Exercise Hints and Solutions for Chapter 23

Agent-based and Individual-Based Modeling: *A Practical Introduction, 2nd Edition*

Exercise 2

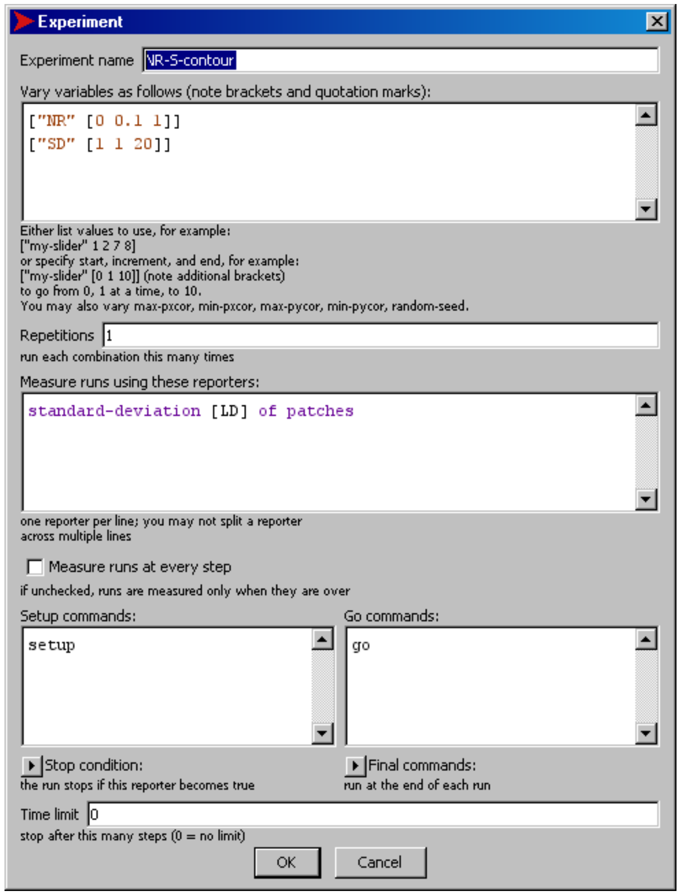
Points that should be addressed in the answer to this exercise include:

- The model result that is contoured in these plots is the fraction of simulations in which the simulated wild dog population went extinct within 100 years. This result can be interpreted as an indicator of the risk that the real dog population would go extinct. Therefore, darker regions of the contour plots represent good management outcomes: lower risk of extinction of the wild dog population.
- The standard values of carrying capacity (60) and adult mortality probability (0.2) are represented by the central contour plot. Under those values, the risk of extinction varies much more with the probability of mortality for dispersers (X axis of the plots) than with the disperser meeting rate (Y axis). The contour lines of equal risk are almost vertical, with risk being only a little less at low values of disperser meeting rate (contours bend a little to the left at the bottom of the plot).
- Parameter interactions would be indicated if the basic pattern of relationship between disperser mortality and disperser meeting rate changed among the different contour plots. Such changes would mean that the relative effect of disperser mortality and meeting rate depends on the values of adult mortality and carrying capacity. However, this general pattern—extinction risk depending mainly on disperser mortality and being a little less at low values of disperser meeting rate—does not change among the nine contour plots.
- Two exceptions to the previous point are the upper left and lower right plots. When carrying capacity is low (40) and adult mortality high (0.25), the disperser mortality and meeting rate parameters are unimportant because the population always goes extinct. The opposite is true when carrying capacity is high (80) and adult mortality low (0.15): disperser parameters are unimportant because the population rarely goes extinct.

