PROGRAM SPACECAP Version 1.0

Suggested Citation:

Pallavi Singh, Arjun M. Gopalaswamy, Andrew J. Royle, N. Samba Kumar and K. Ullas Karanth (2010). SPACECAP: A Program to Estimate Animal Abundance and Density using Bayesian Spatially-Explicit Capture-Recapture Models. Wildlife Conservation Society - India Program, Centre for Wildlife Studies, Bangalure, India. Version 1.0.

OPERATING MANUAL AND HELP

Introduction

SPACECAP, is a user-friendly software package for estimating animal densities using closed model capture-recapture sampling specifically based on photographic captures. It implements a new generation of spatially explicit capture-recapture models developed in a recent paper titled "Bayesian Inference in Camera Trapping Studies for a Class of Spatially Explicit Capture Models" by J. A Royle, K. U. Karanth, A. M. Gopalaswamy, and N. S. Kumar, published in *Ecology,* 90(11), 2009, pp. 3233-3244.

The conventional approach to the analysis of animal density from camera trap surveys is to apply closed capture-recapture model analyses, and, then convert resulting estimates of abundances to densities using a wide range of essentially ad hoc methods. For example, a standard practice in camera trapping studies has been to use auxiliary capture location information to estimate the mean or maximum distance moved by the study species as a basis to apply a "buffer width" around the trap array or, to use various other heuristic "adjustments" (Wilson and Anderson 1985a, b, Karanth and Nichols 1998, Parmenter et al. 2003, Trolle and Kery 2003) to estimate the effectively sampled area. While these approaches appear to work adequately in practice, little had been known about the range of conditions under which they work well. This is because most real world study situations involve study area of odd shapes and sizes and difficult terrain that makes setting camera traps challenging and conditions assumed by ad hoc approaches may not apply.

Spatially explicit capture-recapture models implemented in SPACECAP package directly estimate animal density by explicitly using the information on capture histories in combination with spatial locations of captures under a unified Bayesian modeling framework. This approach offers advantage such as: substantially dealing with problems posed by individual heterogeneity in capture probabilities in conventional capture-recapture analyses. It also offers non-asymptotic inferences which are more appropriate for small samples of capture data typical of photo-capture studies. Further technical details about the models and analyses can be obtained from the paper by Royle et al. (2009), cited above. We also note another spatially explicit modeling and estimation approach based on conventional likelihood based inference that is available (Borchers and Efford 2008).

The Spatial Capture-Recapture Model and its Parameters

The model considered in the current version of SPACECAP applies to binary observations $y(i,j,k)$ for individual "i", trap "j" and sample occasion (e.g., night) "k". The model is a type of binary regression model, similar to logistic regression, in which
$$y(i,j,k) \sim \text{Bernoulli}(p(i,j,k))$$
Here $p(I, j, k)$ is the probability of detecting an individual "i", at trap "j" and sample occasion "k". The probability $p(i,j,k)$ is then related to suitable covariates of interest by applying a suitable transformation. In SPACECAP we make use of the complementary log-log link transformation. Thus, the simplest possible model is expressed

cloglog(p(i,j,k)) = b0

where b0 is a parameter to be estimated. By taking the inverse of the cloglog transformation

p(i,j,k) = 1-exp(-exp(b0))

This particular function arises by considering the binary observations to be formally reduced from a Poisson encounter frequency model (see Royle et al 2009 for details). That is, p(i,j,k) is the Pr(at least 1 encounter in a trap) under the Poisson model. The most general model that SPACECAP presently allows is a model in which

cloglog(p(i,j,k)) = b0 + b1*x(i,j,k) + b2*dist(s(i),u(j))^2

In this expression x(i,j,k) is an indicator of previous capture of individual i in trap j. Thus "b1" is a measure of the behavioral response (see below for more discussion of this; also see Royle et al 2009 for more details). The parameter b2 is a regression coefficient on the effect of distance between individual activity center s(i) and the location of trap j, u(j). b2 is constrained to be < 0 implying that the probability of encountering an individual in a trap decreases as the distance between the individual and the trap increases. In the model, s(i) are the collection of animal locations within a prescribed region S (this is provided by the user – see below). S is referred to subsequently as the state-space of the random variables s(i).

