Island biosecurity risk

Software for predicting risk of non-indigenous species arriving and establishing on islands.

User manual for Biosecurity BBN Software Version 1, released 24-07-2017

Cheryl A. Lohr[1], Owen Woodberry[2], and Amelia S. Wenger

[1]Department of Biodiversity, Conservation and Attractions, Science and Conservation Division
[2]Bayesian Intelligence Pty Ltd
[3]School of Earth and Environmental Sciences, University of Queensland, St. Lucia, Queensland 4072

Acknowledgements

This project was funded by the Gorgon Barrow Island Net Conservation Benefits Fund. Many experts contributed knowledge to the processes used to develop this software notably Peter Kendrick, Andrew Burbidge, Allan Burbidge, Dave Algar, Russell Palmer, Carolyn Williams, David Pearson, Corrin Everitt, Georgia Ward-Fear, Ruchira Somaweera, Neil Thomas, Martin Dziminski, Mark Cowan, Judy Dunlop, John Angus, Andy Edwards, Lesley Gibson, Fran Stanley, Michael Lohr, Greg Keighery, Vicki Long, Alicia Whittington, Nicole Godfrey, Kellie Passeretto, John Scott, Andrew Mitchell.

Citation

Lohr, C.A., Woodberry, O., and Wenger, A.S (2017) *Island biosecurity risk: Software for predicting risk of non-indigenous species arriving and establishing on islands.* Department of Biodiversity Conservation and Attractions, Perth, Western Australia.

Executive Summary

Background

Around the globe, islands are the last refuge for many threatened and endemic species. Islands are also important sites for recreation, cultural activities, and industrial development, all of which facilitate the establishment of non-indigenous and/or invasive species (NIS). Surveillance is employed to detect invasive species after their arrival, and to generate data for decisions about follow-up actions. Unless surveillance is prioritised according to risk of establishment of NIS, it may be infeasible to implement efficiently over large tracts of publicly accessible land.

The key biosecurity problem for many regions is one of prioritising sites for surveillance activities and identifying NIS most likely to disperse to, establish, and proliferate on those sites. Common challenges to prioritising sites include a lack of data on the dispersal behaviour of NIS and limited funding or expertise for site-specific modelling in management agencies.

Methods

We created software that automates the creation and computation of species- and site-specific biosecurity Bayesian Belief Networks (BBNs). The *Biosecurity BBN Software* consists of a series of default Bayesian Belief Networks that describe 8 potential dispersal pathways, linked by Java computing code to several comma delimited spreadsheets (.csv) and the freely available GeNIe application. It allows software users to convert relatively straightforward data on NIS and sites into predictions of the risk of NIS arrival and establishment. The *Biosecurity BBN Software* has been designed to be generically applicable to any archipelago, and to any non-parasitic flora or fauna NIS. The *Biosecurity BBN Software* has also been designed to be readily updated and re-run in light of new data.

User Manual

This report acts as a software user manual and provides instructions for downloading and using the *Biosecurity BBN Software*. The authors used the *Biosecurity BBN Software* to predict biosecurity risk for fauna NIS on the Pilbara Islands of Western Australia (Lohr, et al., 2017). We have used this example to illustrate instructions for using the *Biosecurity BBN Software*. In summary the *Biosecurity BBN Software* requires data on site/island attributes, recreational or industrial visitor load, infrastructure, habitat availability, and NIS dispersal behaviours via swimming or flying, industrial or recreational human movement, temporary tidal land bridges, flood plumes and rafts, ocean currents, and wind currents. The *Biosecurity BBN Software* generates spreadsheets with raw and average estimates of the number of NIS arriving on site via each of the dispersal pathways, plus average estimates of the risk of NIS establishing a population on each of the sites.

Recommendations

The outputs of the *Biosecurity BBN Software* were validated extensively against historical data and anecdotes for the Pilbara islands. Parameter uncertainty, particularly in regards to propagule pressure will influence estimates of NIS arrival and establishment (Lohr, et al., 2017). It is recommended that software users refer to the relative risk of NIS arrival and establishment across their sites of interest prior to making any recommendations for NIS management.

ACKNOWLEDGEMENTS **2**

EXECUTIVE SUMMARY **3**

INTRODUCTION **6**

COMPUTER REQUIREMENTS **9**

Hyperlinks to software files *9*
List of input spreadsheets *9*
List of output spreadsheets *9*
List of temporary data calculation spreadsheets *10*

DISPERSAL SUB-MODELS AND INPUT SPREADSHEETS **10**

HUMAN-MEDIATED DISPERSAL BY RECREATIONAL VISITORS 10
SITE OR ISLAND DESIRABILITY 10
island.csv *11*
visitation.csv *13*
RECREATION FAUNA 14
present.csv *14*
threat.csv *15*
RECREATION FLORA 16
region.csv *17*
regionThreatDensity.csv *18*
HUMAN-MEDIATED DISPERSAL BY INDUSTRIAL ACTIVITIES 18
islandCarry.csv *19*
DISPERSAL BY SELF-POWERED SWIMMING OR FLYING 19
islandThreatDensity.csv *20*
distance.csv *21*
DISPERSAL ACROSS TEMPORARY TIDAL-LAND BRIDGES 21
regionBridge.csv *22*
islandBridge.csv *22*
DISPERSAL VIA CATASTROPHIC FLOOD WATERS 22
floodCurrentCarry.csv *23*
DISPERSAL VIA CATASTROPHIC FLOOD WATERS WITH RAFT 23
raftCarry.csv *24*
DISPERSAL VIA NORMAL OCEAN CURRENTS 24
regionCurrentCarry.csv *25*
islandCurrentCarry.csv *26*
DISPERSAL VIA WIND 26
windCarry.csv *27*
regionFavourableWind.csv *27*
islandFavourableWind.csv *28*
islandFavourableWind-temp.csv *28*

SOFTWARE OUTPUTS **28**

ESTIMATED NUMBER OF ARRIVALS 28
arrivals.csv *29*
PROBABILITY OF ESTABLISHMENT 30
establishmentRate.csv *31*
habitat.csv *31*
predcomp.csv *31*
establishment.csv *32*

POTENTIAL APPLICATIONS AND EXTENSIONS OF THE *BIOSECURITY BBN SOFTWARE* **32**

Multi-year dispersal projections via Bernoulli distribution *32*

TROUBLE-SHOOTING **33**

GLOSSARY **34**

BIBLIOGRAPHY **34**

Introduction

Invasive species are one of the leading causes of animal extinctions globally (Clavero & Garcia-Berthou, 2005). The introduction of invasive non-indigenous species (NIS) to natural environments can affect the structure and function of established ecosystems. Ecological interactions between native and invasive species may be direct (e.g., predation, herbivory, parasitism, competition, mutualism) or indirect (e.g., habitat alteration and nutrient cycles, cascading trophic interactions) and result in severely disrupted ecosystem dynamics (Ives & Carpenter, 2007; Sakai, et al., 2001).

As native species have declined, islands have become refuges for native species against threats present on the mainland. Both natural islands and fenced reserves have become the last refuge for several endangered mammals in Western Australia (Morris, et al., 2015). While the small size and relative isolation of islands buffer them from global NIS transport (Cope, et al., 2016), islands remain particularly vulnerable to NIS because their low diversity of native species compared with continents means that they are poorly protected by biotic barriers to naturalization and invasion (Mack & Lonsdale, 2002).

Quarantine and surveillance are two tools employed to restrict the spread of invasive species to new areas. Quarantine programs aim to reduce the risk of NIS arriving and typically focus on the most likely source populations of NIS or dispersal pathway to focus the search efforts of quarantine personnel (Cope, et al., 2016; Garcia-Diaz, et al., 2016; Faulkner, et al., 2017). Quarantine programs are difficult to implement on publicly accessible natural lands because, as the number of source populations and dispersal pathways available to NIS increases, so too does the number of locations and material to be searched and treated, which quickly makes quarantine cost-prohibitive. Domestic quarantine programs tend to rely heavily on education and outreach to raise awareness among the users of natural lands of the potential to spread NIS (Boser, et al., 2014). Surveillance programs, on the other hand, aim to find NIS after their arrival but before their populations become large and potentially uncontrollable, and hence are designed to search high-risk areas more often (Moore, et al., 2010; Whittle, et al., 2013).

The key biosecurity problem for most islands is one of prioritising islands for surveillance activities and identifying NIS most likely to disperse to, establish, and proliferate on islands. Biosecurity risk assessments frequently appear in the form of ranked lists of species that identify the greatest threats for an area (Gordon, et al., 2011; Lohr, et al., 2015; Pheloung, et al., 1999). Many risk assessments, however, are built on limited data, and a limited number of poorly defined species characteristics, or measures related to NIS dispersal (Dahlstrom, et al., 2011). Pest risk maps are a recommended tool for identifying high risk NIS and surveillance sites (Venette, et al., 2010) because modern computational power allows analysts to combine complex models with spatial data sets to produce refined risk-assessment results (Koch, et al., 2013; Kearney, et al., 2008). Unfortunately, these pest risk maps are frequently not designed to be used in data-deficient areas, on poorly studied NIS, or by land managers. Many of these maps are also generated for entire continents, and hence produce results at a resolution that is uninformative for site-specific surveillance programs (Kearney, et al., 2008). Additionally, classical two-dimensional spatial spread models usually depict the spread of invasive species as a diffusion across the landscape, which has limited applicability to long-distance dispersal events which are better depicted by networking models (Koch, et al., 2013; Tingley, et al., 2013), especially across a hostile matrix. These deficiencies may explain some of the failures to implement risk assessments (Dahlstrom, et al., 2011).

Bayesian Belief Networks (BBNs) are graphical models that allow users to represent and reason about an uncertain domain. The nodes in a BBN represent a set of random variables. Directed arcs link pairs of nodes representing the direct relationships between variables (Figure 1). The strength of the relationships is quantified by conditional probability distributions associated with each node

(Korb & Nicholson, 2010). The key advantages of BBNs are that they can be used: to synthesise almost any type of data including, binary, categorical, or continuous data derived from anecdotes of uncertain quality and/or high-resolution experimental data; for scenario analysis and rerun with different alternatives and assumptions (Smith, et al., 2007); and to highlight the sources and implications of uncertainties and help users identify where further research would lead to the greatest improvements in model confidence (Marcot, et al., 2001; Marcot, 2006). BBNs have been used to identify potential translocation sites for endangered species (Laws & Kesler, 2012); predict an endangered species distribution (Smith, et al., 2007), and assess species interactions (Martin, et al., 2005). The disadvantage of most modelling systems, including BBNs is that they can take considerable resources to construct.

In this document we describe new software that automates the creation and computation of species- and site-specific biosecurity Bayesian Belief Networks (BBNs). The *Biosecurity BBN Software* consists of a series of default Bayesian Belief Networks that describe 8 potential dispersal pathways (Figure 1), linked by purpose-built Java computing code to several comma delimited spreadsheets (.csv) and the GeNIe Modeler application. The *Biosecurity BBN Software* allows users to convert relatively straightforward data on NIS and sites into high-resolution predictions of the risk of NIS arrival and establishment. More specifically, the models estimate for each island or discrete site, the number of individuals of each NIS arriving, the number of individuals using each dispersal pathway, and the annual risk of establishment for each NIS, despite uncertainty in data inputs.

Data input spreadsheets define threat and island attributes, along with other characteristics of dispersal pathways depicted in Figure 1. For each NIS, the eight dispersal pathway sub-models (Figure 1-9) are created and run (n = 1000) via stochastic sampling, to generate a distribution of values for the number of arrivals per species per island. For the sake of reporting, these distributions are summarized by a mean and standard deviation in an output spreadsheet that can be used immediately or linked to a geographic information system (GIS) to create a pest risk map. The arrival distributions for each of the dispersal pathways are aggregated for each threat and run through the establishment sub-model (via stochastic sampling, as above) to generate a probability of establishment for each species on each island.

The eight dispersal pathway sub-models are presented individually under the heading *Dispersal sub-models*.

The *Biosecurity BBN Software* has been designed to be generically applicable to any archipelago, and to any non-parasitic flora or fauna NIS. The *Biosecurity BBN Software* has also been designed to be readily updated and re-run in light of new data.

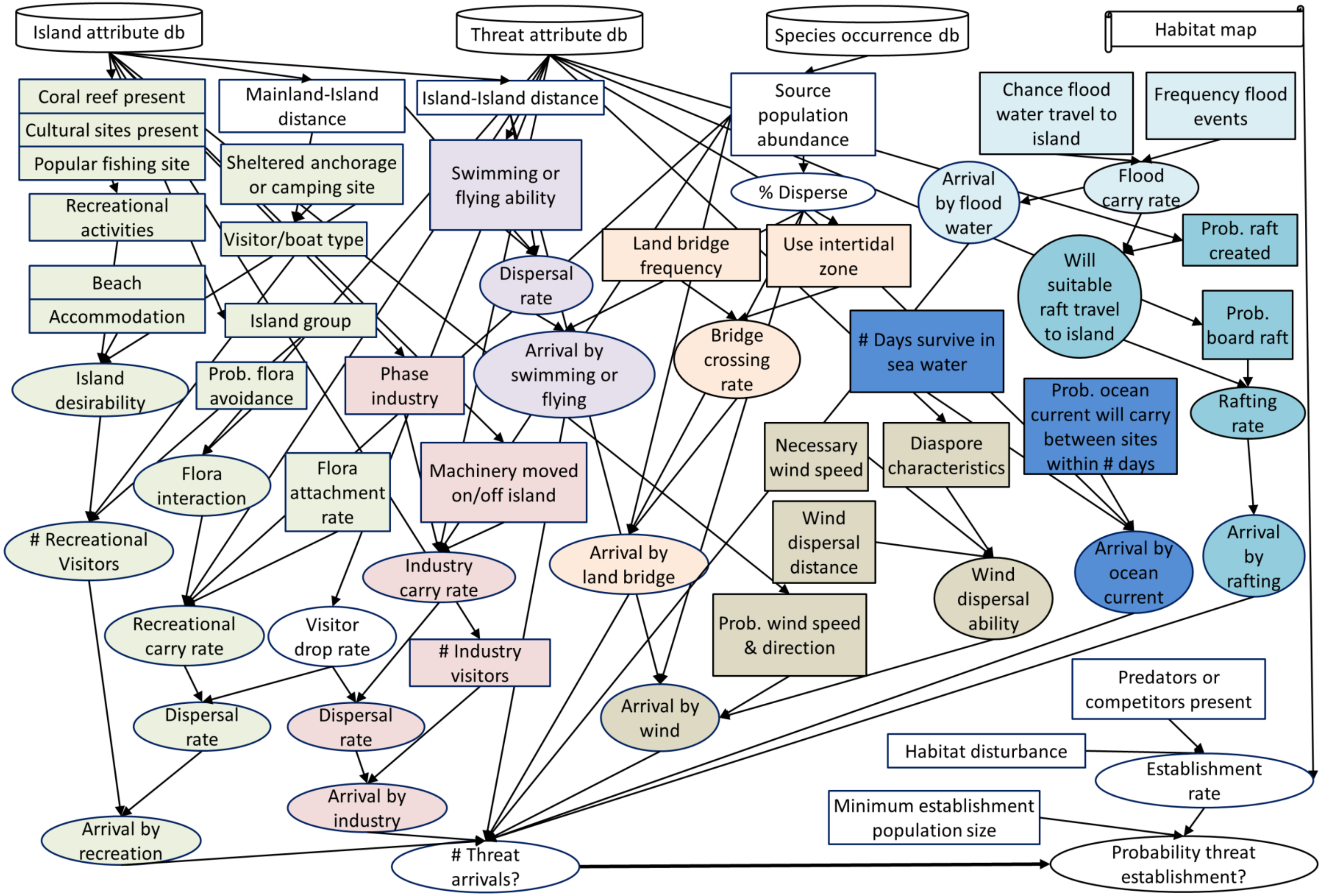


Figure 1: Schematic of the *Biosecurity BBN Software* for the dispersal of NIS to discrete sites. Input nodes are represented by rectangles; calculated nodes are represented by ovals. The eight dispersal pathways (sub-models) are colour-coded. Nodes shared by multiple pathways are white. The sub-model concerning establishment is on the lower right.

