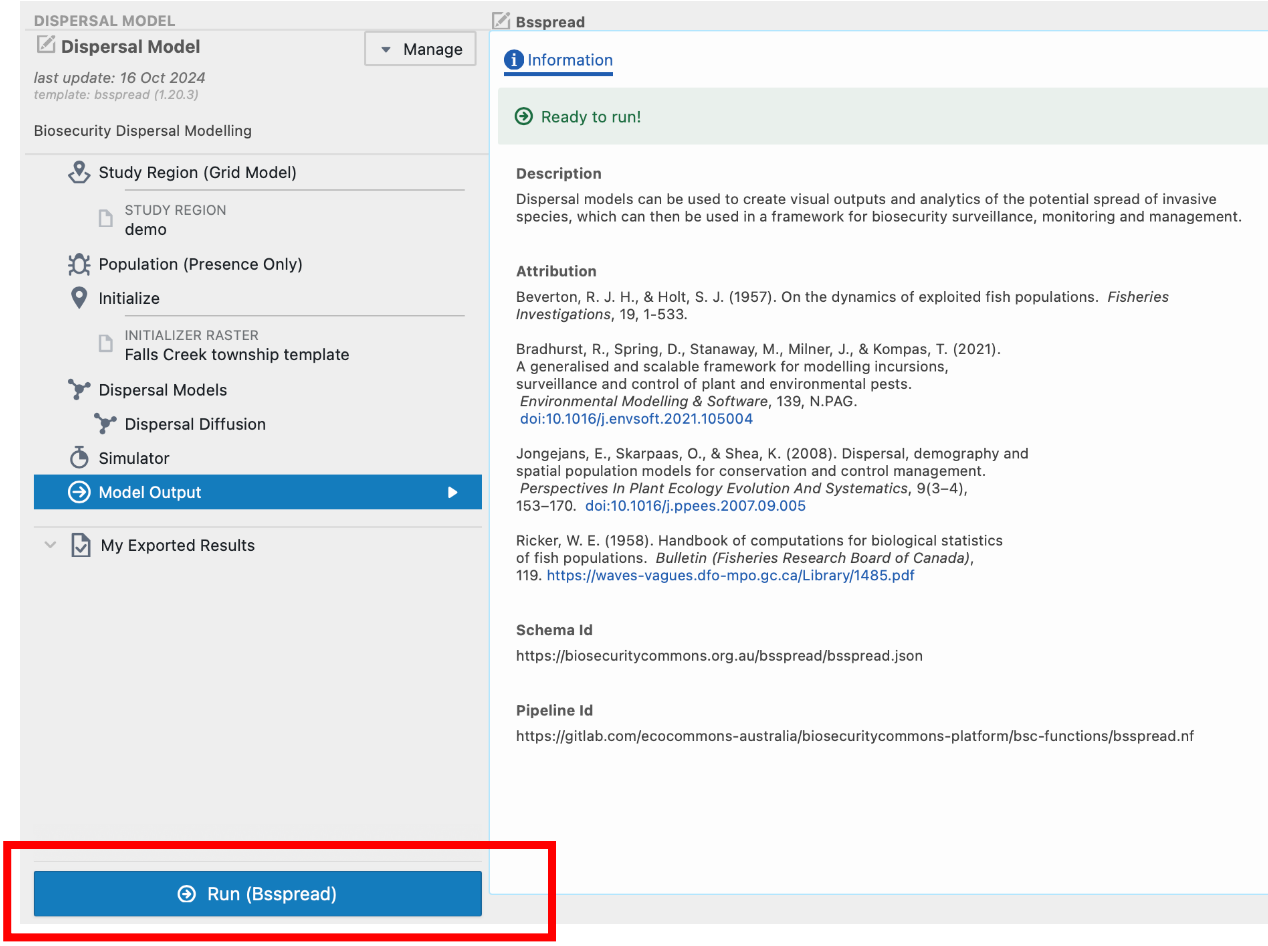
TIP: Setting a high number of time steps and replicates can result in a long running simulation. Try conservative values initially until you are happy with your parameters.

When you are ready click Save and move to the “Model Output” step to complete the experiment.

Step 5. Model Output

This is the last step in the tree. If all your exclamation marks have disappeared, then the model is ready.

Click the blue “Run” button in the bottom left-hand corner to run the workflow (see screenshot below).



last update: 16 Oct 2024
template: bsspread (1.20.3)

Biosecurity Dispersal Modelling

STUDY REGION
demo

INITIALIZER RASTER
Falls Creek township template

Ready to run!

Description

Dispersal models can be used to create visual outputs and analytics of the potential spread of invasive species, which can then be used in a framework for biosecurity surveillance, monitoring and management.