In SPACECAP, the parameter N is the population size of individuals – the number of activity centers located in S (see further discussion of this below).
In SPACECAP, N is reported under the name Nsuper.
Density, D = N/||S|| where ||S|| is the area of the state-space. The units for Density, D is animals/100 km^2.

SPACECAP also reports several derived parameters:
lam0 = exp(b0), is the intercept in terms of expected encounter frequency; this may be thought of as the expected encounter rate of an individual "i" in trap location "j" at sampling occasion "k", whose home-range centre is exactly at the trap location.
sigma = sqrt(1/b2), is the scale parameter of a bivariate normal encounter function (or an exponential, but this option isn't provided in SPACECAP yet). This may also be viewed as a "range parameter" of an animal. For a highly mobile animal, this value will tend to be large (eg:- tigers will have a higher sigma value in comparison to civets)
beta = b1, is the regression coefficient that measures the behavioral response
psi is the ratio of the number of animals actually present within S to the maximum allowable number (set by the user during data augmentation – see below).

When the "behavioral response" option (see below) is chosen, SPACECAP estimates two more derived parameters:
p1 = 1-exp(-lam0), is the encounter probability of an individual before initial encounter.
p2 = 1-exp(-beta), is the encounter probability of an individual after initial encounter.

Installing PROGRAM SPACECAP Version 1.0

STEP 1: Download latest version of program R (R Development Core Team)

SPACECAP works within the R programming environment, so your very first step will be to connect to the internet, go to the website http://www.r-project.org, download and install the latest version of R (R 2.9.2 or higher) from the nearest CRAN mirror.

STEP 2: Download the package "SPACECAP" to your computer.

When you launch R, go to Packages→Install package(s), once again select your nearest CRAN mirror and select package SPACECAP for installation.

STEP 3: OPEN the package SPACECAP

In the R environment, go to Packages→Load package, select SPACECAP and load the package.

STEP 4: Launching SPACECAP

In the R environment, at the prompt ">", type the command SPACECAP().
This will launch the Graphic-User Interface of SPACECAP and you are now set to begin the Bayesian Spatially-Explicit Capture-Recapture (SECR) analysis of your camera trap survey data.

IF YOU FACE ANY PROBLEM IN INSTALLING THIS PACKAGE, PLEASE WRITE TO <pallavi.cws@gmail.com>

SECR Analysis using SPACECAP

Running an SECR Analysis in SPACECAP essentially involves four simple steps:

1. Setting up the input files
2. Selecting the appropriate model combination
3. Selecting the Markov chain Monte Carlo (MCMC) settings
4. Hitting the "RUN" button.

STEP 1: SETTING UP THE INPUT FILES FOR ANALYSIS

SPACECAP requires you to first create **three** input files and store these files at a suitable location on your computer. These files are:

1. Animal Capture Details File (Animal ID no., Trap Location no., Sampling Occasion no.)

2. Trap Deployment details File (Trap Spatial Location, Deployment Activity, Sampling Occasion no.)

3. State-space details File (describing the Potential Animal Home Range Center Details in terms of their spatial location and habitat suitability indicator for these home range centers)

These three raw data files can most easily be created using spreadsheet applications like Microsoft EXCEL, OpenOffice or other software you are comfortable with. However, eventually, all files must be saved in an ASCII comma separated format (.csv), because SPACECAP can only read these types of input files.

Upon launching SPACECAP, you will notice on the input data panel, three separate buttons to load the three input data files. Pressing of these buttons will help you locate the corresponding input files using your desktop browser (like Windows Explorer or Finder).