Computer requirements

The island biosecurity risk software has the following requirements:

- Microsoft Windows
- Java 7 or later (JRE should be sufficient)
- GeNIe Modeler (https://www.bayesfusion.com/genie-modeler)

The installation package includes libraries for:

- GeNIe & Smile (https://www.bayesfusion.com/genie-modeler)
- Sqlite
- Opencsv
- javajson

Any users of the *Biosecurity BBN Software* may alter value, site names, and NIS names in the input spreadsheets. Only experienced BBN programmers should consider altering the default dispersal sub-models.

List of input spreadsheets

distance.csv
establishmentRate.csv
floodCurrentCarry.csv
habitat.csv
industryCarry.csv
island.csv (columns A:W)
islandBridge.csv
islandCurrentCarryRate1.csv
islandCurrentCarryRate3.csv
islandCurrentCarryRate7.csv
islandFavourableWind.csv
islandThreatDensity.csv
predcomp.csv
present.csv
raftCarry.csv
region.csv
regionBridge.csv
regionCurrentCarryRate1.csv
regionCurrentCarryRate3.csv
regionCurrentCarryRate7.csv
ragionFavourableWind.csv
threat.csv
visitation.csv
windCarry.csv

List of output spreadsheets

island.csv (columns X, Y, Z, and AA)
arrivals.csv
establishment.csv

Dispersal sub-models and input spreadsheets

Default dispersal sub-models are coded with GeNIe Modeler software, a graphical user interface (GUI) which allows for interactive model building and learning, which is owned by BayesFusion, LLC (https://www.bayesfusion.com/genie-modeler). The dispersal sub-models may be viewed by opening the *WAB* sub-folder *nets* and double-clicking on one of the sub-models which will open the sub-model in the GeNIe Modeler. Similarly, sub-models may be modified in the GeNIe Modeler, but caution and experience in BBN programming is required as appropriate nodes, headings, and data categories are required to be present within the sub-models for the Java code to automatically create species- and site-specific BBNs. Most users should limit edits to the data input spreadsheets.

By entering probability values of zero into data input spreadsheets users of the *Biosecurity BBN Software* can effectively turn off selected sub-models. Dispersing NIS (see *Input spreadsheets*) will be forced to travel through any remaining pathways. Regardless of which sub-models users wish to activate, all sites and all NIS must be listed in the NAME columns of all data input spreadsheets. We recommend using copy, pastes, and if necessary transpose functions in spreadsheet software to ensure site names and NIS names are identical and recognisable by the Java code. Any threats and sites listed in island.csv or threat.csv may be turned off without losing data by changing information in the ACTIVE column to 'FALSE'.

Human-mediated dispersal by recreational visitors

Three BBN sub-models are actually required to calculate the probability that recreational site users may introduce NIS: separate models for flora and fauna NIS introductions, per a recreational visitor, and a common model calculating the desirability of the island to determine the number of recreational visitors travelling to/from each site.

Site or Island desirability

All of the nodes in the site or island *Desirability* sub-model are categorical. Users may edit the conditional probabilities by opening the sub-model in GeNIe (Figure 2), followed by opening specific nodes. For example, users may open the node *Desirability* by double-clicking the top-left corner of the node. Under the tab *Definition* users may redefine the characteristics of highly desirable, moderately desirable, and undesirable sites for recreational visitors within the constraints of the pre-defined categories (Figure 3). Similarly, the type of vehicle/boat used to access the site is described using categorical data. In our case study (Lohr, et al., 2017) we separated recreational boats into three different sizes: small (<5m length), medium (5-10m in length), and big (>10m in length) (Figure 4). Some boats cannot access particular types of islands (binary data=0) due to the benthic topography. Values within these nodes may be edited without expertise in BBN programming. User should not edit node names or category names unless they have experience in BBN programming.

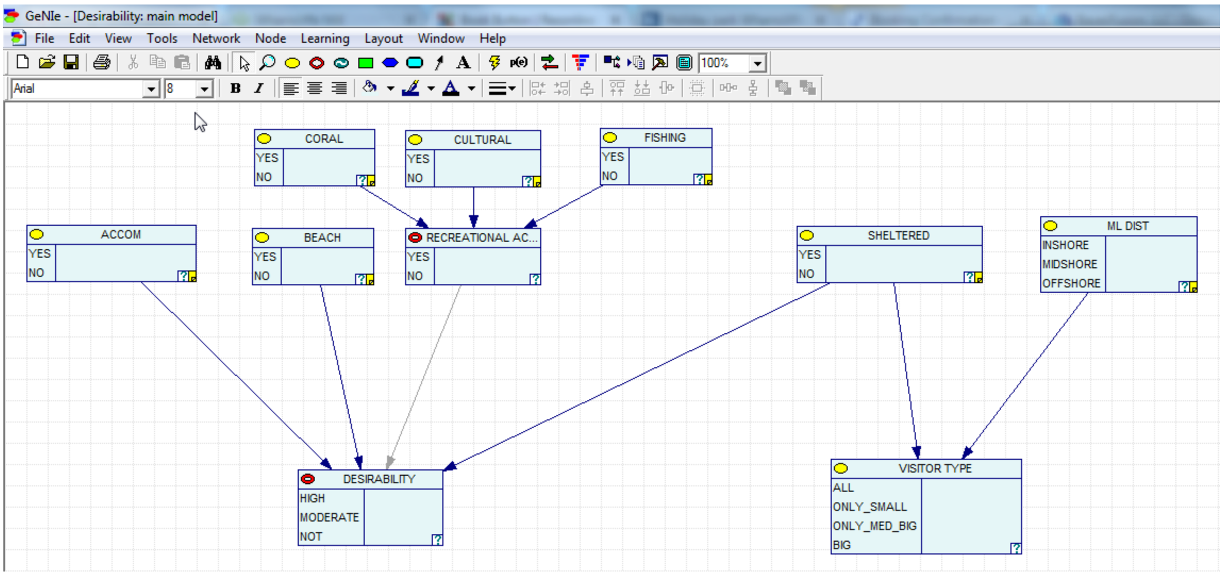


Figure 2: Screen-capture of the island desirability sub-model when opened in GeNIe. Rectangular nodes contain categorical conditional probabilities.

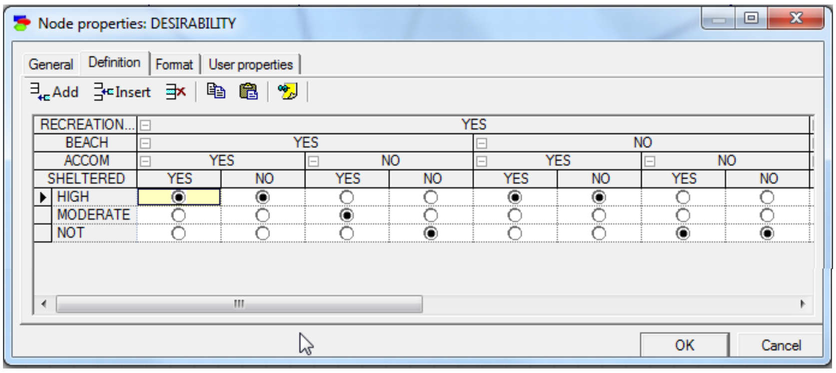
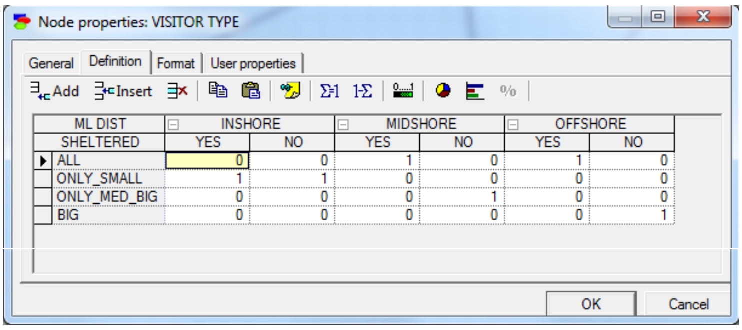


Figure 3: Screen-capture of the node *Desirability* from within the sub-model *Desirability.*

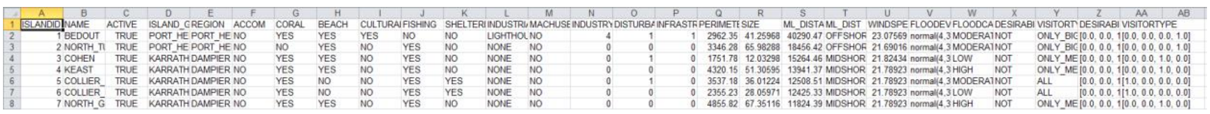


ML DIST	INSHORE		MIDSHORE		OFFSHORE	
SHELTERED	YES	NO	YES	NO	YES	NO
ALL	0	0	1	0	1	0
ONLY_SMALL	1	1	0	0	0	0
ONLY_MED_BIG	0	0	0	1	0	0
BIG	0	0	0	0	0	1

Figure 4: Binary data defining what type of vehicle or boat may be used to access each type of site. In this example, only small boats (<5m in length) may access inshore islands along the Pilbara coastline.

Two data input spreadsheets are used to estimate the number of visitors to each site: *island* and *visitation*.

island.csv



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	ISLANDID	NAME	ACTIVE	ISLAND_G	REGION	ACCOM	CORAL	BEACH	CULTURAI	FISHING	SHELTERI	INDUSTRI	MACHUSE	INDUSTRY
2	1	BEDOUT	TRUE	PORT_HE	PORT_HE	NO	YES	YES	YES	NO	NO	LIGHTHOU	NO	4
3	2	NORTH_TI	TRUE	PORT_HE	PORT_HE	NO	NO	YES	NO	YES	NO	NONE	NO	0
4	3	COHEN	TRUE	KARRATH	DAMPIER	NO	YES	YES	NO	YES	NO	NONE	NO	0
5	4	KEAST	TRUE	KARRATH	DAMPIER	NO	YES	YES	NO	YES	NO	NONE	NO	0
6	5	COLLIER_	TRUE	KARRATH	DAMPIER	NO	YES	NO	NO	YES	YES	NONE	NO	0
7	6	COLLIER_	TRUE	KARRATH	DAMPIER	NO	YES	NO	NO	YES	YES	NONE	NO	0
8	7	NORTH_G	TRUE	KARRATH	DAMPIER	NO	YES	YES	NO	YES	NO	NONE	NO	0

	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB
1	DISTURBA	INFRASTR	PERIMETE	SIZE	ML_DISTA	ML_DIST	WINDSPE	FLOODEV	FLOODCA	DESIRABI	VISITORT	DESIRABI	VISITORTYPE	
2	1	1	2962.35	41.25968	40290.47	OFFSHOR	23.07569	normal(4,3	MODERAT	NOT	ONLY_BIG	[0.0, 0.0, 1	[0.0, 0.0, 0.0, 1.0]	
3	0	0	3346.28	65.98288	18456.42	OFFSHOR	21.69016	normal(4,3	MODERAT	NOT	ONLY_BIG	[0.0, 0.0, 1	[0.0, 0.0, 0.0, 1.0]	
4	1	0	1751.78	12.03298	15264.46	MIDSHOR	21.82434	normal(4,3	LOW	NOT	ONLY_ME	[0.0, 0.0, 1	[0.0, 0.0, 1.0, 0.0]	
5	0	0	4320.15	51.30595	13941.37	MIDSHOR	21.78923	normal(4,3	HIGH	NOT	ONLY_ME	[0.0, 0.0, 1	[0.0, 0.0, 1.0, 0.0]	
6	1	0	3537.18	36.01224	12508.51	MIDSHOR	21.78923	normal(4,3	MODERAT	NOT	ALL	[0.0, 0.0, 1	[1.0, 0.0, 0.0, 0.0]	
7	0	0	2355.23	28.05971	12425.33	MIDSHOR	21.78923	normal(4,3	LOW	NOT	ALL	[0.0, 0.0, 1	[1.0, 0.0, 0.0, 0.0]	
8	0	0	4855.82	67.35116	11824.39	MIDSHOR	21.78923	normal(4,3	HIGH	NOT	ONLY_ME	[0.0, 0.0, 1	[0.0, 0.0, 1.0, 0.0]	

Column A - ISLANDID: Enter an integer that acts as an identification number for each site. This number must be unique and consistently identify one specific site across any input spreadsheets with the column header *ISLANDID*.

