Supplement S1: Decision Support Scheme

The following table shows the data sets and parameters that are required for each model.

Table S8-1.1. Data sets and parameters required for each model.

Model		Model A	Model B	Model C	Model D
Data sets	CLIMEX output (or suitability model)	X	X	X	X
	Economic data (GIS map)	X			
Parameters	Carrying capacity P_{max}			X	X
	Yearly multiplication factor λ_{max}			X	X
	Relative rate of spatial increase r	X			
	Spread rate c		X		
	Scale parameter u				X
	Shape parameter ν				X

Decision support scheme on quantitative spread modelling

1. ***'Is it appropriate to map spread?'***

   1.1 Is the qualitative rating of spread within the Pest Risk Analysis (PRA) Scheme insufficient and are quantitative data needed to indicate when the pest is expected to arrive at particular location (depending on the entry/starting point)? **Yes/No**

   1.2 Is more detailed information required to answer questions within the PRA Scheme? **Yes /No**

   1.3 Are detailed estimates required to quantify impacts as they build up over time (e.g. infested area per year)? **Yes/No**

   1.4 Is detailed dynamic information on spread over time required to design and target surveillance campaigns, contingency plans or phytosanitary measures? **Yes/No**

   1.5 Are quantitative data required for the analysis of costs and benefits of phytosanitary measures (e.g. comparison of spread with and without measures)? **Yes/No**

If at least one question has been answered with yes, go to 2

2. ***'What kind of maps/information/data are already available for the suitability of the PRA area for the organism?'***

   2.1 Is a map of endangered areas already available?

   If YES, go to 3
   If NO, go to 2.2

   2.2 Is a CLIMEX model or another climate suitability model giving indices similar to Ecoclimatic Index (EI) and Growth Index (GI) already available?

   If YES and available for the PRA area, go to 3
   If YES but needs to be adjusted for the PRA area, return to the model and go to 2.3
   If NO, attempt to construct a CLIMEX model or another climate suitability model and go to 2.3

   2.3 Has a CLIMEX model or another climate suitability model giving indices similar to EI and GI been successfully developed?

   If YES, go to 3
   If NO, STOP or use. You cannot apply the spread module

2.4 Is a map of hosts (or habitats) already available?

If YES, you can apply the spread module, go to 3
If NO, you can apply the spread module but must assume that suitable hosts (or habitats) are available everywhere, go to 3

3. *'What are the key factors affecting spread of the pest and how much information is available?'*

Collect available information and try to either derive a spread rate (km/year) or to indicate whether short- or long-distance dispersal is relevant for the species.

Spread mechanism	Yes/No	Specify	Short distance (within a range of 1 m–10 km)	Long distance (move/jump to new areas)
			Yes/No (and/or spread rate)	Yes/No (and/or spread rate)
Active movement		e.g. flight of adult beetles		
Passive movement with wind, water, etc.		e.g. spore dispersal with splash water		
Human assistance		e.g. with vehicles, trade, hitchhiking		
Vector needed		e.g. beetle, psyllid, aphid		

Go to 4

4. *'How much information is available on the key data for parameter estimation?'*

Consider the available information and indicate also the uncertainties of this information

Key data for parameters (for pest and/or vector)	Specify	Information available Yes/No	Uncertainty of available data (low – medium – high)
4.1 Spread rate(s) (question 3)	(already done in 3)		
4.2 Entry points, locations of observations	e.g. airports and harbours		
4.3 Data on pest densities observed for at least 2 different years (without eradication or containment measures if possible)	e.g. 5 trees infested in 1 year, 500 trees infested 3 years later		
4.4 Data on highest observed pest densities (carrying capacity of host plant or habitat)	e.g. 20 *Anoplophora* exit holes per tree; 2–2.5 million *Diabrotica* beetles per hectare		
4.5 Data on the lowest pest density to establish a new population	e.g. one male and one female		
4.6 Distribution maps for two different time steps			

Go to 5

5. ***'Based on the information, data and maps available for the organism (question 4.1–4.6) how easy will the parameter estimation and therewith the modelling be?'***

   - **Only yes answers, with**
     - low to medium uncertainty
       → very easy and straightforward, all the spread models can be used; go to 6
     - medium to high uncertainty
       → very easy, all the spread models can be used but best, likely and worst case scenarios should be simulated to evaluate the spread uncertainty; go to 6
   - **More yes than no answers**, **with**
     - low to medium uncertainty
       → quite easy, some spread models can be used; go to 6
     - medium to high uncertainty
       → quite easy, some spread models can be used but best, likely and worst case scenarios should be simulated to evaluate the spread uncertainty; go to 6
   - **Same number of yes and no answers**, **with**
     - low to medium uncertainty
       → possible, some spread models can be used; go to 6
     - medium to high uncertainty
       → possible, some spread models can be used but best, likely and worst case scenarios should be simulated to evaluate the spread uncertainty; go to 6 or STOP
   - **More no than yes answers, with**
     - low to medium uncertainty
       → it might be possible to apply one or two models; go to 6 or STOP
     - medium to high uncertainty
       → Impossible, STOP
   - **Only no answers**
     → Impossible, STOP

6. ***Modelling and mapping the spread potential***

   Before modelling and mapping spread, the area of potential establishment must have already been identified and mapped (see Baker *et al*., 2012). A CLIMEX EI map is also required (to use outputs from other models further work is required).