CREATING INPUT FILES

INPUT FILE 1: Animal Capture Details

The input file containing individual animal capture histories and locations consists of a 3-column table, each column representing the Location Number, the Animal Identity Number and the Sampling Occasion number, strictly in that order. Please note that these are all "number" fields. Please not enter labels containing alpha-numeric characters such as "CAT-123", "PLACE-100" or "January 2009" etc. for these fields, because SPACECAP will not recognize them. Use simple integer numbers. Each unique individual captured during camera trap sampling should be given a

unique identification number, ranging from 1 to n, where n is the total number of unique individuals caught during the camera trap survey.

Duration of the overall survey is determined by species biology in order to meet the assumption of demographic closure. The duration of sampling occasions (or periods) will, in turn, be based on how many such occasions are there in the survey duration.

Because of shortage or camera traps or logistical reasons in many camera trap surveys, the study area is quite often divided into blocks or sub-units and camera trapping is conducted sequentially in these (Karanth and Nichols 2002). With data from conventional capture-recapture analyses the resulting capture history matrix is artificially constructed to ensure that assumption that camera trap survey simultaneously covers the entire area on each sampling occasion (See Karanth and Nichols 2002 page 132-133).

However, since spatial locations of traps are explicitly incorporated into capture-recapture modeling in SPACECAP this artificial construction of sampling occasions is not required, which greatly enhances flexibility in survey and analyses. Each sampling occasion must have a unique identity number, ranging from 1: T, where T is the total number of sampling occasions. We note that each sampling occasion need not cover the entire survey area.

Each camera trap location must be given a unique identification number, ranging from 1: J, where J is the total number of camera trap locations used in the survey.

Assuming a camera trapping survey was conducted from 10 Jan 2009 to 30 Jan 2009, and, we treat each day as a sampling occasion, we end up with 20 sampling occasions. Let us assume there were 16 camera trap locations used in this survey and for logistical reasons the study area was partitioned into 4 blocks and each block contained 4 camera trap locations. Let us assume further that camera traps were deployed for 5 successive days in each block covering each of the four blocks successively in 20 days. Assuming only 6 animals were photo-captured and identified, the **INPUT FILE 1** for SPACECAP would look like:

LOC_ID	ANIMAL_ID	SO
1	4	17
5	5	13
7	6	17
8	6	16
9	3	20
11	1	14
15	2	12

The first data row tells us that Animal ID no 4 was captured at Location ID 1 on the 17th sampling occasion. In your spreadsheet application, you can build up the animal capture details data as per the above format (please make sure that you INCLUDE the header row with the exact titles on column headings as shown above). The file must eventually be saved as a comma delimited ASCII file with an appropriate file name and .csv extension (for example, "captures.csv") and saved in the working directory.

INPUT FILE 2: Trap Deployment Details

In many camera trap sample surveys of animals (as in the above example), all camera trap stations in the study area may not be operational simultaneously for logistical reasons (for example: limited number of cameras or manpower). Therefore, the trap deployment details input file provides SPACECAP with the information on the dates when each camera trap location was active and operational during the survey. Some trap-nights or trap-days of capture data may be "lost" as a result of camera trap failure, theft, vandalism or animal-damage. This type of trap

activity/passivity information can also be effectively fed into and used in SPACECAP. The trap deployment details file records both these types of information, thus accurately accounting for trapping effort.

The trap deployment data are stored in a two dimensional matrix of camera trap locations and sampling occasions in a binary, 0/1 format, where 0 indicates that a particular camera trap station was NOT operational on a particular sampling occasion, and 1 indicates that it was operational. The trap location is denoted in 3 columns in the table representing the Trap Location ID no., the spatial location expressed in X and Y-coordinates (in Universal Transverse Mercator UTM projection system in GIS). It is important that these coordinates are represented in the UTM projection system, because it is used for all distance measurements and computations in SPACECAP.