Column B - NAME: Enter an alpha-numeric name for each site. Use an underscore rather than blank spaces. This name must be unique and consistently identify one specific site as the names are used as column headers in the 6 spreadsheets that refer to dispersal via ocean currents.

Column C - ACTIVE: Enter either TRUE or FALSE. This column acts as an on/off switch and allows the user to limit the number of sites that will be run through the *Biosecurity BBN Software* code. The code will calculate the biosecurity risk estimates for fewer sites faster. TRUE=ON; FALSE=OFF.

Column D – ISLAND_GROUP: Data in this column must match column headers in the spreadsheet *visitation.csv*. Sites should be grouped according to the number and distribution of visitors (see example (Lohr, et al., 2017)).

Column E - REGION: Data in this column indicate the nearest source of NIS that may disperse via non-human mediated pathways. Data in this column must match row headings in *region.csv*, *regionCurrentCarryRate1.csv*, *regionCurrentCarryRate3.csv*, *regionCurrentCarryRate7.csv*, *regionFavourableWind.csv*, *regionBridge.csv*, and *regionThreatDensity.csv*.

Column F - ACCOM: This column contains binary data (YES/NO) indicating whether a site has permanent accommodation for visitors.

Column G - CORAL: This column contains binary data (YES/NO) indicating whether a site has nearby coral reefs that may be attractive for tourists.

Column H - BEACH: This column contains binary data (YES/NO) indicating whether a site has a beach that may be attractive for tourists.

Column I - CULTURAL: This column contains binary data (YES/NO) indicating whether a site has cultural sites that may be attractive for tourists. We included nearshore shipwrecks, European historical sites, and indigenous rock art as cultural sites that may be attractive for tourists.

Column J - FISHING: This column contains binary data (YES/NO) indicating whether a site has nearby fishing grounds that may be attractive for tourists.

Column K - SHELTERED: This column contains binary data (YES/NO) indicating whether a site has sheltered boat landing or camping sites that may be attractive to tourists.

Column L - INDUSTRYPHASE: This column contains data categorising the type of industrial activity that is occurring on a site. Data must match categories listed in the spreadsheet *industryCarry.csv* under the column INDUSTRYPHASE. In our example we have five types of industrial activity with markedly different levels of human activity, movement of pre-fabricated machinery, and quarantine: active; construction; decommissioned; lighthouse; and none. Entering 'none' for all sites will effectively turn off the industrial dispersal pathway.

Column M - MACHUSE: This column contains data indicating whether vehicles or large machinery may move onto a site. In our example, some industrial sites are connected to mainland regions via a causeway and hence large vehicles may move onto a site daily. Whereas other offshore industrial sites have limited amounts of machinery move onto a site once construction is complete.

Column N - INDUSTRYVISITATION: Enter an integer enumerating the number of visits made to a site by industry employees per annum. We multiplied the number of employees by an estimated number of shift changes that cause employees to move on and off a site per annum.

Column O - DISTURBANCE: This column contains binary data (1/0) indicating whether a site has disturbed ground that may be more vulnerable to invasion by weeds. We included the presence of digging fauna as a source of disturbed ground.

Column P - INFRASTRUCTURE: This column contains binary data (1/0) indicating whether a site has human infrastructure (including but not necessarily accommodation) that may facilitate the establishment of some NIS.

Column Q - PERIMETER: Enter a real number describing the perimeter of the site.

Column R - SIZE: Enter a real number describing the area of a site in hectares.

Column S – ML_DISTANCE: Enter a real number describing the distance (metres) between a site and the nearest source region of NIS.

Column T – ML_DIST: Enter a category describing the distance between a site and the mainland. Data in this column must match column headings in the GeNIe node VISITOR TYPE (Figure 4).

Column U - WINDSPEED: Enter a real number describing the average wind speed passing over a site. An if-else statement within the *Wind* sub-model, node *Speed Category*, will automatically assign a wind speed category to each site.

Column V - FLOODEVENTS: Enter an integer or a probability distribution describing the frequency of flood causing rainfall events in the source regions. We entered 'normal(4,3)', indicating that there is a normal distribution with an average of 4 flood causing rainfall events per year and a standard deviation of 3. The *Biosecurity BBN Software* runs 1000 iterations of the model and will randomly select numbers for the *Flood Plume Events* node within the *FloodCurrent* sub-model from within this normal distribution.

Column W - FLOODCATEGORY: This column contains categorical data describing the likelihood that flood waters will reach the perimeter of each site. Categories must match data in column 2 of *floodCurrentCarry.csv*.

Column X - DESIRABILITYMOSTLIKELY: The *Biosecurity BBN Software* will automatically enter data into the column at the end of each run. Pre-existing data will be over-written.

Column Y - VISITORTYPEMOSTLIKELY: The *Biosecurity BBN Software* will automatically enter data into the column at the end of each run. Pre-existing data will be over-written.

Column Z - DESIRABILITY: The *Biosecurity BBN Software* will automatically enter data into the column at the end of each run. Pre-existing data will be over-written.

Column AA - VISITORTYPE: The *Biosecurity BBN Software* will automatically enter data into the column at the end of each run. Pre-existing data will be over-written.

visitation.csv

	A	B	C	D	E	F	G	H	I	J	K	L
1	VISITATIONID	DESIRABILITY	VISITORTYPE	EXMOUTH	KARRATHA	ONSLOW	PORT_HEDLAND	MONTEBELLO	PORT_WALCOTT_OR_JOHNS_CREEK			
2	0	HIGH	ALL	300	3000	800	120	400	120			
3	1	HIGH	ONLY_SMALL	108.7	1039.22	277.1	41.6	1.1	41.6			
4	2	HIGH	ONLY_MED_BIG	181.2	1764.71	470.6	70.6	99.7	70.6			
5	3	HIGH	ONLY_BIG	10.1	196.08	52.3	7.8	299.2	7.8			
6	4	MODERATE	ALL	60	500	60	24	150	24			
7	5	MODERATE	ONLY_SMALL	20.8	173.2	20.8	8.3	0.8	8.3			
8	6	MODERATE	ONLY_MED_BIG	35.3	294.1	35.3	14.1	37.8	14.1			
9	7	MODERATE	ONLY_BIG	3.9	32.7	3.9	1.6	111.3	1.6			
10	8	NOT	ALL	1	1	1	1	30	1			
11	9	NOT	ONLY_SMALL	0.3	0.3	0.3	0.3	0.2	0.3			
12	10	NOT	ONLY_MED_BIG	0.6	0.6	0.6	0.6	7.6	0.6			
13	11	NOT	ONLY_BIG	0.1	0.1	0.1	0.1	22.3	0.1			

Column A - VISITATIONID: Enter a unique integer for a particular combination of visitation criteria.

Column B - DESIRABILITY: Enter categorical data describing level of site desirability for visitors. Data must match row headings available in *Desirability* sub-model, node *Desirability* (Figure 3). We have used HIGH, MODERATE, and NOT.

Column C - VISITORTYPE: Enter categorical data describing the type of visitors or vehicle that may access each site. Data must match row headings available in *Desirability* sub-model, node *Visitor Type* (Figure 4). We have used ALL, ONLY_SMALL, ONLY_MED_BIG, and ONLY_BIG.

Column D-onwards: Enter one column heading per ISLAND_GROUP listed in the *island.csv*. In the cells below enter the estimated number of visitors per annum to sites within corresponding groups identified by column heading and categories in columns B and C. For example, highly desirable islands that may be accessed by any recreational craft in the Karratha group (e.g., West Lewis Island) are thought to have approximately 3000 visitors per annum. A probability distribution may be entered here and into similar cells in other spreadsheets.

Recreation Fauna

The results of the island *Desirability* sub-model feed into the *Island Visitation* node of the *Recreation Fauna* sub-model (Figure 5). The estimated number of visitors per annum is combined with a calculated probability that any one of those visitors may be dispersing fauna NIS. The probability that visitors may be dispersing fauna NIS depends on the presence of NIS at the visitor's home-town or travelling start point; the propensity visitors have to carry NIS; and the propensity visitors have to drop a threat at the site they are visiting. Two data input spreadsheets contain data necessary for this sub-model: present.csv and threat.csv. The hypothesis underlying the Recreation Fauna sub-model is that fauna NIS will hitch a ride on visitor's vehicles or boats (White & Shine, 2009) or in visitor's luggage (Koch, et al., 2013). Once NIS has hitched a ride it may or may not exit the visitor's vehicle or luggage at the site or island. If a threat is already established on a site it will continue to be present on the site regardless of the number of newly arriving NIS.

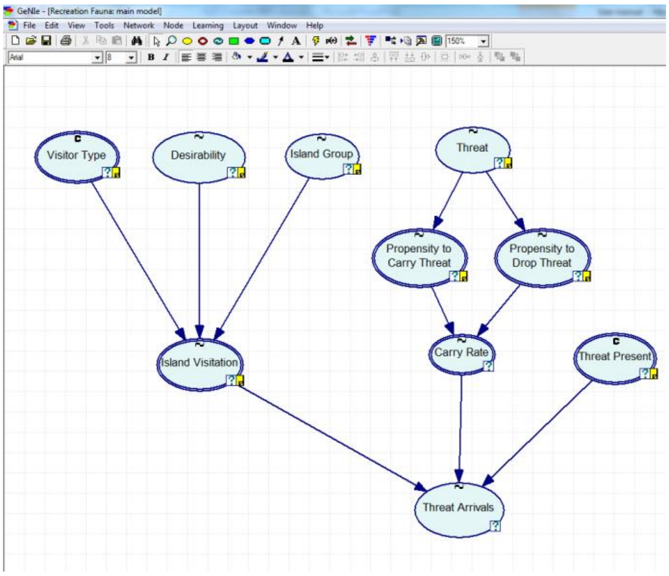


Figure 5: Screen-capture of the *Recreation Fauna* sub-model as viewed in GeNIe.

present.csv

	A	B	C	D	E	F	G	H
1	ISLANDREGIONID	BELLY_ACHE_BUSH	BUFFEL	CENCHRUS_SP	CROWNBEARD	RUBY_DOCK	INDIGOFERA	KAPOK
2	ML	1	1	0	0	0	1	1
3	EXMOUTH	0	1	0	0	0	1	0
4	ONSLOW	0	1	0	0	0	1	1
5	DAMPIER	1	1	0	0	0	1	1
6	PORT_WALCOTT	1	1	0	0	0	1	1
7	PORT_HEDLAND	0	1	0	0	0	1	0
8	BEDOUT	0	1	0	0	0	0	1

Column A –ISLANDREGIONID: This column contains the unique alpha-numeric labels used to name sites. Data must match *island.csv*, column B – NAME or *region.csv*, column B – NAME.
Columns B onwards: Column headings must match *threat.csv*, column B – NAME. There must be one column per threat listed in *threat.csv*.
Cells contain binary data (1/0) indicating whether each NIS is already present on each source region or each site.

threat.csv

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
1	THREATID	NAME	ACTIVE	TYPE	DISPERSA	WINDDISF	FLORAAT	FLORAAV	FAUNACA	DROPRAT	CURRENT	FLOODCU	CURRENT	WINDPRO	TEMP	ESTABLIS	BRIDGEUS	SWIMAVE	SWIMLON	SWIMAVE	SWIMLON	SWIMMAX	RAFTCRE	RAFTBOARDRATE	
2	0	BELLY_AC	FALSE	FLORA	NONE	LOW	0.55	0	0	50	0	0	0	0	0	1	0	0	0	0	0	0	0	0	
3	1	BUFFEL	FALSE	FLORA	WIND	MODERAT	12.5	0	0	50	0	0	0	0.001	0	1	0	0	0	0	0	0	0	0	
4	2	CENCHRU	FALSE	FLORA	WIND	LOW	12	0	0	50	0	0	0	0.001	0	1	0	0	0	0	0	0	0	0	
5	3	CROWNBI	FALSE	FLORA	WIND	MODERAT	6	0.1	0	50	0	0	0	0.001	0	1	0	0	0	0	0	0	0	0	
6	4	INDIGOFE	FALSE	FLORA	NONE	LOW	0	0	0	50	0	0	0	0	0	1	0	0	0	0	0	0	0	0	
7	5	KAPOK	FALSE	FLORA	WIND	HIGHLY	6.5	0	0	50	0	0	0	0.001	0	1	0	0	0	0	0	0	0	0	
8	6	MESQUITE	FALSE	FLORA	CURRENT	LOW	0	1	0	50	4	0.001	0.001	0	0	1	0	0	0	0	0	0	0	0	
9	7	MOTHER_	FALSE	FLORA	NONE	LOW	0	0	0	50	0	0	0	0	0	1	0	0	0	0	0	0	0	0	

Column A - THREATID: Enter an integer that acts as an identification number for each NIS. This number must be unique and consistently identify one specific NIS across any input spreadsheets with the column header *THREATID*.

Column B - NAME: Enter an alpha-numeric name for each NIS. Use an underscore rather than blank spaces. This name must be unique and consistently identify one specific NIS as the names are used as column headers in the spreadsheets *present.csv*, *predcomp.csv*, *islandThreatDensity.csv*, *regionThreatDensity.csv*, *industryCarry.csv*, *raftCarry.csv*, *habitat.csv*, *floodCurrentCarry.csv*, and *establishmentRate.csv*.

Column C - ACTIVE: Enter either TRUE or FALSE. This column acts as an on/off switch and allows the user to limit the number of NIS that will be run through the *Biosecurity BBN Software* code. The code will calculate the biosecurity risk estimates for fewer NIS faster. TRUE=ON; FALSE=OFF.

Column D - TYPE: Enter either FLORA or FAUNA. This data identifies which recreation dispersal sub-model should be used to estimate biosecurity risk to sites for each NIS.

Column E - DISPERSALMODE: For flora NIS, identify the types of natural dispersal each NIS may use: NONE; WIND; or CURRENT. Enter NONE for fauna NIS.

Column F - WINDDISPERSABILITY: Enter categorical data describing the flora NIS ability to disperse via wind. Data must match categories listed in *windCarry.csv*, column B – SeedType. Flora NIS diaspore morphology may be used to estimate the wind dispersal potential of seeds under average wind conditions (Hintze, et al., 2013; Tackenberg, et al., 2003).