   One or more spread models can be applied. The models are designed to help to map changes in the endangered area through time and space. A description of the spread models is given in the main text of Chapter 8. The practical application in R is described in the online supplement S2 to Chapter 8.

Supplement S2: R commands to apply the spread module

Step 1: Calling the code and visualizing the input data sets

At the beginning of the modelling work and before applying the models, it is necessary to type the following commands in R to install necessary libraries and read input files and specify the format of the input files (note that the working directory in R should be the workspace of your study species):

```
library(sp)             # calls a library necessary for using spatial data
library(raster)         # calls a library necessary for using raster layers
library(rgdal)          # calls a library necessary for using geospatial data
climexcsv=F             # the CLIMEX output is not given in a ".csv" file
                        # (F means false) but in ".txt" file
                        # otherwise it should be climexcsv=T (T means true)
elevmax=F               # no elevation limit is given, otherwise
                        # elevation should be given in meters, e.g. elevmax=2000.
habitatfile=T           # a habitat file is provided, otherwise habitatfile=F.
habitatformat=".tif"    # the format of the file is ".tif".
                        # if the format was ascii then habitatformat= ".asc"
                        # if it was GRID then it should be habitatformat= NULL
```

Then, **the R code should be loaded**. Click in the R menu on 'File', 'Open a R source code', find the file 'Prog-SpreadModule.R', and click 'Open'. This step ensures that all functions necessary for running the models are known to R. Several commands are available to plot the CLIMEX and habitat maps that act as key inputs:

```
printinfo()          # provides information about the code and the simulations
plotRA()             # plots the suitable area (where EI>0 within the habitat area)
plotGI()             # plots the growth index over the suitable area. The index is
                     # rescaled so that the maximum is 100
plothabitat()        # plots the habitat distribution or host density
```

Step 2: Command lines and arguments of the function to run model A

To run the model, the following line should be typed in R:

```
res=lgecon(N0=0.03,r=0.33,t=20,econraster=T,econformat=”.tif”,mult=1,sim=10,
export=T, name=”OutputA”)
```

This line tells R to call the function *lgecon* with the seven arguments included in the list in round brackets ( ). These arguments are explained in Table S8-2.1. The results of model A are stored in the object *res*, given in the left-hand side of the assignment.

Table S8-2.1. Arguments of the function to simulate spread with model A.

Arguments	Description
N0	The value for n_0
r	The value for *r*
t	The time horizon for the simulations, since $t = 0$ in 1992, then $t = 20$ gives the output of the model for the year 2012. This value can be changed to see the spatio-temporal dynamics
econraster	Whether an economic file is provided (yes = T or no = F)
econformat	The format of the economic file (if provided)
mult	Multiplicative factor (if needed), for instance if the economic value is given per hectare instead of per km^2
valperhost	Multiplicative factor if the economic value is derived from the abundance of the host (in ‘habitat’ file).
sim	Number of replicates for the random case scenario
export	Whether the output of the three scenarios should be exported in GIS files (.tif) (yes = T or no = F).
name	Name of the files that will be exported (if requested), it should be given inside quotes (“Name”)

Outputs of model A are retrieved by typing the object name given to the output (in this case: res) followed by the $ sign, and a name for the specific kind of output desired. Available outputs are listed in Table S8-2.2.

Table S8-2.2. Output values for model A.

Commands	Description
res$ntot	Number of cells considered in the model
res$nRA	Number of cells within the suitable area
res$econmem	List of economic values used in the model
res$ninv	Number of invaded cells at time *t*
res$pniche	Percentage of invaded cells within the suitable area
res$worst	Sum of the economic value over all the invaded cells following the worst case scenario
res$best	Sum of the economic value over all the invaded cells following the best case scenario
res$rand	Summary (quantiles) of the sum of economic value over all the invaded cells across the replicate simulations of the random case scenario

Step 3: Command lines and arguments of the function to run model B

To run the model, the following line should be typed in R:

```
res=radial(RR=80,t=20,coord=c(20.30,44.82),figkm=F,figdd=T,export=T,name="OutputB")
```

This calls the function *radial* with the arguments listed in brackets, and places the results in the object *res*. The meaning of the arguments of the function *radial* is given in Table S8-2.3.

Table S8-2.3. Arguments of the function to simulate spread with model B.