We illustrate trap deployment data entry using a Table for the same example of **animal capture details** described above. Please recall that the camera trapping survey was conducted with a camera trap array of 16 trap locations, deployed in 4 blocks with 4 camera trap sites in each block. The camera trap survey was conducted over 20 sampling occasions, during which each trapping block was sampled over 5 sampling occasions. The resulting TRAP DEPLOYMENT DATA file would look like the one below:

LOC_ID	X_Coord	Y_Coord	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	619303	1325966	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	624151	1325013	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	624722	1323864	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	621806	1322453	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	622451	1320137	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
6	622599	1317937	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
7	623179	1315941	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
8	625156	1315587	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
9	626022	1314224	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0
10	627568	1315494	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0
11	619604	1324739	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1	0	0	0	0	0
12	621478	1324515	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0
13	623317	1323989	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1
14	624406	1321603	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1
15	624482	1320577	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1
16	629229	1319793	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1

The table shows that the camera trap sites (Loc 1-4) were operational during sampling occasions 1-5 but weren't operational on the remaining sampling occasions 6-20. Additionally, an odd '0' corresponding to Loc 3 and sampling occasion 3 indicates that a trap night was "lost" here.

The trap deployment details file should be constructed exactly as above in your spreadsheet application. It must then be converted to a comma separated ASCII file (.csv file) with an appropriate filename (e.g. traps .csv) in the working directory accessed by SPACECAP.

INPUT FILE 3: Potential Home-Range Centers

In SPACECAP analyses, the surveyed area containing the camera trap array combined with an extended area surrounding it, known as the "state-space" of the underlying point process, say S, which is represented by a large number of equally spaced points in the form of a very fine mesh. These points are visualized as representing all possible potential activity centers (or home range centers) of all the animals in the animal population being surveyed. This fine grid or mesh of

points can be easily generated using a GIS software (like ArcView, MAPINFO etc.) as briefly described below. We view this as an approximation to an underlying continuous state-space which, in practice, would normally be difficult to characterize for computational purposes except in very basic situations where regular polygons might be reasonable. While estimates of population size, N, will be sensitive to the size and extent of the state-space, the estimated density $D=N/\|S\|$ is invariant as the extent of the state-space increases. Thus, S should be chosen to be sufficiently large so as to ensure stability of the density estimate. Conceptually, this occurs (under the models fitted in SPACECAP) by choosing S to buffer the trap array by 2 or 3 times the encounter probability scale parameter.

First a rectangle is formed by connecting the outermost camera trap locations using the GIS software of your choice. This rectangle is called the "Minimum area rectangle". A buffer distance (which is sufficiently large to ensure that no individual animal outside of the buffered region has any probability of being photo-captured by the camera traps in the array during the survey) is added to the rectangle around encompassing the trap array. Thereafter, using GIS numerous equally spaced points representing home range centers are generated for this extended area. In practice, some of these potential home range centers in the mesh may end up in habitats known to be entirely unsuitable for the study species (say in the middle of a village, for tiger data). SPACECAP appropriately deals with this problem because "Grid Cells" input file clearly specifies which of these potential home range/activity centers lie within suitable habitat and which do not.

The potential home-range centers data file essentially consists of a 3 column table. The first two columns are the X and Y coordinates (both in UTM projection system) of all the potential activity centers (the equally spaced points generated from your GIS software) and the third column is a habitat suitability indicator column, indicated with 1s if the potential activity centers lies within suitable species habitat or with 0s otherwise. The format of this file is as given below (this example is a small subset of the actual data set, purely for illustration):

Example of potential home-range centers input file format:

X_COORD	Y_COORD	HABITAT
611734	1299581	1
612301.6	1299583	1
612869.2	1299585	0
613436.9	1299587	0
614004.5	1299589	0
614572.2	1299591	0
615139.8	1299593	0
615707.5	1299595	0
616275.1	1299597	1
616842.8	1299600	0
617410.4	1299602	1
617978.1	1299604	1
618545.7	1299606	1

Such a table must be created for all the potential activity centers (this will be a very large table) and saved as a comma delimited ASCII file (.csv file) with an appropriate name, for example, centers.csv, and saved in the working directory.