Column G - FLORAATTACHMENTRATE: Enter a real number describing the percentage of visitors that a flora NIS diaspore may attached to.

Column H - FLORAAVOIDANCE: Enter a real number describing the probability (values 0 to 1) that visitors will avoid walking through a patch of flora NIS. For example, for prickly pear (*Opuntia sp.*) we entered 1 as people will avoid walking through a patch of cacti.

Column I - FAUNACARRYRATE: Enter a percent value (0 to 100) describing the percentage of visitors that may be carrying fauna NIS. A probability distribution may be entered instead of a single value (e.g., 'normal(12.5,9)' will cause the software to randomly select a number from a normal distribution with an average of 12.5% and a standard deviation of 9%)

Column J - DROPRATE: Enter a percent value (0 to 100) describing the percentage of visitors that may drop fauna or flora NIS they are carrying on a site. A probability distribution may be entered instead of a single value.

Column K – CURRENTCARRYDAYS: Enter an integer describing the maximum number of days a flora NIS diaspore may survive in seawater. The tool uses this number to determine which of the *islandCurrentCarryRate.csv files to use, in a similar fashion to the TYPE variable*. Possible values are 1, 3, or 7 (days), corresponding to the input spreadsheets: *islandCurrentCarryRate1.csv*, *islandCurrentCarryRate3.csv*, and *islandCurrentCarryRate7.csv*.

Column L - FLOODCURRENTPROPAGULEPRESSURE: Enter a real number describing the number of flora NIS propagules, per a geographic unit, that may disperse out of a population via flood waters.

Column M - CURRENTPROPAGULEPRESSURE: Enter a real number describing the number of flora NIS propagules, per a geographic unit, that may disperse out of a population via ocean currents.

Column N - WINDPROPAGULEPRESSURE: Enter a real number describing the number of flora NIS propagules, per a geographic unit, that may disperse out of a population via wind currents.

Column O - TEMP: Enter zero in all cells. This column is required for the software to calculate dispersal via the wind sub-model.

Column P - ESTABLISHMENTSIZE: Enter an integer describing the minimum number of individuals for each NIS required for a new population to establish. Self-fertilizing or parthenogenic species may only require 1 individual to establish a new population.

Column Q - BRIDGEUSE: Enter a real number describing the probability (values 0 to 1) that fauna NIS may cross a land bridge when it is exposed. Land bridges may be permanent (e.g., causeway or man-made bridge) or temporary tidal land bridge. The probability that a land bridge exists is listed in *regionBridge.csv* or *islandBridge.csv*.

Column R - SWIMAVERAGEDIST: Enter an integer describing the distance (metres) that a high proportion of individuals within a fauna NIS could reliably swim or fly. The number acts as a lower bound for distance that some individuals within a fauna NIS may be able to swim or fly. See example Table 1.

Column S - SWIMLONGDIST: Enter an integer describing the distance (metres) that matches the maximum distance most individuals within a fauna NIS will be able to swim or fly. Few individuals should be willing or able to exceed this distance. See example Table 1.

Column T - SWIMAVERAGERATE: Enter a real number describing the probability (values 0 to 1) that fauna NIS may survive swimming or flying a distance less than or equal to the distance listed in SWIMAVERAGEDIST. See example Table 1.

Column U - SWIMLONGRATE: Enter a real number describing the probability (values 0 to 1) that fauna NIS may survive swimming or flying a distance between the values listed in SWIMAVERAGEDIST and SWIMLONGDIST. See example Table 1.

Column V - SWIMMAXRATE: Enter a real number describing the probability (values 0 to 1) that fauna NIS may survive swimming or flying further than the distance listed in SWIMLONGDIST. See example Table 1.

Column W - RAFTCREATIONRATE: Enter a real number describing the probability (values 0 to 1) that a suitably buoyant raft would be created for each fauna NIS listed. Some species may raft on a fresh water lens, others require buoyant debris, whereas some may create their own raft out of conspecifics (Adams, et al., 2011).

Column X - RAFTBOARDRATE: Enter a real number describing the probability (values 0 to 1) that fauna NIS will board a raft when present. Consider whether or not a fauna NIS lives near a floodway and is likely to be caught in flood waters.

Table 1: *Rattus rattus* as an example for calculating fauna NIS dispersal via swimming or flying across a hostile matrix: 72% (0.72 probability) of rats will survive swimming ≤ 250m; 40% (0.4 probability) of rats will survive swimming ≥250m but ≤ 500m; and 3% (0.03 probability) of rats will survive swimming > 500m. Probabilities were derived from (Tabak, et al., 2015) and distance categories for *R. rattus* were derived from (Russell, et al., 2008).

NAME	SWIMAVERAGEDIST	SWIMLONGDIST	SWIMAVERAGERATE	SWIMLONGRATE	SWIMMAXRATE
RATTUS_RATTUS	250	500	0.72	0.4	0.03

Recreation Flora

The results of the island *Desirability* sub-model also feed into the *Island Visitation* node of the *Recreation Flora* sub-model (Figure 6). Unlike fauna which may actively seek to hitch a ride with visitors, flora NIS may passively attach to visitors vehicles, luggage or clothing when visitors pass through an NIS population (Pickering & Mount, 2010). The passive attachment of flora diaspores to visitors first requires that visitors interact with flora NIS prior to or at their travelling start point. Users of the *Biosecurity BBN Software* may enter values for the propensity for visitors to avoid flora NIS (e.g. tourists might avoid picking up cacti pads) and the propensity for flora NIS to attach diaspores to visitors alongside values describing the propensity for visitors to drop diaspores on the visited site in the *threat.csv* spreadsheet. Two additional data input spreadsheets are required to describe the quantity of flora NIS that may be producing propagules that could attach to recreational visitors; *region.csv*, and *regionThreatDensity.csv*.

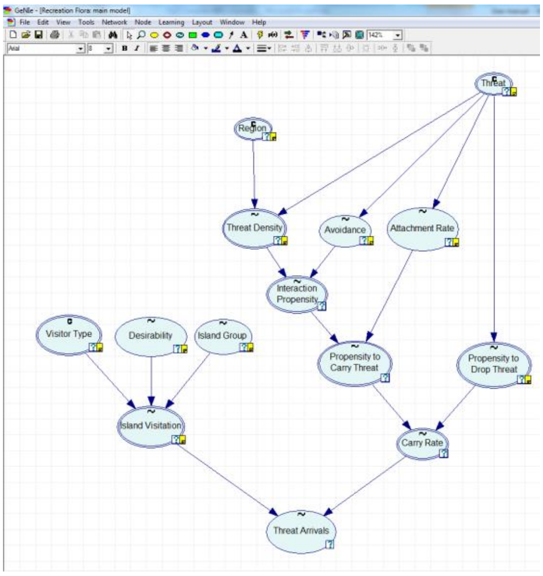


Figure 6: Screen-capture of the *Recreation Flora* sub-model as viewed in GeNIe.

region.csv

	A	B	C	D	E	F
1	REGIONID	NAME	PERIMETER	SIZE	WINDDIR	WINDSPEED
2	0	ML	2814668.135	220674.1648	50	22
3	1	EXMOUTH	542087.05	24157.03548	50	22
4	2	ONSLOW	649815.995	53887.47605	50	20
5	3	DAMPIER	275696.47	64847.55338	50	16
6	4	PORT_WALC(	601972.03	10695.41969	50	24
7	5	PORT_HEDLA	745096.59	67086.68017	50	17

Column A – REGIONID: Enter an integer that acts as an identification number for each NIS source region. This number must be unique and consistently identify one specific region across any input spreadsheets with the column header *REGIONID* (e.g., *regionThreatDensity.csv*).
Column B – NAME: Enter an alpha-numeric name for each NIS source region. Use an underscore rather than blank spaces. This name must be unique and consistently identify one specific region as the names are used as column headers in other spreadsheets. The row identified as ML refers to the mainland as a whole and is used in the floodCurrent sub-model as flood waters originate on the entireity of the mainland rather than discrete regions.
Column C – PERIMETER: Enter a real number describing the length (metres) of perimeter of the region. In our case study, all of our NIS source regions were a 1 km wide band of mainland coastline (Lohr, et al., 2017). Propagules will only disperse out to islands from the coastal side of the region; therefore we entered 0.5*perimeter of each region.
Column D – SIZE: Enter a real number describing the area (hectares) of the NIS source region.
Column E – WINDDIR: The proportion of time the wind direction was towards the islands. 50 was set as the default in the absence of more specific information.
Column F – WINDSPEED: average annual wind speed (kmph) at the region.

regionThreatDensity.csv

	A	B	C	D	E	F	G
1	REGIONID	BELLY_ACHE_BUSH	BUFFEL	CENCHRUS_S	CROWNBEARD	INDIGOFERA	KAPOK
2	ML	2.8	28	0	0	0	28
3	EXMOUTH	0	28	0	0	0	28
4	ONSLOW	0	28	0	0	0	28
5	DAMPIER	7	28	0	0	0	28
6	PORT_WALCOTT	7	28	0	0	0	28
7	PORT_HEDLAND	0	28	0	0	0	28

Column A – REGIONID: Enter an integer that acts as an identification number for each NIS source region. This number must be unique and consistently identify one specific region across any input spreadsheets with the column header *REGIONID* (e.g., *region.csv*).
Column B onwards: Column headings contain the name of each NIS and must match the data in *threat.csv*, column B – NAME.
Cells: Enter a real number describing the percentage of each region covered with each flora NIS. This number will be multiplied by data in *region.csv*, column D – SIZE to estimate the total size of the flora NIS population in each region. For fauna NIS enter the number of individuals (may be a real number or probability distribution) expected to disperse out of each region per annum. We estimated that approximately 1% of the fauna NIS population would disperse out of each region per annum.

Human-mediated dispersal by industrial activities

Industry is subject to far stricter quarantine protocols than recreational island users. Therefore, a separate sub-model is used to describe the probability that NIS may disperse via industrial activities (Figure 7). Unlike the recreational visitors sub-model, users of the *Biosecurity BBN Software* are required to define the number of visitors to industrial sites in the *island.csv* spreadsheet. In our case study of the Pilbara Islands we multiplied the number of employees working at a given site by an expected number of shift changes per year (Lohr, et al., 2017). Industrial sites are divided into categories defining the *Industrial Phase* or level of activity at a site, and whether or not large pieces of pre-fabricated machinery that may hide NIS from quarantine inspectors are moving onto the site (see *island.csv*). A propensity for visitors to industrial sites to carry NIS (Scott, et al., 2017) may then be defined by users of the *Biosecurity BBN Software* in the *industryCarry.csv.*

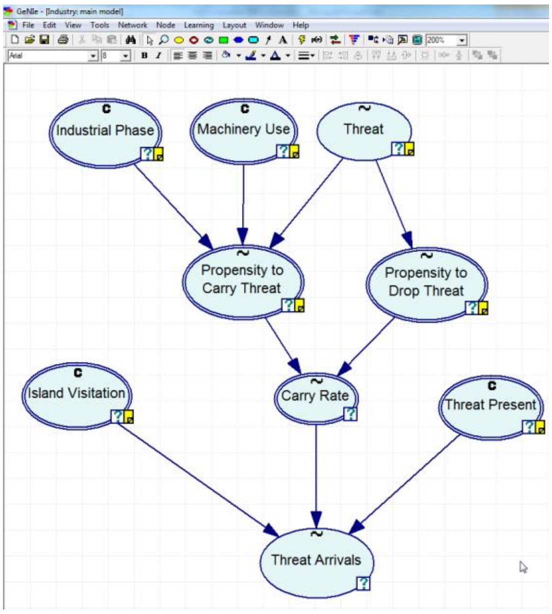


Figure 7: Screen-capture of the *Industry* sub-model as viewed in GeNIe.

islandCarry.csv

	A	B	C	D	E	F	G	H	I	J	K
1	INDUSTRYCARRYID	INDUSTRYPHASE	MACHINEUSE	BUFFEL	CENCHRUS_	CROWNBEARD	KAPOK	HEMIDACTYLUS_	PHEIDOLE_	RATTUS_RATTUS	
2	0	ACTIVE	YES	15	15	12.8	10.5	Normal(0.34,0.14)	Uniform(0,0)	Normal(0.34,0.14)	
3	1	ACTIVE	NO	12.6	12.6	0	10.8	Normal(0.25,0.14)	Uniform(0,0)	Normal(0.25,0.14)	
4	2	CONSTRUCTION	YES	23.75	23.75	24	23	Normal(0.49,0.14)	Uniform(0,0)	Normal(0.49,0.14)	
5	3	CONSTRUCTION	NO	17	17	0	13.4	Normal(0.49,0.14)	Uniform(0,0)	Normal(0.49,0.14)	
6	4	DECOMMISSIONED	YES	6.25	6.25	11.4	4	Normal(0.25,0.14)	Uniform(0,0)	Normal(0.25,0.14)	
7	5	DECOMMISSIONED	NO	7.6	7.6	0	4.2	Normal(0.25,0.14)	Uniform(0,0)	Normal(0.25,0.14)	
8	6	LIGHTHOUSE	YES	0	0	0	0	Normal(0.16,0.14)	Uniform(0,0)	Normal(0.16,0.14)	
9	7	LIGHTHOUSE	NO	0	0	0	0	Normal(0.16,0.14)	Uniform(0,0)	Normal(0.16,0.14)	
10	8	NONE	YES	0	0	0	0	Uniform(0,0)	Uniform(0,0)	Uniform(0,0)	
11	9	NONE	NO	0	0	0	0	Uniform(0,0)	Uniform(0,0)	Uniform(0,0)	

Column A – INDUSTRYCARRYID: Enter an integer that acts as an identification number for each industry dispersal combination. This number must be unique and consistently identify one specific industry dispersal combination across any input spreadsheets with the column header *INDUSTRYCARRYID*.

Column B – INDUSTRYPHASE: This column contains data categorising the type of industrial activity that is occurring on a site. Data must match categories listed in the spreadsheet *island.csv* under the column INDUSTRYPHASE. In our example we have five types of industrial activity with markedly different levels of human activity, movement of pre-fabricated machinery, and quarantine: active; construction; decommissioned; lighthouse; and none.

Column C – MACHINEUSE: This column contains binary data indicating whether vehicles or large machinery may move onto a site. Data in this column is combined with *island.csv*, column M - MACHUSE.