Exercise 3

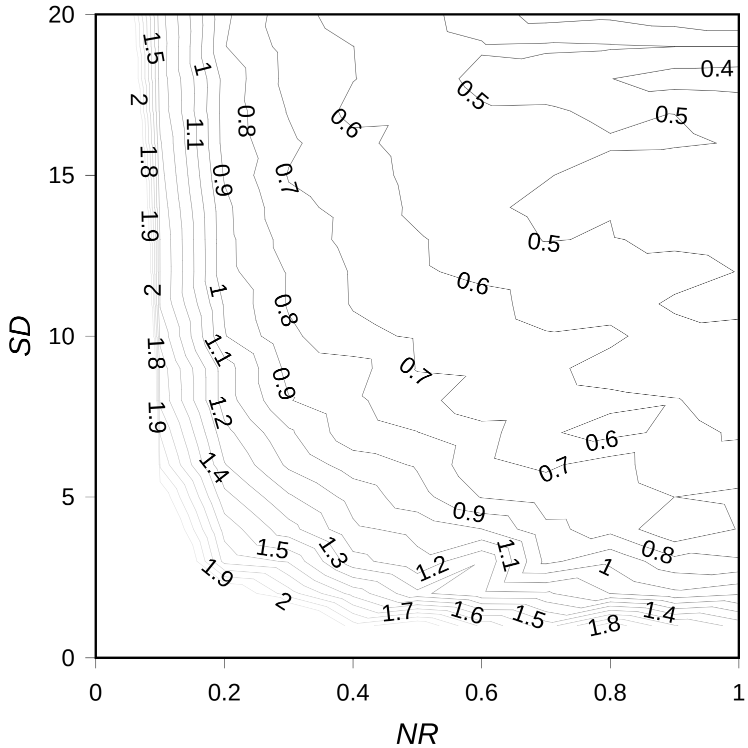
The instructor materials for Chapter 23 include the original NetLogo program for the Breeding Synchrony model. If students use it, they are likely to notice several unusual things about how it is programmed. For example, it does not use `tick` and `ticks`; that is because it was originally written before NetLogo had a built-in tick counter.

Using the NetLogo program we provide, the contour plot for this exercise can be generated using this BehaviorSpace experiment:



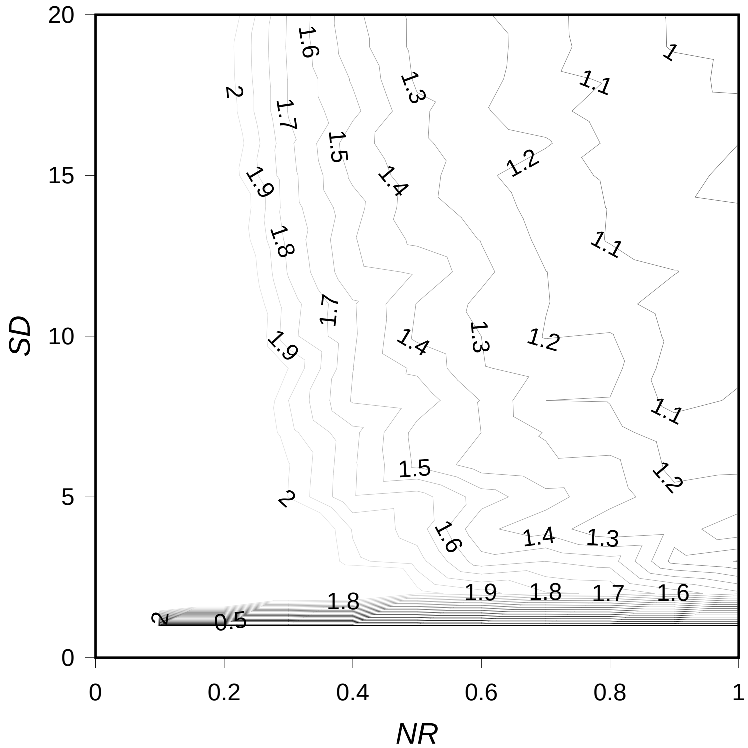
This experiment requires putting the parameter *SD* in a slider. Values of zero for *SD* should be avoided because the birds never breed in that case, so results are meaningless.

The resulting contour plot should look something like this:



This contour is characteristic of a model that is sensitive to the two parameters only at extreme (in this case, low) values. When either *NR* or *SD* is very low, the contour lines are quite close together and the model result varies most strongly with the low parameter. However, when both parameters are high the model is not sensitive to them. The standard deviation in laying date is close to 0.5 days whenever *NR* is above about 0.6 and *SD* is above about 10.