In the input data panel, there is a text box for specifying the area of potential home-range centers. This area may be imagined to be composed of a point at the centre of a square, which we call "pixel". Please enter the pixel size area used by the GIS software in this box. The "fineness of the mesh" determined by the pixel size used for spacing of potential home range centers is dictated by species biology (e.g. a few hundred meters for tigers but only a few meters for civets). Caution

should be exercised here to not specify a state-space that is too fine because the MCMC algorithm run time increases linearly with the size of the state-space grid. Regarding this as a discrete approximation to some underlying continuous state-space probably justifies a reasonably coarse state-space grid. Regardless of the desired state-space dimension, we recommend carrying out a trial run with a coarse grid to evaluate the performance.

Please load these files using on screen 'buttons' provided in the input data panel of SPACECAP. Please specify the pixel size area of a potential home-range center (as created in input file 3 above) in square kilometers and then click on "OK". Please check the frame at the bottom for status or error messages. In case you wish to edit your selection, please click on the "Edit" button and start the selection all over again.

Proceed to the Model Definition frame.

STEP 2: SELECTING THE APPROPRIATE MODEL COMBINATION FOR ANALYSIS

The Model Definition panel of SPACECAP consists of a set of options to select an appropriate model combination for the Spatial-Capture Recapture Analysis. These are simple radio buttons indicating each model choice. Some of these model options that are "grayed out" are expected to be made available in future developments of SPACECAP.

The model choices are:

1. Trap response present OR Trap response absent
Selecting "**Trap response present** " option runs the behavioral response option (equivalent to Model "Mb"). Select the "**Trap response absent**" option if you decide otherwise. This model implements a local or "trap-specific" behavioral response under which the probability of encounter in a trap increases subsequent to initial capture in that trap. This is in contrast to the conventional "global" behavioral response which parameterizes a constant increase in encounter probability (on the logit scale, usually) that is not trap specific.

2. **Spatial Capture-Recapture OR Non-spatial Capture-Recapture**
Select "**Spatial Capture-Recapture**" for running a spatially explicit capture-recapture analysis, or "**Non-spatial Capture-Recapture**" for running a conventional capture-recapture analysis (this is equivalent to the Null Model "Mo" in non-spatial CR analysis)

3. Half Normal OR Negative Exponential
Currently SPACECAP analyses SECR models with only the Half-Normal detection function.

4. Bernoulli (binary) OR Poisson encounter process
 Currently the analysis is run with the Bernoulli encounter model in which the probability of success is derived as the probability of a positive response under a Poisson encounter rate model. This motivates use of the complementary log-log link which relates encounter probability to distance and other covariates. After the model definition is complete, please click on "OK". Please check the frame at the bottom for status or error messages. In case you wish to edit your selections, please click on the "Edit" button, change your model definition and click on "OK" again. Proceed to the MCMC simulation settings frame.

STEP 3: Setting the Markov-Chain Monte Carlo (MCMC) parameters (for advanced users)

SPACECAP, uses the Markov-Chain Monte Carlo simulation algorithm written in Program R (R Development Core Team) to estimate the parameters of the Spatially-explicit Capture Recapture models of Royle et al 2009). The relevant settings can be set in the MCMC simulation settings panel of SPACECAP.

No of iterations – This defines the number of MCMC iterations for the analysis (if you aren't sure what this means please set this to a value of about 50,000)

Burn-in – This defines the number of initial values to discard during the MCMC analysis. Setting this at about 1000 usually works well in our experience. Note that some evaluation of whether this is sufficient should be carried out using conventional methods (a topic we will address in subsequent releases).

Thinning – This defines the thinning rate. Only iteration numbers defined by the thinning rate are stored during the analysis (if you aren't sure of what this means, you may set this at a value of 1 – that is, no thinning. If you are setting it to 1, please make sure that the number of iterations specified is less than 65,000. This is because Microsoft Excel, which in most versions has a maximum limit of 65000 rows, will not be able to open our output file).