Column D onwards: Column headings must contain the same data as threat.csv, column B – NAME. All NIS must be listed with one NIS per column.

Cells: Enter a real number, or probability distribution describing the percentage (values 0 to 100) of industry site visitors that may be carrying each NIS.

Dispersal by self-powered swimming or flying

Many NIS may disperse by swimming, flying or walking across an inhospitable matrix. The success of these individuals at crossing the matrix depends on propagule pressure, the capabilities of the species, and the distance individuals must cross (Figure 8). Users of the *Biosecurity BBN Software* define propagule pressure by entering values into three spreadsheets: *regionThreatDensity.csv*, and *present.csv,* and *islandThreatDensity.csv*. The capabilities of a NIS at crossing an inhospitable matrix are entered into *threat.csv*. The distance between NIS source regions and sites is entered into *distance.csv*.

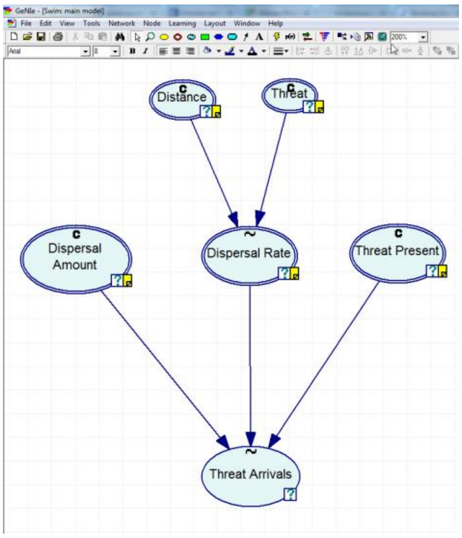


Figure 8: Screen-capture of the *Swim* sub-model as viewed in GeNIe.

islandThreatDensity.csv

	A	B	C	D	E	F	G	H	I	J
1	ISLANDID	BUFFEL	KAPOK	COLUMBA_I	FELIS_CA	HEMIDACTYLUS_	MUS_MUS	RHINELLA_	VULPES_	VULPES
2	TENT	15	1	0	0.02	0	0.5	0	0.01	
3	BURNSIDE	5	0	0.1	0	0	0.5	0	0.01	
4	SIMPSON	0	0	0.1	0	0	0.5	0	0	
5	WHALEBONE	0	0	0	0	0	0	0	0	
6	WHITMORE	10	0	0	0	0	0	0	0	
7	LITTLE_ROBERTS_SOUT	0	0	0	0	0	0	0	0	
8	ROBERTS	10	15	0.1	0	0	0	0	0	
9	ISLAND_93	0	0	0	0	0	0	0	0	
10	SOUTH	0	0	0	0	0	0	0	0	
11	DOOLE	20	5.4	0.1	0	0	0	0	0	
12	BEACON	0	0	0	0	0	0	0	0	

Column A – ISLANDID: Enter an alpha-numeric name that acts as an identifier for each site. This name must be unique and consistently identify one specific site across any input spreadsheets with the column header *ISLANDID* (e.g., *island.csv*).
Column B onwards: Column headings contain the name of each NIS and must match the data in *threat.csv*, column B – NAME.
Cells: Enter a real number describing the percentage of each region covered with each flora NIS. This number will be multiplied by data in *island.csv*, column R – SIZE to estimate the total size of the flora NIS population in each region. For fauna NIS enter the number of individuals (may be a real number or probability distribution) expected to disperse out of each site per annum. We estimated that approximately 1% of the fauna NIS population would disperse out of each site per annum.

distance.csv

	A	B	C	D	E	F
1	SOURCEID	BEDOUT	NORTH_TURTLE	COHEN	KEAST	COLLIER_ROCKS_5
2	BEDOUT	0	38490.30514	255372.1	252532.7	250475.6259
3	NORTH_TURTLE	38490.31	0	224850.6	221992.9	219536.9231
4	COHEN	255372.1	224850.642	0	2044.279	5214.548842
5	KEAST	252532.7	221992.875	2044.279	0	2643.459758
6	COLLIER_ROCKS_5	250475.6	219536.9231	5214.549	2643.46	0

Column A – SOURCEID: The column heading SOURCEID must be maintained. Rows below contain the same data as island.csv, column B – NAME.

Column B onwards: Column headings contain the same data as island.csv, column B – NAME.

Cells: Enter real numbers describing the distance between any two sites in metres.

Dispersal across temporary tidal-land bridges

In the Pilbara, some islands are connected to each other or to the mainland via temporary tidal land bridges. It is thought that animals rarely cross intertidal land bridges deliberately. Instead animals forage in the inter-tidal zone (Carlton & Hodder, 2003), or are trapped by large tidal movements and encouraged to disperse towards nearby islands by moving water. Regardless, if tidal land bridges or permanent causeways link sites to NIS source regions then users may need to estimate the probability that a land bridge exists and the probability that NIS may disperse across the bridges. The bridge sub-model (Figure 9) is designed to estimate dispersal through temporarily opened dispersal pathways. In addition to previously described data input spreadsheets, users will need to add data to *regionBridge.csv* and *islandBridge.csv*.

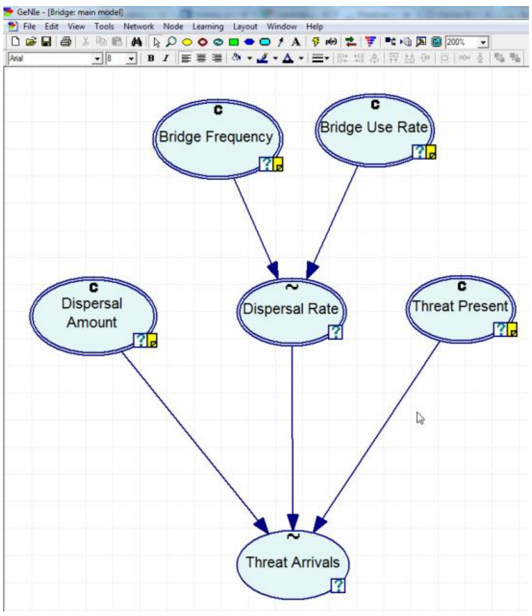


Figure 9: Screen-capture of the *Bridge* sub-model as viewed in GeNIe.

regionBridge.csv

	A	B	C	D	E	F	G	H
1	SOURCEID	CAREY	TENT	SIMPSON	DIXON	FINUCANE	FLY	HOPE
2	EXMOUTH	0	0.114486	0.018383	0	0	0	0.424078
3	ONSLOW	0.424078	0	0	0	0	0	0
4	DAMPIER	0	0	0	0	0	0	0
5	PORT_WALCOTT	0	0	0	0.424078	0	0	0
6	PORT_HEDLAND	0	0	0	0	1	0	0

Column A – SOURCEID: The column heading SOURCEID must be maintained. Rows below contain the same data as *region.csv*, column B – NAME, rows 1 onwards.

Column B onwards: Column headings contain the same data as *island.csv*, column B – NAME.

Cells: Enter real numbers describing the probability (values 0 to 1) that a land bridge exists. Value of 1 indicates a permanent land bridge or causeway.

islandBridge.csv

	A	B	C	D	E	F
1	SOURCEID	KEAST	COLLIER_5	COLLIER_6	NORTH_GIDLEY	MIDDLE_GIDLEY
2	KEAST	1	0	0	0	0
3	COLLIER_5	0	1	0.42407795	0	0
4	COLLIER_6	0	0.42407795	1	0	0
5	NORTH_GIDLEY	0	0	0	1	0.114486009
6	MIDDLE_GIDLEY	0	0	0	0.114486009	1
7	HAUY	0	0	0	0	0
8	DELAMBRE	0	0	0	0	0
9	GIDLEY	0	0	0	0	0.018382696

Column A – SOURCEID: The column heading SOURCEID must be maintained. Rows below contain the same data as *island.csv*, column B – NAME.

Column B onwards: Column headings contain the same data as *island.csv*, column B – NAME.

Cells: Enter real numbers describing the probability (values 0 to 1) that a land bridge exists between any two sites. Value of 1 indicates a permanent land bridge or causeway.

Dispersal via catastrophic flood waters

NIS may be dispersed during catastrophic flood waters as individuals or diaspores are caught up in moving water, or debris. The FloodCurrent sub-model (Figure 10) was designed for predicting the dispersal of flora NIS and hence uses estimates of region-SIZE (*region.csv*) multiplied by percent of landscape covered in flora NIS (*regionThreatDensity.csv*) to estimate the number of flora NIS propagules that may be dispersed by flood waters (Figure 10, node Dispersal amount). The frequency of flood events over NIS source regions is captured in *island.csv*, Column V – FLOODEVENTS. The categorical likelihood that flood waters will reach each site is captured in *island.csv*, Column W – FLOODCATEGORY. One additional data input spreadsheet is required for the FloodCurrent sub-model, *floodCurrentCarry.csv*.

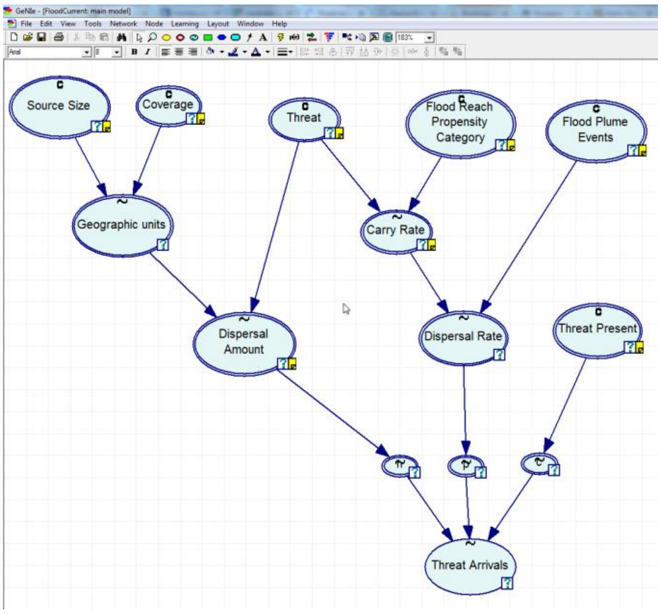


Figure 10: Screen-capture of the *FloodCurrent* sub-model as viewed in GeNIe.

floodCurrentCarry.csv

	A	B	C	D	E	F	G	H
1	FLOODCURRENTCARRYID	FLOODCATEGORY	BELLY_ACHE_BUSH	BUFFEL	CENCHRUS_	CROWNBI	KAPOK	MESQUITE
2	0	HIGH	0.03	0.2	0	0	0.2	0.1
3	1	MODERATE	0.005	0.1	0	0	0.1	0.05
4	2	LOW	0	0.05	0	0	0.05	0.01
5	3	NONE	0	0	0	0	0	0

Column A – FLOODCURRENTCARRYID: The column heading FLOODCURRENTCARRYID must be maintained. Rows below contain a unique numeric identifier.

Column B – FLOODCATEGORY: List the categories describing the likelihood that flood water will reach each site. Categories must be the same as *island.csv*, column W – FLOODCATEGORY.

Column C onwards: Column headings must contain the same data as *threat.csv*, column B – NAME.

Cells: Enter a real number describing the probability (values 0 to 1) that diaspores or individuals will survive being dispersed via flood waters without a raft (see Dispersal via catastrophic flood waters with raft) given the likelihood that flood waters will reach a site. Also consider the likelihood that NIS reside near flood ways.

Dispersal via catastrophic flood waters with raft

Most fauna NIS are going to require a raft to survive dispersal via flood waters and the number of individuals caught in floods may not relate to their density across the NIS source region as with flora NIS. Therefore, a separate Raft sub-model is used (Figure 11). The number of fauna NIS that may be available to disperse via rafting is entered into *regionThreatDensity.csv*. The frequency of flood events over NIS source regions is captured in *island.csv*, Column V – FLOODEVENTS. The categorical likelihood that flood waters will reach each site is captured in *island.csv*, Column W – FLOODCATEGORY. The probability that a suitably sized raft may be created during a flood event, and the probability that fauna NIS may board available rafts is captured in threat.csv, column W – RAFTCREATIONRATE, and column X – RAFTBOARDRATE. One additional data input spreadsheet is required for the *Raft* sub-model, *raftCarry.csv*. Ultimately, the probability that NIS disperse via a raft on flood waters is a combination of the probability that a flood will occur, by the probability a flood will reach a site, by the probability that a raft is created, by the probability that NIS board a raft, by the probability that NIS will survive riding a raft all the way to a site.

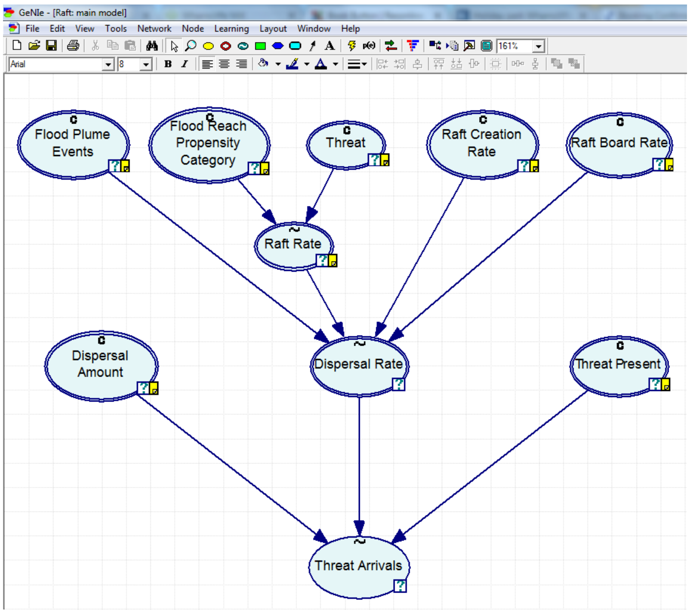


Figure 11: Screen-capture of the *Raft* sub-model as viewed in GeNIe.

raftCarry.csv

	A	B	C	D	E	F	G	H	I	J	K	L
1	RAFTCARRYID	FLOODCATEGORY	BUFFEL	KAPOK	BOS_	CANIS_	FELIS_	HEMIDACTYLUS_	RATTUS_	RHINELLA_	VULPES_VULPES	
2	0	HIGH	0.01	0.01	0.5	0.5	0.5	0.75	0.75	0.75	0.5	
3	1	MODERATE	0.005	0.005	0.1	0.1	0.1	0.3	0.3	0.3	0.1	
4	2	LOW	0.001	0.001	0.01	0.01	0.01	0.05	0.05	0.05	0.01	
5	3	NONE	0	0	0	0	0	0	0	0	0	

Column A – RAFTCARRYID: The column heading RAFTCARRYID must be maintained. Rows below contain a unique numeric identifier.