As a robustness analysis example, we modified the model to add noise to the SD term, as Jovani and Grimm (2008) did in the on-line appendix to their publication. We interpreted their description as meaning that, each time step, the model determines stochastically whether each bird has noise added to its stress level (instead of having the same birds have noise each tick). The resulting contour plot with p = 25% of birds having noise in their stress levels is:



The main observation from this plot is that, while noise in stress levels reduced the level of breeding synchrony (the standard deviation in laying dates is approximately doubled), the birds still laid their eggs close together in time. Standard deviation in laying dates was less than 1.4 days over much of the parameter range, which means that 95% of birds laid their eggs within ± 2.8 days of the mean date. The basic sensitivity pattern observed from the previous contour plot (model results sensitive to *NR* and *SD* only when their values are low) appears robust to this change.

One difference is apparent: when *SD* is extremely low (around 1), the bird stress levels do not get low enough for breeding to start over the 200-day duration of the simulation (which causes the extremely close contours at the bottom of the plot).

Exercise 5

Among the methods that should be documented in a solution to this exercise are:

- What currencies were used as key model results to analyze sensitivity of.
- The high and low values for each parameter analyzed.
- The number of values used for each parameter.

Two model outputs that seem important as "currencies" for evaluating disease severity are (1) the fraction of individuals who get the disease—the fraction of turtles that have become resistant by the end of the simulation, and (2) how long the outbreak lasts, which is the tick at which the outbreak stops.

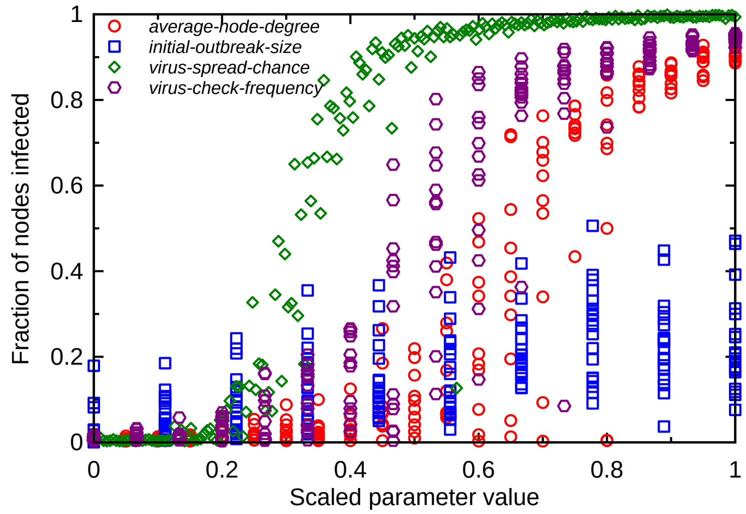
Results should include graphs of how the currencies responded to each parameter, identification of which parameters produced relatively linear responses, and regression statistics for those relatively linear responses.

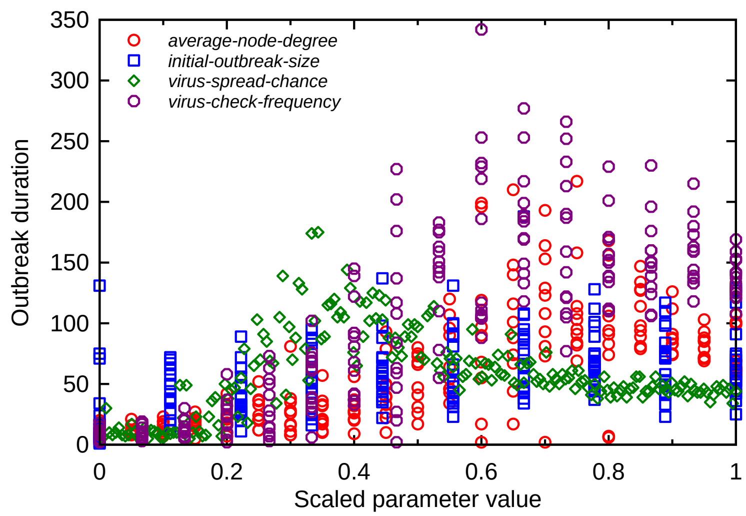
We used the sensitivity analysis design and got the results in the following table, which is based on table 23.1. Note that this analysis differs from that of the Wild Dog model because three of the four parameters analyzed can have only integer values, which limited the number of values that can be simulated. Therefore, we ran enough replicates of each parameter value to bring the total number of simulations close to 200. The parameter ranges are also skewed, with the highest value being much farther from the standard value than the lowest value is, so the standard parameter value does not have a scaled parameter value of 0.5.