Data augmentation – Since we are uncertain about the total number of animals, which is likely to be larger than the minimum number caught during your camera trapping survey, you will need to "augment" this value by a certain amount. Ideally, you would like this to be a very large number relative to the number you have caught, but setting it up to be very high will cause the analysis to run for a very long time. As a rule of thumb, you could set this to a value of about 5-10 times the number of animals you have photo-captured during the survey. Data augmentation (DA) is a computational device that enables a convenient Bayesian analysis of capture-recapture models where N is unknown. In the context of SECR models, N is the population of individuals having their activity centers on the prescribed state-space. The basic idea of DA is to provide an upper bound on N, say M, which is equal to N plus the number of augmented individuals. Technically, M is the upper limit of a uniform(0,M) prior for N, which is a customary "noninformative" prior for N in this context. As a practical matter, data augmentation creates a list of pseudo-individuals that are always available for the MCMC algorithm to "use" if necessary. That is, these pseudo-individuals leave and enter the population depending on the current values of the model parameters. See Royle, Dorazio and Link (2007) for some general context and Royle et al. (2009) for details in the context of spatial capture-recapture models.

After the MCMC simulation values have been specified, please click on "OK". Please check the frame at the bottom for status or error messages. In case you wish to edit these settings, please click on the "Edit" button, edit these values and click on "OK" again. You are now all set to start the analysis.

STEP 4: Running the analysis

The last step will simply involve activating the RUN option in the top menu bar. This will start performing the analysis and you will see a progress bar indicating the status of the analysis. Currently, an analysis involving 50000 iterations takes about 14 hours on a fast computer. We are working hard to make the algorithms run faster – we promise!

Results

The posterior density estimates along with standard errors appear as a table in the output panel upon the completion of the analysis. This table also reports estimates of parameters lam0, sigma, psi and beta. If the analysis was run with trap response present, the estimates of "p1" and "p2" are also reported.

Additionally, all the results are written into a comma separated file (called param_val_<timestamp>.csv) and is saved into the current working directory. All the summary statistics are written into a file called summary_stats_<timestamp>.csv, which is also saved to the current directory. And the posterior density graphs of all parameters are all stored as jpeg files (.jpg) in the current working directory.

Further Developments in progress

We plan to add a number of additional utility modules for various applications, which we will be incorporating in the subsequent versions to be released in next few months. Some of the features we are currently working on are:

1. Include the Poisson encounter process model.
2. Include the Negative-Exponential distance function model.
3. Provide with the Hazard-Rate distance function model.
4. Generate FIT statistics for model selection using the Bayesian p-value.

Literature Cited

Borchers, D. L., and M. G. Efford. 2008. Spatially explicit maximum likelihood methods for capture-recapture studies. Biometrics 64:377-385.

Karanth, K. U., and J. D. Nichols. 1998. Estimation of tiger densities in India using photographic captures and recaptures. Ecology 79:2852-2862.

Parmenter, R., et al. 2003. Small-mammal density estimation: a field comparison of grid-based vs. web-based density estimators. Ecological Monographs 73:1–26.

R Development Core Team.2005. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Royle, J. A., R. M. Dorazio, and W. A. Link. 2007. Analysis of multinomial models with unknown index using data augmentation. Journal of Computational and Graphical Statistics 16:67-85.

Royle, J. A., K. U. Karanth, A. M. Gopalaswamy and N. S. Kumar. 2009. Bayesian inference in camera trapping studies for a class of spatial capture-recapture models. Ecology 90(11), 3233-3244.

Trolle, M., and M. Ke´ry. 2003. Estimation of ocelot density in the Pantanal using capture–recapture analysis of camera trapping data. Journal of Mammalogy 84:607–614.

Wilson, K. R., and D. R. Anderson. 1985a. Evaluation of a nested grid approach for estimating density. Journal of Wildlife Management 49:675-678.

Wilson, K. R., and D. R. Anderson. 1985b. Evaluation of two density estimators of small mammal population size. Journal of Mammalogy 66:13-21.