Column B – FLOODCATEGORY: List the categories describing the likelihood that flood water will reach each site. Categories must be the same as *island.csv*, column W – FLOODCATEGORY.

Column C onwards: Column headings must contain the same data as *threat.csv*, column B – NAME.

Cells: Enter a real number describing the probability (values 0 to 1) that diaspores or individuals will survive being dispersed via flood waters with a raft given the likelihood that flood waters will reach a site.

Dispersal via normal ocean currents

Some NIS, particularly flora NIS, may be dispersed by normal ocean currents if their diaspores fall into the ocean. In the Pilbara islands for example, prickly pear (*Opuntia stricta*) paddles are typically found establishing in the high-water mark, which suggests paddles are being dispersed by ocean currents. Data entered into *threat.csv*, column M – CURRENTPROPAGULEPRESSURE, *present.csv*, *regionThreatDensity.csv*, and *region.csv*, column C - PERIMETER are used to determine the number of diaspores that may enter and disperse via ocean currents. Data entered into *threat.csv*, column K – CURRENTCARRYDAYS refers to the ability for diaspores to survive in seawater which depends on several physiological traits including buoyancy, seed coatings, and dormancy (Guja, et al., 2010). The program uses the CURRENTCARRYDAYS to determine which currentcarry(rate) spreadsheet to use: *regionCurrentCarry[*CURRENTCARRYDAYS*].csv*, and *islandCurrentCarry[*CURRENTCARRYDAYS]*.csv.* These inputs specify the probability that ocean currents will disperse diaspores from one NIS source region or source site to another site within a set time frame.

Default files contain 6 ocean current data input spreadsheets: *regionCurrentCarry1.csv, regionCurrentCarry3.csv, regionCurrentCarry7.csv, islandCurrentCarry1.csv, islandCurrentCarry3.csv, islandCurrentCarry7.csv.* The Pilbara island project had access to high-resolution ocean current modelling for the Northwest/Pilbara coastline (http://www.csiro.au/connie/) that would allow us to estimate the dispersal of particles over variable time periods. Ocean current models are also available for other regions of the globe at (http://www.csiro.au/connie/).

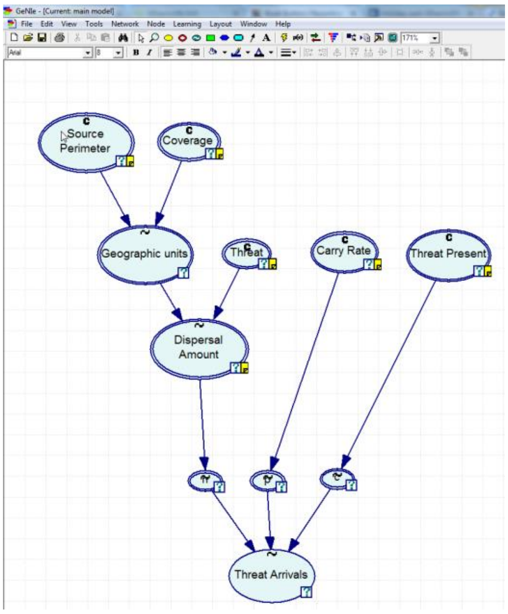


Figure 12: Screen-capture of the *Current* sub-model as viewed in GeNIe.

regionCurrentCarry.csv

	A	B	C	D	E	F	G
1	SOURCEID	WEST_LEWIS_SOUTH	COWIE	LOCKER	SANDY_	SANDY_F	WEST_LEWIS_NORTH
2	EXMOUTH	0	0	0	0	0	0
3	ONSLOW	0	0.5	0.5	0	0	0
4	DAMPIER	0.5	0	0	0.5	0	0
5	PORT_WALCOTT	0	0	0	0	0.5	0
6	PORT_HEDLAND	0	0	0	0	0	0

Column A – SOURCEID: The column heading SOURCEID must be maintained. Rows below contain the same data as *region.csv*, column B - NAME.

Column B onwards –Column headings must contain the same data as *island.csv*, column B – NAME.

Cells: Enter a real number describing the probability (values 0 to 1) that ocean currents will travel from a source regions coastline to each site within the time period specified in the file name.

islandCurrentCarry.csv

	A	B	C	D	E	F	G
1	SOURCEID	BEDOUT	COHEN	KEAST	COLLIER_ROCKS_5	COLLIER_ROCKS_6	MIDDLE_GIDLEY
2	BEDOUT	1	0	0	0	0	0
3	COHEN	0	1	0.014	0.014	0.007	0.028
4	KEAST	0	0.014	1	0.014	0.014	0.028
5	COLLIER_ROCKS_5	0	0.007	0.014	1	0.021	0.028
6	COLLIER_ROCKS_6	0	0.007	0.014	0.021	1	0.028
7	MIDDLE_GIDLEY	0	0.007	0.014	0.014	0.014	1
8	HAUY	0	0	0	0	0	0
9	DELAMBRE	0	0	0	0	0	0
10	GIDLEY	0	0.007	0.014	0.007	0.007	0.028

Column A – SOURCEID: The column heading SOURCEID must be maintained. Rows below contain the same data as *island.csv*, column B - NAME.

Column B onwards –Column headings must contain the same data as *island.csv*, column B – NAME.

Cells: Enter a real number describing the probability (values 0 to 1) that ocean currents will travel from one site to each other site within the time period specified in the file name.

Dispersal via wind

The wind dispersal sub-model was primarily designed for estimating the probability of dispersal of flora NIS diaspores. Hence, data entered into *threat.csv*, column F – WINDDISPERSABILITY, and column N - WINDPROPAGULEPRESSURE, *present.csv*, *regionThreatDensity.csv*, and *region.csv*, column D - SIZE are used to determine the number of diaspores (Dispersal amount) that may enter and disperse via wind currents (Figure 13). The spreadsheet region.csv also contains the columns WINDSPEED, which is used to determine if average wind speeds over the NIS source regions will be sufficient to raise diaspores into the air column and WINDDIR, which is used to determine the proportion of time (on an annual basis) that the wind direction is blowing toward each island from every island. Flora NIS diaspore morphology may be used to estimate the wind dispersal potential of seeds under average wind conditions (Hintze, et al., 2013; Tackenberg, et al., 2003). Three additional data input spreadsheets are used by the wind sub-model: *windCarry.csv*, *regionFavourableWind.csv*, and *islandFavourableWind.csv*.

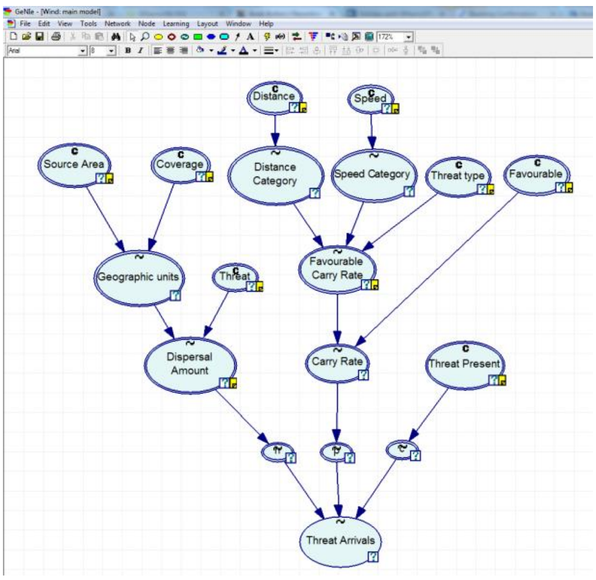


Figure 13: Screen-capture of the *Wind* sub-model as viewed in GeNIe.

windCarry.csv

	A	B	C	D	E
1	WINDCARRYRATEID	SeedType	Distance	Speed	CarryRate
2	0	HIGHLY	LONG	HIGH	0.00517
3	1	HIGHLY	LONG	MEDIUM	0.005193
4	2	HIGHLY	LONG	LOW	0.008817
5	3	HIGHLY	SHORT	HIGH	0.006581
6	4	HIGHLY	SHORT	MEDIUM	0.000387
7	5	HIGHLY	SHORT	LOW	0.002867
8	6	MODERATE	LONG	HIGH	0.001141
9	7	MODERATE	LONG	MEDIUM	0.001735
10	8	MODERATE	LONG	LOW	0.003872
11	9	MODERATE	SHORT	HIGH	0.006053
12	10	MODERATE	SHORT	MEDIUM	0.006751
13	11	MODERATE	SHORT	LOW	0.000337
14	12	LOW	LONG	HIGH	0.002588
15	13	LOW	LONG	MEDIUM	0.002034
16	14	LOW	LONG	LOW	0.00593
17	15	LOW	SHORT	HIGH	0.000912
18	16	LOW	SHORT	MEDIUM	0.007728
19	26	LOW	SHORT	LOW	0.008499

Column A – WINDCARRYID: The column heading WINDCARRYID must be maintained. Rows below contain a unique numeric identifier. All combinations of SeedType, Distance, and Speed must be entered and given a unique numeric identifier.

Column B – SeedType: This data corresponds to entries in threat.csv, column F – WINDDISPERSABILITY.

Column C – Distance: The distances between NIS source regions or source sites and other sites is placed into one of two categories, long or short distance. SHORT distance is ≤600m. LONG distance is >600m. The distance between NIS source regions and sites is listed in island.csv, column S – ML_DISTANCE. The distance between island source sites and other sites is listed in *distance.csv*.

Column D – Speed: Wind speed has been divided into three categories: HIGH, MEDIUM, and LOW. HIGH speed is ≥27kmph. MEDIUM speed is ≥18kmph. LOW speed is <18kmph. These values are coded into the GeNIe sub-model, Speed Category node.

Column E – CarryRate: Enter a real number describing the probability (values 0 to 1) that flora NIS diaspores will be carried by wind with the characteristics listed in columns B through D.

regionFavourableWind.csv

	A	B	C	D	E	F	G
1	SOURCEID	BEDOUT	NORTH_TI	COHEN	KEAST	COLLIER_	COLLIER_
2	EXMOUTH	0	0	0	0	0	0
3	ONSLOW	0	0	0	0	0	0
4	DAMPIER	0	0	0	0	0	0
5	PORT_WALCOTT	0	0	0	0	0	0
6	PORT_HEDLAND	0	0	16.66667	16.66667	16.66667	16.66667

Column A – SOURCEID: The column heading SOURCEID must be maintained. Rows below contain the same data as *region.csv*, column B - NAME.

Column B onwards –Column headings must contain the same data as *island.csv*, column B – NAME.

Cells: Enter a real number describing the proportion of time on an annual basis (values 0 to 1) that the wind will be travelling in the direction of each of the sites.

islandFavourableWind.csv

	A	B	C	D	E	F	G	H
1	SOURCEID	BEDOUT	NORTH_TI	COHEN	KEAST	COLLIER_	COLLIER_	NORTH_G
2	BEDOUT	0	8.333333	16.66667	25	25	25	25
3	NORTH_TURTLE	0	0	8.333333	8.333333	8.333333	8.333333	8.333333
4	COHEN	0	0	0	0	16.66667	25	25
5	KEAST	0	0	0	0	33.33333	41.66667	16.66667
6	COLLIER_ROCKS_5	0	0	0	0	0	25	8.333333
7	COLLIER_ROCKS_6	0	0	0	0	0	0	8.333333
8	NORTH_GIDLEY	0	0	0	0	0	0	0

Column A – SOURCEID: The column heading SOURCEID must be maintained. Rows below contain the same data as *island.csv*, column B - NAME.

Column B onwards –Column headings must contain the same data as *island.csv*, column B – NAME.

Cells: Enter a real number describing the probability (values 0 to 1) that the wind will be travelling from the site listed in column A to each of the sites listed in the headings of column B onwards. 0.5 was set as the default in the absence of more specific information.

Software outputs

The dispersal sub-models described above are linked by a purpose built tool in Java. Ultimately, the Java code connects nodes across the various sub-models, uses stochastic sampling (n = 1000) to run the sub-models with variable input parameters, and generate a distribution of values for the number of arrivals per NIS per site. Only parameters with probability distributions rather than static inputs will vary in the 1000 iterations of the *Biosecurity BBN software*. An output spreadsheet, *arrivals.csv* is automatically generated and saved in the WAB/DB folder. This spreadsheet will be overwritten by subsequent runs of the *Biosecurity BBN software* unless copies are saved elsewhere.