Attribution

Beverton, R. J. H., & Holt, S. J. (1957). On the dynamics of exploited fish populations. *Fisheries Investigations*, 19, 1-533.

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 & Software*, 139, N.PAG. doi:10.1016/j.envsoft.2021.105004

Jongejans, E., Skarpaas, O., & Shea, K. (2008). Dispersal, demography and spatial population models for conservation and control management. *Perspectives In Plant Ecology Evolution And Systematics*, 9(3–4), 153–170. doi:10.1016/j.ppees.2007.09.005

Ricker, W. E. (1958). Handbook of computations for biological statistics of fish populations. *Bulletin (Fisheries Research Board of Canada)*, 119. https://waves-vagues.dfo-mpo.gc.ca/Library/1485.pdf

Schema Id

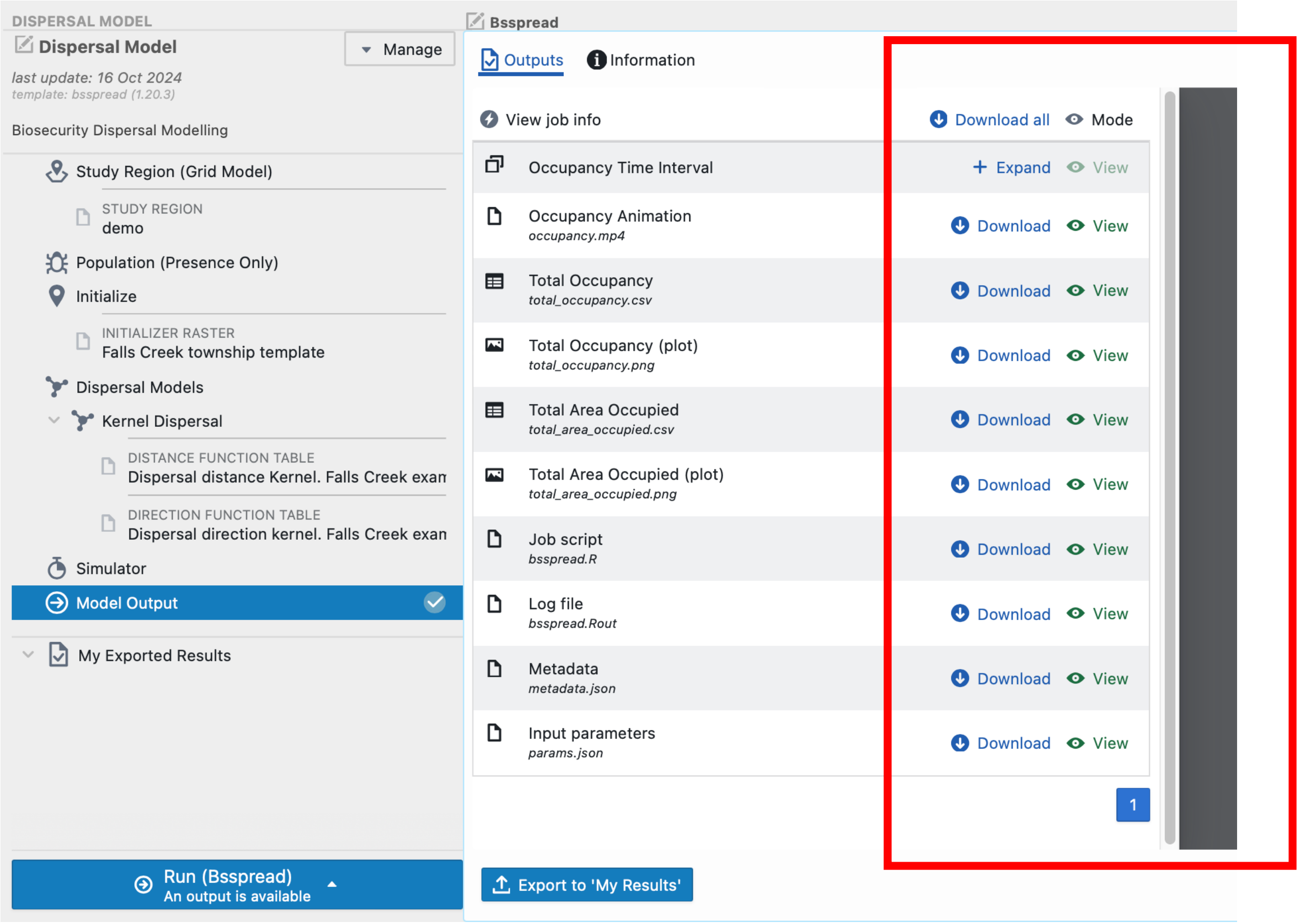
https://biosecuritycommons.org.au/bsspread/bsspread.json

Pipeline Id

https://gitlab.com/ecocommons-australia/biosecuritycommons-platform/bsc-functions/bsspread.nf

The output page will be updated as the job progresses from “Created”, “Submitted”, “Started” and “Success”.