Arguments	Description
RR	Rate of radial range expansion (corresponding to the parameter *c*)
coord	The geographical coordinates (WGS 1984) of the entry point(s) in decimal degrees
t	The time horizon for the simulations, since $t = 0$ in 1992, then $t = 20$ gives the output of the model for the year 2012. This value can be changed to see the spatio-temporal dynamics
figkm	Whether the output map should be plotted in a projected metric space (yes = T or no = F)
figdd	Whether the output map should be plotted in a decimal-degree space (yes = T or no = F)
export	Whether the output should be exported in GIS files (.tif) (yes = T or no = F)
name	Name of the file that will be exported (if requested), it should be given inside quotes ("Name")

The numerical results are retrieved by typing one or more of the commands listed in Table S8-2.4.

Table S8-2.4. Output values for model B.

Commands	Description
res$ntot	Number of cells considered in the model
res$nRA	Number of cells within the suitable area
res$ninv	Number of invaded cells at time *t*
res$pniche	Percentage of invaded cells within the suitable area
res$radial	Coordinates in decimal degrees of the centre of invaded cells (1st column is latitude and 2nd column is longitude)

Step 4: Command lines and arguments of the function to run model C

To run the model, the following line should be typed in R:

```
res = slg(N0= 1.6E-7,lmax=40,t=20,movie=F, export=T, name="OutputC")
```

The meaning of the arguments of the function *slg* is given in Table S8-2.5.

Table S8-2.5. Arguments of the function to simulate spread with model C.

Arguments	Description
N0	Population density (expressed as a percentage of the carrying capacity) at time t = 0 in all suitable cells (it corresponds to the parameter p_0)
lmax	The maximum yearly multiplication factor of population density (λ_{max})
t	The time horizon for the simulations, since t = 0 in 1992, then t = 20 gives the output of the model for the year 2012. This value can be changed to see the spatio-temporal dynamics
movie	Whether a series of spread maps should be plotted for t = 1 to the given time horizon (yes = T or no = F). Click on the map to shift to the following map
export	Whether the output of the three scenarios should be exported in GIS files (.tif) (yes = T or no = F)
name	Name of the files that will be exported (if requested), it should be given inside quotes ("Name").

Several values are calculated by the function *slg* (see Table S8-2.6). To retrieve them, the command in Table S8-2.6 can be used.

Table S8-2.6. Output values for model C.

Commands	Description
res$ntot	Number of cells considered in the model
res$nRA	Number of cells within the suitable area
res$slg	Population density (expressed as the percentage of the carrying capacity) over the grid points
res$sum0	Number of cells with a population density = 0
res$sum25	Number of cells with a population density in [0, 25%]
res$sum50	Number of cells with a population density in [25%, 50%]
res$sum75	Number of cells with a population density in [50%, 75%]
res$sum100	Number of cells with a population density in [75%, 100%]

Step 5: Command lines and arguments of the function to run model D

To run the model, the following line should be typed in R:

```
res = dispk(N0=NULL, lmax=40, p=5, u=80, presencefile=T,figkm=F,figdd=T, t=20)
```

Arguments for *dispk* are given in Table S8-2.7.

Table S8-2.7. Arguments of the function to simulate spread with model D.

Arguments	Description
N0	Population density (expressed as a percentage of the carrying capacity) at time $t = 0$ in all suitable cells (it corresponds to the parameter p_0) = NULL if coordinates and initial population density are given in the file 'presence.txt' = the p_0 *value* if entry points should be randomly selected
lmax	The maximum yearly multiplication factor of population density (λ_{max})
P	The proportion engaged in dispersal (always tested for $P = 1$)
p	Shape parameter of the 2Dt distribution (corresponding to the parameter ν)
u	Scale parameter of the 2Dt distribution (corresponding to the parameter u)
presencefile	Whether a separate file called 'presence.txt' is given to provide the coordinates of entry points and the initial population density (yes = T or no = F).
nentry	Number of entry points to be selected at random if no presence file is given and a value if given for N0.
figkm	Whether the output map should be plotted in a projected metric space (yes = T or no = F)
figdd	Whether the output map should be plotted in a decimal-degree space (yes = T or no = F).
t	The time horizon for the simulations, since $t = 0$ in 1992, then $t = 20$ gives the output of the model for the year 2012. This value can be changed to see the spatio-temporal dynamics

To export the simulated spread in a GIS file, the following command should be written:

```
exportkernel(res$dispk,name="OutputD")
```

To define a threshold above which the species is supposed to be present (e.g. 10%) and plot the potential distribution area, the following command should be typed:

```
plotkernel(res$dispk,presence=res$presence,figkm=F,figdd=T,t=NULL,threshold=10)
```

Table S8-2.8. Output values for model D.

Commands	Description
res$ntot	Number of cells considered in the model
res$nRA	Number of cells within the suitable area
res$ninv	Number of invaded cells at time *t*
res$pniche	Percentage of invaded cells within the suitable area
res$dispk	Population density (expressed as the percentage of the carrying capacity) over the grid points (extended grid)
res$presence	Data from the file 'presence.txt' or randomly generated data