Estimated number of arrivals

The estimated number of arrivals spreadsheet lists results for each NIS in the *threat.csv* spreadsheet by each site listed in the *islands.csv* spreadsheet on a separate row. The 600 islands by 11 NIS in our case study (Lohr, et al., 2017) resulted in 6600 rows of data in the *arrivals.csv*. The number of arrivals travelling through each of the 8 dispersal sub-models is presented in 9 columns (2 columns for dispersal by recreational visitors). Each column contained comma-delimited data housed within square brackets. The first number within those brackets is the average estimated number of arrivals travelling through each pathway. Subsequent values are the results generated by the 1000 iterations of the *Biosecurity BBN Software*. Four columns (CURRENTSUMMARY, WINDSUMMARY, SWIMSUMMARY, and BRIDGESUMMARY) break down the average number of arrivals according to source location. In these four dispersal sub-model NIS could travel or island-hop from one site/island to another site/island. These columns give users of the *Biosecurity BBN software* access to island-hopping results. The final four columns give users access to the total estimated average number of arrivals, summed across the dispersal sub-models, the complete histogram of results generated by 1000 iterations of the BBN calculations, and the total estimated standard deviation for those results. The tabulated mean and standard deviation results may be used immediately or linked to a geographic information system (GIS) to create a pest risk map (Smith, et al., 2007).

arrivals.csv

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
1	ARRIVALSID	ISLANDID	THREATID	ISLANDNAME	THREATNAME	RECFLORA	RECFAUNA	INDUSTRY	RAFT	FLOODCURRENT	CURRENT	WIND	SWIM	BRIDGE	CURRENTSUMMARY	WINDSUMMARY	SWIMSUMMARY	BRIDGESUMMARY	TOTAL	TOTALHISTO	TOTALMEAN	TOTALSD
2	0	1	27	BEDOUT	VULPES_VULP	[0.0, 0.0]	[0.0, 0.0]	[0.0, 0.0]	[0.001000	[0.0, 0.0]	[0.0, 0.0]	[0.0, 0.	[0.221	[0.0, 0.0]	{}	{}	{TENT=0.01100000	{}	[0.2220000000000001,		0.222	1.27178
3	1	2	27	NORTH_TURTL	VULPES_VULP	[0.0, 0.0]	[0.0, 0.0]	[0.0, 0.0]	[0.005999	[0.0, 0.0]	[0.0, 0.0]	[0.0, 0.	[0.220	[0.0, 0.0]	{}	{}	{TENT=0.01199999	{}	[0.22699999999999998		0.227	1.338489
4	2	3	27	COHEN	VULPES_VULP	[0.0, 0.0]	[0.0, 0.0]	[0.0, 0.0]	[0.001000	[0.0, 0.0]	[0.0, 0.0]	[0.0, 0.	[0.193	[0.0, 0.0]	{}	{}	{TENT=0.01000000	{}	[0.19499999999999995		0.195	1.180808
5	3	4	27	KEAST	VULPES_VULP	[0.0, 0.0]	[0.0, 0.0]	[0.0, 0.0]	[0.015, 0.1	[0.0, 0.0]	[0.0, 0.0]	[0.0, 0.	[0.220	[0.0, 0.0]	{}	{}	{TENT=0.00500000	{}	[0.23500000000000004		0.235	1.330706
6	4	5	27	COLLIER_ROC	VULPES_VULP	[0.0, 0.0]	[0.0, 0.0]	[0.0, 0.0]	[0.002000	[0.0, 0.0]	[0.0, 0.0]	[0.0, 0.	[0.214	[0.0, 0.0]	{}	{}	{TENT=0.00800000	{}	[0.21699999999999997		0.217	1.222531
7	5	6	27	COLLIER_ROC	VULPES_VULP	[0.0, 0.0]	[0.0, 0.0]	[0.0, 0.0]	[0.0, 0.0]	[0.0, 0.0]	[0.0, 0.0]	[0.0, 0.	[0.192,	[0.0, 0.0]	{}	{}	{TENT=0.01299999	{}	[0.192, 1.17080551982		0.192	1.170806

Column A – ARRIVALSID: Unique numeric identifier for each NIS-Site combination.

Column B – ISLANDID: Unique numeric identifier that corresponds with *island.csv*, column A – ISLANDID.

Column C – THREATID: Unique numeric identifier that corresponds with *threat.csv*, column A – THREATID.

Column D – ISLANDNAME: Unique alpha-numeric identifier that corresponds with *island.csv*, column B – NAME.

Column E – THREATNAME: Unique alpha-numeric identifier that corresponds with *threat.csv*, column B – NAME.

Column F – RECFLORA: Contains the results of the dispersal of flora NIS via recreational visitors. First value is the average of 1000 iterations of the BBN calculations. Subsequent values are the result of an individual run of the BBN. [0.0, 0.0] = no dispersal through the sub-model or the sub-model was effectively deactivated.

Column G – RECFAUNA: Contains the results of the dispersal of fauna NIS via recreational visitors. First value is the average of 1000 iterations of the BBN calculations. Subsequent values are the result of an individual run of the BBN. [0.0, 0.0] = no dispersal through the sub-model or the sub-model was effectively deactivated for a given NIS-Site combination.

Column H – INDUSTRY: Contains the results of the dispersal of NIS via industrial visitors. First value is the average of 1000 iterations of the BBN calculations. Subsequent values are the result of an individual run of the BBN. [0.0, 0.0] = no dispersal through the sub-model or the sub-model was effectively deactivated for a given NIS-Site combination.

Column I – RAFT: Contains the results of the dispersal of fauna NIS via rafts travelling on flood waters. First value is the average of 1000 iterations of the BBN calculations. Subsequent values are the result of an individual run of the BBN. [0.0, 0.0] = no dispersal through the sub-model or the sub-model was effectively deactivated for a given NIS-Site combination.

Column J – FLOODCURRENT: Contains the results of the dispersal of flora NIS via flood waters (no raft). First value is the average of 1000 iterations of the BBN calculations. Subsequent values are the result of an individual run of the BBN. [0.0, 0.0] = no dispersal through the sub-model or the sub-model was effectively deactivated for a given NIS-Site combination.

Column K – CURRENT: Contains the results of the dispersal of flora NIS via normal ocean currents. First value is the average of 1000 iterations of the BBN calculations. Subsequent values are the result of an individual run of the BBN. [0.0, 0.0] = no dispersal through the sub-model or the sub-model was effectively deactivated for a given NIS-Site combination.

Column L – WIND: Contains the results of the dispersal of flora NIS via average wind currents. First value is the average of 1000 iterations of the BBN calculations. Subsequent values are the result of an individual run of the BBN. [0.0, 0.0] = no dispersal through the sub-model or the sub-model was effectively deactivated for a given NIS-Site combination.

Column M – SWIM: Contains the results of the dispersal of fauna NIS via self-powered swimming or flying across a hostile matrix. First value is the average of 1000 iterations of the BBN calculations. Subsequent values are the result of an individual run of the BBN. [0.0, 0.0] = no dispersal through the sub-model or the sub-model was effectively deactivated for a given NIS-Site combination.

Column N – BRIDGE: Contains the results of the dispersal of fauna NIS over land bridges that may be temporary. First value is the average of 1000 iterations of the BBN calculations. Subsequent values are the result of an individual run of the BBN. [0.0, 0.0] = no dispersal through the sub-model or the sub-model was effectively deactivated for a given NIS-Site combination.

Column O – CURRENTSUMMARY: Provides a breakdown of the number of arrivals travelling from each source via the normal ocean current sub-model. Data is present as ISLANDNAME=average arrival from this source. Data is comma-delimited.

Column P – WINDSUMMARY: Provides a breakdown of the number of arrivals travelling from each source via the average wind sub-model. Data is present as ISLANDNAME=average arrival from this source. Data is comma-delimited.

Column Q – SWIMSUMMARY: Provides a breakdown of the number of arrivals travelling from each source via the self-powered swimming/flying sub-model. Data is present as ISLANDNAME=average arrival from this source. Data is comma-delimited.

Column R – BRIDGESUMMARY: Provides a breakdown of the number of arrivals travelling from each source via the temporary land bridge sub-model. Data is present as ISLANDNAME=average arrival from this source. Data is comma-delimited.

Column S – TOTAL: Contains the results of the dispersal of NIS summed across all of the dispersal sub-models. First value is the average of 1000 iterations of the BBN calculations. Subsequent values are the result of an individual run of the BBN. [0.0, 0.0] = no dispersal through the sub-model or the sub-model was effectively deactivated for a given NIS-Site combination.

Column T – TOTALMEAN: Contains the average number of NIS arriving by all of the dispersal sub-models.

Column U – TOTALSD: Contains the standard deviation of the number of NIS arriving by all of the dispersal sub-models.

Probability of establishment

The estimated number of arrival for each of the dispersal pathways are aggregated for each threat and run through the establishment sub-model (via stochastic sampling, as above) to generate a probability of establishment for each NIS on each site (Figure 14). Data for the nodes Present, Establishment Size, and Threat Arrivals (Figure 14), are derived from the sub-models and data input spreadsheets described above. Three additional data input sheets are required to calculated the risk of new NIS populations becoming established on each site: *establishmentRate.csv*, *habitat.csv*, and *predcomp.csv*.

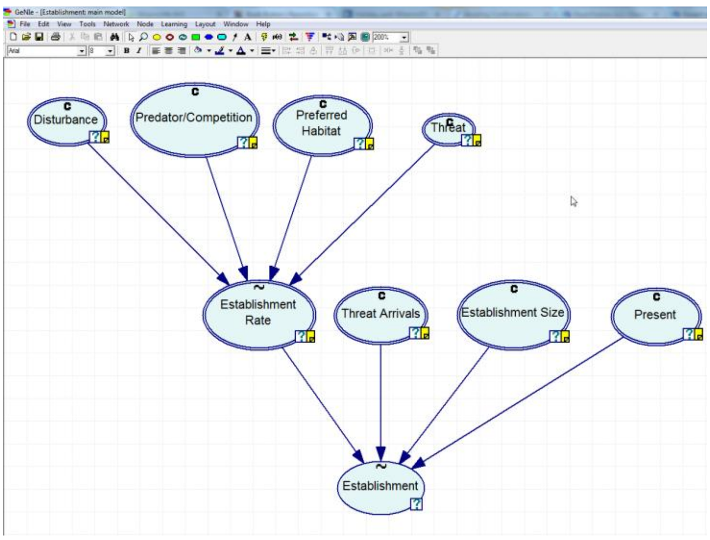


Figure 14: Screen-capture of the *Establishment* sub-model as viewed in GeNIe.

establishmentRate.csv

	A	B	C	D	E	F	G	H	I	J	K	L
1	ESTABLISHMENTRATEID	DISTURBANCE	PREDATORCOMPETITION	PREFERRED_HABITAT	BUFFEL	KAPOK	FELIS_(	HEMIDACTYLUS_	RATTUS_	RHINELLA_	VULPES_VULPES	
2	0	TRUE	TRUE	TRUE	50	50	0.25	0.4	0.5	0.95	0.5	
3	1	TRUE	TRUE	FALSE	0.01	0.01	0	0	0	0	0	
4	2	TRUE	FALSE	TRUE	50	50	0.95	0.87	0.9	0.95	0.95	
5	3	TRUE	FALSE	FALSE	0.01	0.01	0	0	0.1	0	0	
6	4	FALSE	TRUE	TRUE	20	10	0.25	0.4	0.5	0.95	0.5	
7	5	FALSE	TRUE	FALSE	0.01	0.01	0	0.05	0.1	0	0	
8	6	FALSE	FALSE	TRUE	20	10	0.95	0.87	0.9	0.95	0.95	
9	7	FALSE	FALSE	FALSE	0.01	0.01	0	0.1	0.1	0	0	

Column A – ESTABLISHMENTRATEID: The column heading ESTABLISHMENTRATEID must be maintained. Rows below contain a unique numeric identifier. All possible combinations of categories in columns B through D must be listed with a numeric identifier in column A.

Column B – DISTURBANCE: Binary data (TRUE/FALSE) describing whether significant habitat disturbance that may facilitate the establishment of weeds is present. We included the presence of digging fauna as a significant source of disturbance on islands (Lohr, et al., 2017).

Column C – PREDATORCOMPETITION: Binary data (TRUE/FALSE) describing whether there are any predators or competitors present on sites that may hinder the establishment of NIS (Palmer, et al., 2013; Kennedy, et al., 2002).

Column D – PREFERRED_HABITAT: Binary data (TRUE/FALSE) describing whether the habitat preferred or essential for each NIS species is present.

Column E onwards: Column headings must contain the same data as *threat.csv*, column B – NAME.

Cells: Enter a real number describing the probability (values 0 to 1) that each NIS will establish a new population on a site with the criteria listed in columns B through D.

habitat.csv

	A	B	C	D	E	F	G	H	I
1	ISLANDID	BUFFEL	CROWNBEARD	BOS_TAURUS	FELIS_CATUS	RATTUS_RATTUS	RHINELLA_	VULPES_VULPES	
2	BEDOUT	0	bernoulli(0.5)	0	1	1	1	1	
3	NORTH_TURTLE	1	bernoulli(0.5)	0	1	1	0	1	
4	COHEN	1	bernoulli(0.5)	0	0	1	0	0	
5	KEAST	1	bernoulli(0.5)	0	1	1	0	1	
6	COLLIER_ROCKS_5	1	bernoulli(0.5)	0	1	1	1	1	
7	COLLIER_ROCKS_6	1	bernoulli(0.5)	0	0	1	0	0	
8	NORTH_GIDLEY	1	bernoulli(0.5)	0	1	1	1	1	

Column A – ISLANDID: Unique alpha-numeric identifier that corresponds with *island.csv*, column B – NAME.

Column B onwards: Column headings must contain the same data as *threat.csv*, column B – NAME.

Cells: Enter a real number or probability distribution describing the probability (values 0 to 1) that the preferred habitat for each NIS is present on each site. Bernoulli(0.8) equates to 80% chance that the preferred habitat type is present (20% not).

predcomp.csv

	A	B	C	D	E	F	G	H	I	J	K
1	ISLANDID	BUFFEL	CROWNBEARD	BOS_	COLUMBA_	FELIS_	HEMIDACTYLUS_	RATTUS_	RHINELLA_	VULPES_VULPES	
2	BEDOUT	0	1	0	1	0	0	0	0	0	
3	NORTH_TURTLE	0	1	0	1	1	1	1	0	1	
4	COHEN	0	1	0	1	1	1	1	0	1	
5	KEAST	0	1	0	1	1	1	1	0	1	
6	COLLIER_ROCKS_5	0	1	0	1	1	1	1	0	1	
7	COLLIER_ROCKS_6	0	1	0	1	1	1	1	0	1	

Column A – ISLANDID: Unique alpha-numeric identifier that corresponds with *island.csv*, column B – NAME.

Column B onwards: Column headings must contain the same data as *threat.csv*, column B – NAME.

Cells: Enter a real number or probability distribution describing the probability (values 0 to 1) that predators or competitors are present on each site. Bernoulli(0.8) equates to 80% chance that predators or competitors are present (20% not).

establishment.csv

The establishment.csv is a results file saved in the folder WAB/DB. It will be overwritten by subsequent runs of the *Biosecurity BBN software* unless a copy is saved elsewhere.