Parameter	Standard value	Minimum, step, maximum value; number of replicates	Regression statistics for fraction infected		Regression statistics for outbreak duration	
			Slope	R^2	Slope	R^2
average-node-degree	15 links	[5 1 25] 10	1.03	0.76	106	0.47
initial-outbreak-size	3 people	[1 1 10] 20	0.23	0.41	44	0.21
virus-spread-chance	2.5%	[0.1 0.05 10] 1	1.26	0.78	16	0.02
virus-check-frequency	10 ticks	[5 1 20] 13	1.19	0.85	186	0.59

The most important results are the graphical ones, illustrated below. One important conclusion is that three of the four parameters caused highly nonlinear, logistic-curve-shaped responses in the fraction of the population that was infected; each of the parameters except *initial-outbreak-size* is capable by itself of tipping the simulation from a short, minor outbreak to complete infection of the population. Therefore, the linear regression statistics are not a good way to quantify sensitivity of fraction infected to parameter values.

A second conclusion is that the two “currencies” responded very differently to parameter values. While fraction infected generally rose as parameter values increased, the outbreak duration tended to peak and then decrease. The cause of this difference should be clear: as infection spreads more rapidly, outbreaks initially last longer because more nodes are infected; but when infection is very rapid, outbreaks end sooner because the whole population is rapidly infected. The linear regression statistics are also not a good way to quantify sensitivity of outbreak duration to parameter values.





Exercise 6

Some hints for this exercise:

- Identifying reasonable distributions for each parameter would be much easier if the model were applied to a specific real disease in a specific place or culture. If it were, the research literature on that disease could be used to identify typical values of parameters and ranges of uncertainty.
- The importance of uncertainty in predicting the relative benefits of alternative intervention strategies can be evaluated using the methods discussed at the end of Section 23.3. Identify parameter scenarios that represent a few such strategies: perhaps a lower value of *virus-spread-chance* to represent a strategy of emphasizing sanitation to reduce virus spread when people meet, a lower value of *average-node-degree* to represent a

strategy of reducing contact among people, and a lower value of *virus-check-frequency* to represent treatment to speed recovery after infection. Simulate each of these scenarios many times with the same range of deviation in parameter values. Are any of the alternative strategies consistently best or worst over all the parameter combinations?

Exercise 8

The approach described in this exercise could work, keeping in mind that the BehaviorSpace experiment would also need 500 replicates of each parameter combination to evaluate the frequency of extinction.

One limitation of this approach is the large number of model runs. Even with only 4-5 values per parameter, this experiment has 12,500 parameter combinations. With 500 replicates, this is 6,250,000 runs—potentially feasible for the Wild Dog model but not for bigger models.

A second limitation is that it gives equal weight to all values of each parameter. The method used in Section 23.3 assumes (because it draws parameter values from a normal distribution) that parameter values nearer the mean value are more likely and hence have more weight in the uncertainty analysis. (More of the model runs have parameter values near the mean than at the extremes of their distributions.) The normal distribution method of Section 23.3 makes more sense *if* the modeler has reason to believe that the "true" parameter value is more likely to be close to the value chosen as its mean. The BehaviorSpace experiment described in this exercise makes more sense if the modeler is less confident in the parameter's standard value and instead is only confident that the "true" value falls within some range. (The BehaviorSpace experiment uses a uniform distribution of parameter values, which does not assume values near the mean are more likely.)

An important advantage of the approach described in this exercise over the stochastic selection of parameter values used in Section 23.3 is that it guarantees that all combinations of ranges for all parameters are evaluated. For example, the experiment is guaranteed to include runs in which all parameters have very low values and runs in which all parameters have high values. With random selection of parameter values, far more model runs would be needed to have confidence that rare combinations (e.g., extremely low values of all six parameters) are included.

(The "Latin hypercube" technique mentioned in Section 24.5 is designed to combine the advantages of both approaches. It generates sets of parameter values that can give more weight to values closer to the mean, while including combinations of all ranges of all parameters, while keeping the number of parameter sets relatively small. This technique is described in general modeling and uncertainty analysis references, and by: Rose, K. A. 1989. Sensitivity analysis in ecological simulation models. Pages 4230-4234 in M. G. Singh, editor. *Systems and Control Encyclopedia*. Pergamon Press, New York. It is also available in the R-based analysis tools we discuss in Chapter 24.)