When the job is complete users will be able to view and download all the outputs (example screenshot below).



The exact outputs provided depend on the model type and population model selected by the user. Below is a summary of outputs across all model combinations:

- **Population Time Interval** (*Grid models only*): A folder containing point estimate (if single replicate used) or mean and standard deviation (*if multiple replicates specified*) population size geoTIFFs for each collated time step
- **Population Stage Time Interval** (Stage structured *grid models only with separate outputs for each class*): A folder containing point estimate (if single replicate used) or mean and standard deviation (*if multiple replicates specified*) population size geoTIFFs for each stage class and each collated time step
- **Population Time Series** (*Unstructured and Stage structured models only, Non-Grid):* A csv file containing point estimate (if single replicate used) or mean and standard deviation (*if multiple replicates specified*) population size across each collated time step
- **Population Stage Time Series** (*Stage structured grid models only with separate outputs for each class, Non-Grid*): A csv file containing point estimate (if single

replicate used) or mean and standard deviation (*if multiple replicates specified*) population size for each class across each collated time step

- **Population Animation** (*Grid models only*)***:*** A folder animation of .mp4 files containing point estimate (if single replicate used) or mean and standard deviation (*if multiple replicates specified*) population size across collated time steps
- **Total Population** (*Grid & Network models*)**:** A csv containing point estimate (if single replicate used) or mean and standard deviation (*if multiple replicates specified*) population size across nodes or grid cells across collated time step
- **Population (plot)** (*Spatially Implicit only*)**:** A png file of population growth across collated time steps
- **Total Population (plot)** (*Grid and network models, unstructured and stage structured)***:** A png file of population growth summed across all locations
- **Occupancy Time Interval** (*Grid models only*): A folder containing point estimate (if single replicate used) or mean and standard deviation (*if multiple replicates specified*) occupancy geoTIFFs for each collated time step
- **Occupancy Time Series** (*Non-Grid models):* A .csv file containing point estimate (if single replicate used) or mean and standard deviation (*if multiple replicates specified*) occupancy across each collated time step
- **Occupancy Animation** (*Grid models only*)***:*** A folder animation of .mp4 files containing point estimate (*if single replicate used*) or mean and standard deviation (*if multiple replicates specified*) occupancy across time steps
- **Total Occupancy** (*Grid and Network models*)**:** A .csv containing point estimate (*if single replicate used*) or mean and standard deviation (*if multiple replicates specified*) occupancy across nodes or grid cells for each collated time step
- **Occupancy (plot)** (*Spatially Implicit only*)**:** A png file of occupancy across collated time steps. If multiple reps are specified, mean and 2 x SD are included
- **Total Occupancy (plot)** (*Grid and network models):* A png file of population growth summed across all locations. If multiple reps are specified, mean and 2 x SD are included
- **Area Occupied** (*Spatially implicit):* A .csv containing point estimate (*if single replicate used)* or mean and standard deviation (*if multiple replicates specified*) area occupied across collated time step
- **Total Area Occupied** (*Non-Spatially implicit*)**:** A csv containing point estimate (*if single replicate used*) or mean and standard deviation (*if multiple replicates specified*) area occupied across locations across collated time step
- **Area Occupied (plot)** (*Spatially implicit):* A png file of area occupied across collated time steps. If multiple reps are specified, mean and 2 x SD are included

- **Total Area Occupied (plot)** (*Non-Spatially implicit)***:** A png file of area occupied across collated time steps. If multiple reps are specified, mean and 2 x SD are included
- **Rasterized initial population geometry** (*Only included if initial distribution is drawn on platform*)**:** A geoTIFF of the drawn initial population extent
- **Job script** (*All models*)**:** A copy of the R script used to build the dispersal model
- **Log file** (*All models*)**:** A text data file containing processes, messages, and other details associated with model runs
- **Metadata** (*All models*)**:** A .json file containing the metadata required to run the model on Biosecurity Commons
- **Input parameters** (*All models*)**:** Input parameters required to run the Job Script

Step 6. Exporting outputs for use in other workflows

Users may wish to export outputs for use in other projects or other workflows.

To do this, view the output of interest, and select in the bottom left corner of the interactive map “Export to My Results”.

This output will now be discoverable in the users “My results” database, which in turn, makes the layer available for use in other workflows.