	A	B	C	D	E	F	G	H	I	J	K
1	ISLANDREGIONID	BOS_	CANIS_	COLUMBA_	FELIS_(	HEMIDACTYLUS_	MUS_I	RATTUS_	RHINELLA_	VULPES_VULPES	
2	BEDOUT	0	0	0	0	0	0.002	0.02	0	0	
3	NORTH_TURTLE	0	0	0	0	0	0	0.004	0	0.001	
4	COHEN	0	0	0	0	0	0	0.017	0	0	
5	KEAST	0	0	0	0	0	0	0.016	0	0	
6	COLLIER_ROCKS_5	0	0	0	0	0	0.001	1	0	0	
7	COLLIER_ROCKS_6	0	0	0	0	0	0	1	0	0	

Column A – ISLANDID: Unique alpha-numeric identifier that corresponds with *island.csv*, column B – NAME.

Column B onwards: Column headings contain the same data as *threat.csv*, column B – NAME.

Cells: Results: Estimated probability (values 0 to 1) that each NIS will establish a new population on each site. If a NIS is already established on a site (see *present.csv*) then the estimated probability of establishment will be equal to 1.

Potential Applications and Extensions of the *Biosecurity BBN Software*

Multi-year dispersal projections via Bernoulli distribution

The Biosecurity BBN software estimates the annual number of NIS arrivals per site and the annual probability of establishment for each NIS on each site. Users can calculate multi-year 'island-hopping' dispersal projections by feeding results from *establishment.csv* back into the data input spreadsheet *present.csv* as a Bernoulli probability distribution. The Bernoulli distribution is a discrete distribution with two possible outcomes. Bernoulli(0.5) allows for a 50% chance NIS is present and 50% NIS is not present.

Trouble-shooting

Some common problems may occur when trying to run the Biosecurity BBN software on another machine. Problems experienced by the developers are listed below.

Problem: Java is not recognized as an internal or external command (Figure 15)
Solution: Close dialog box. Update Java. Try again.

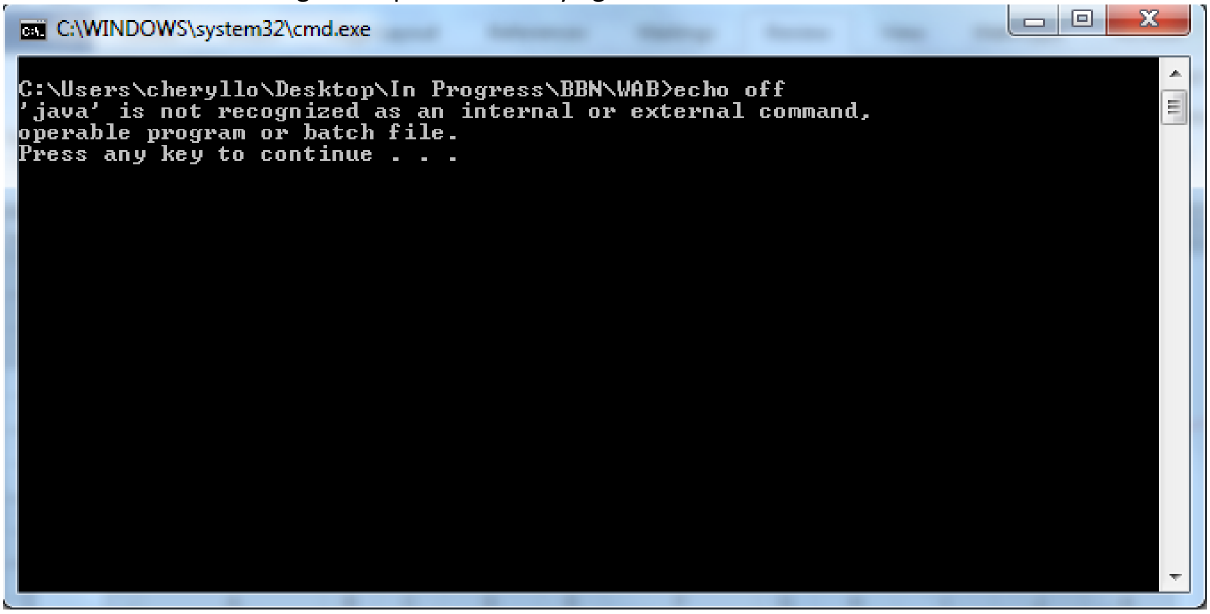


Figure 15: The error message received when you need to update your version of Java.

Problem: SQL exception
Solution: Open all of the data input spreadsheets associated with the *Industry* sub-model. Look for a typing mistake in the category headings. All associated categories and column headings must match.

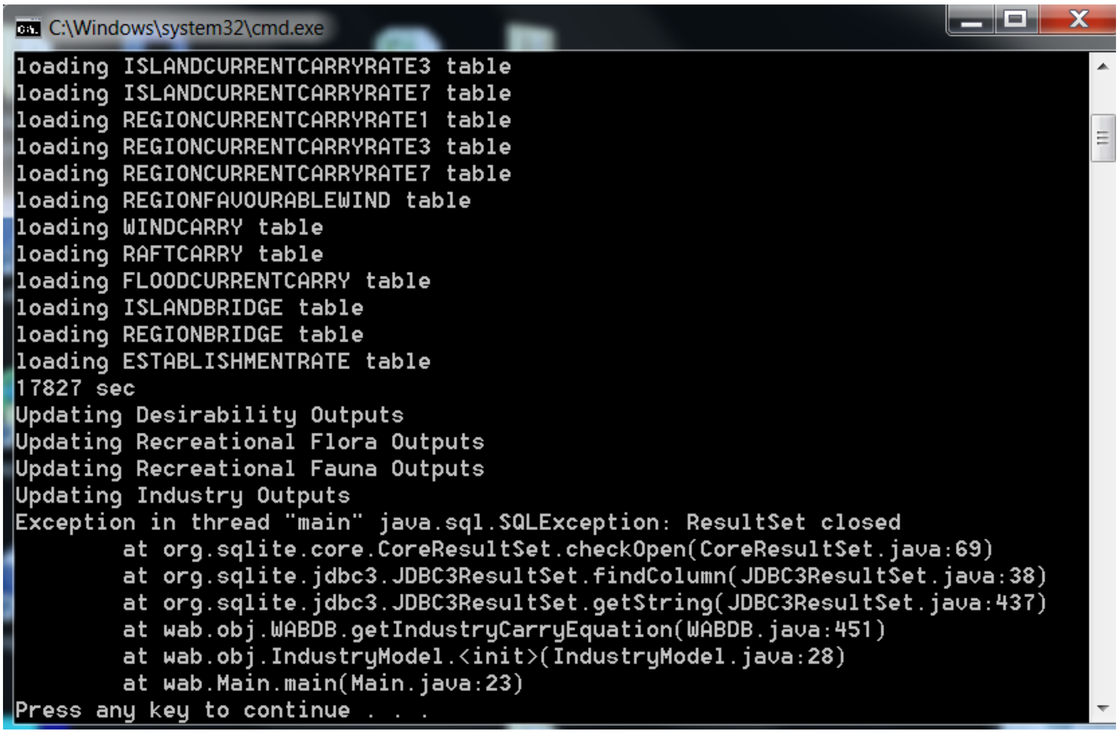


Figure 16: Typical error message when there is a typo in one of your files.

Glossary

Probability distributions that may be entered into data input spreadsheets may be found by opening a sub-model in GeNIe, double click on a node, tab Definition, under Functions and operators, open Probability distributions.

- Normal(average, standard deviation)
- Uniform(minimum, maximum)
- Bernoulli(probability of success)

Bibliography

Adams, B., Hooper-Bui, L., Strecker, R. & O'Brien, D., 2011. Raft formation by the red imported fire ant, Solenopsis invicta. *Journal of Insect Science,* Volume 11, p. 171.

Boser, C. et al., 2014. Strategies for biosecurity on a nearshore island in California. *Monographs of the Western North American Naturalist,* Volume 7, pp. 412-420.

Carlton, J. & Hodder, J., 2003. Maritime mammals: terrestrial mammals as consumers in marine intertidal communities. *Marine Ecology Progress Series,* Volume 256, pp. 271-286.

Clavero, M. & Garcia-Berthou, E., 2005. Invasive species are a leading cause of animal extinctions. *TRENDS in Ecology and Evolution,* 20(3), p. 110.

Cope, R. et al., 2016. Integrative analysis of the physical transport network into Australia. *PLoS One,* Volume 11, p. e0148831.

Dahlstrom, A., Hewitt, C. & Campbell, M., 2011. A review of international, regional and national biosecurity risk assessment frameworks. *Marine Policy,* 35(2), pp. 208-217.

Faulkner, K., Robertson, M., Rouget, M. & Wilson, J., 2017. Prioritising surveillance for alien organisms transported as stowaways on ships travelling to South Africa. *PLoS One,* Volume 12, p. e0173340.

Garcia-Diaz, P., Ross, J., Woolnough, A. & Cassey, P., 2016. Managing the risk of wildlife disease introduction: pathway-level biosecurity for preventing the introduction of alien ranaviruses. *Journal of Applied Ecology,* Volume 54, pp. 234-241.

Gordon, D., Tancig, K., Onderdonk, D. & Gantz, C., 2011. Assessing the invasive potential of biofuel species proposed for Florida and the United States using the Australian Weed Risk Assessment. *Biomass and Bioenergy,* Volume 35, pp. 74-79.

Guja, L., Merritt, D. & Dixon, K., 2010. Buoyancy, salt tolerance and germination of coastal seeds: implications for oceanic hydrochorous dispersal. *Functional Plant Biology,* Volume 37, pp. 1175-1186.

Hintze, C. et al., 2013. D3: The dispersal and diaspore database - Baseline data and statistics on seed dispersal. *Perspectives in Plant Ecology, Evolution and Systematics,* Volume 15, pp. 180-192.

Ives, A. & Carpenter, S., 2007. Stability and diversity of ecosystems. *Science,* Volume 317, p. 10.1126/science.1133258.

Kearney, M. et al., 2008. Modelling species distributions without using species distributions: the cane toad in Australia under current and future climates. *Ecography,* Volume 31, pp. 423-434.

Koch, F., Yamshanov, D. & Haack, R., 2013. Representing uncertainty in a spatial invasion model that incorporates human-mediated dispersal. *NeoBiota,* Volume 18, pp. 173-191.

Korb, K. & Nicholson, A., 2010. *Bayesian artificial intelligence.* 2nd ed. Boca Raton, FL, USA: CRC Press.

Laws, R. & Kesler, D., 2012. A Bayesian network approach for selecting translocation sites for endangered island birds. *Biological Conservation,* Volume 155, pp. 178-185.

Lohr, C., Passeretto, K., Lohr, M. & Keighery, G., 2015. Prioritising weed management activities in a data deficient environment: the Pilbara islands, Western Australia. *Heliyon,* 1(4), p. e00044.

Lohr, C. et al., 2017. Predicting island biosecurity risk from introduced fauna using Bayesian Belief Networks. *Science of the Total Environment,* Volume 601-602, pp. 1173-1181.

Mack, R. & Lonsdale, W., 2002. *Eradicating invasive plants: Hard-won lessons for islands.* Gland, Switzerland, Occasional Paper of the IUCN Species Survival Commission No. 27. International Union for Conservation of Nature and Natural Resources.

Marcot, B., 2006. Characterizing species at risk I: modeling rare species under the Northwest Forest plan. *Ecology and Society,* Volume 11, p. 10.online.

Marcot, B. et al., 2001. Ysing Bayesian belief networks to evaluate fish and wildlife population viability under land management alternatives from an environmental impact statement. *Forest Ecology and Management,* Volume 153, pp. 29-42.

Martin, T., Kuhnert, P., Mengersen, K. & Possingham, H., 2005. The power of expert opinion in ecological models using Bayesian methods: Impact of grazing on birds. *Ecological Applications,* Volume 15, pp. 266-280.

Moore, J. et al., 2010. Protecting islands from pest invasion: optimal allocation of biosecurity resources between quarantine and surveillance. *Biological Conservation,* 143(5), pp. 1068-1078.

Morris, K. et al., 2015. Forty years of fauna translocations in Western Australia: lessons lerned. In: A. DP, M. Hayward, D. Moro & P. Seddon, eds. *Advances in reintroduction biology of Australian and New Zealand fauna.* Clayton South, Victoria, Australia: CSIRO Publishing, pp. 217-236.

Pheloung, P., Williams, P. & Halloy, S., 1999. A weed risk assessment model for use as a biosecurity tool evaluating plant introductions. *Journal of Environmental Management,* Volume 57, pp. 239-251.

Pickering, C. & Mount, A., 2010. Do tourists disperse weed seed? A global review of unintentional human-mediated terrestrial seed dispersal on clothing, vehicles and horses. *Journal of Sustainable Tourism,* Volume 18, pp. 239-256.

Russell, J., Towns, D. & Clout, M., 2008. *Review of rat invasion biology: implications for island biosecurity.* Wellington: Department of Conservation.

Sakai, A. et al., 2001. The population biology of invasive species. *Annual Reviews,* pp. 305-332.

Scott, J. et al., 2017. Zero-tolerance biosecurity protects high-conservation-value island nature reserve. *Scientific Reports,* Volume 7, p. 772.

Smith, C., Howes, A., Price, B. & McAlpine, C., 2007. Using a Bayesian belief network to predict suitable habitat of an endangered mammal - The Julia Creek dunnart (Sminthopsis douglasi). *Biological Conservation,* Volume 139, pp. 333-347.

Tabak, M., Poncet, S., Passfield, K. & Martinez del Rio, C., 2015. Modeling the distribution of Norway rats (Rattus norvegicus) on offshore islands in the Falkland Islands. *NeoBiota,* Volume 24, pp. 33-48.

Tackenberg, O., Poschlod, P. & and Bonn, S., 2003. Assessment of wind dispersal potential in plant species. *Ecological Monographs,* Volume 73, pp. 191-205.

Tingley, R. et al., 2013. Identifying optimal barriers to halt the invasion of cane toads Rhinella marina in arid Australia. *Journal of Applied Ecology,* Volume 50, pp. 129-137.

Venette, R. et al., 2010. Pest risk maps for invasive alien species: A roadmap for improvement. *BioScience,* 60(5), pp. 349-362.

White, A. & Shine, R., 2009. The extra-limital spread of an invasive species via 'stowaway' dispersal: toad to nowhere?. *Animal Conservation,* Volume 12, pp. 38-45.

Whittle, P. et al., 2013. A method for designing complex biosecurity surveillance systems: detecting non-indigenous species of invertebrates on Barrow Island. *Diversity and Distributions,* Volume 19, pp. 629-639.