

A Simulation Model to Evaluate the Efficiency of Adaptive Cluster Sampling

Jesse A. Marks, jmarks@centralmethodist.edu

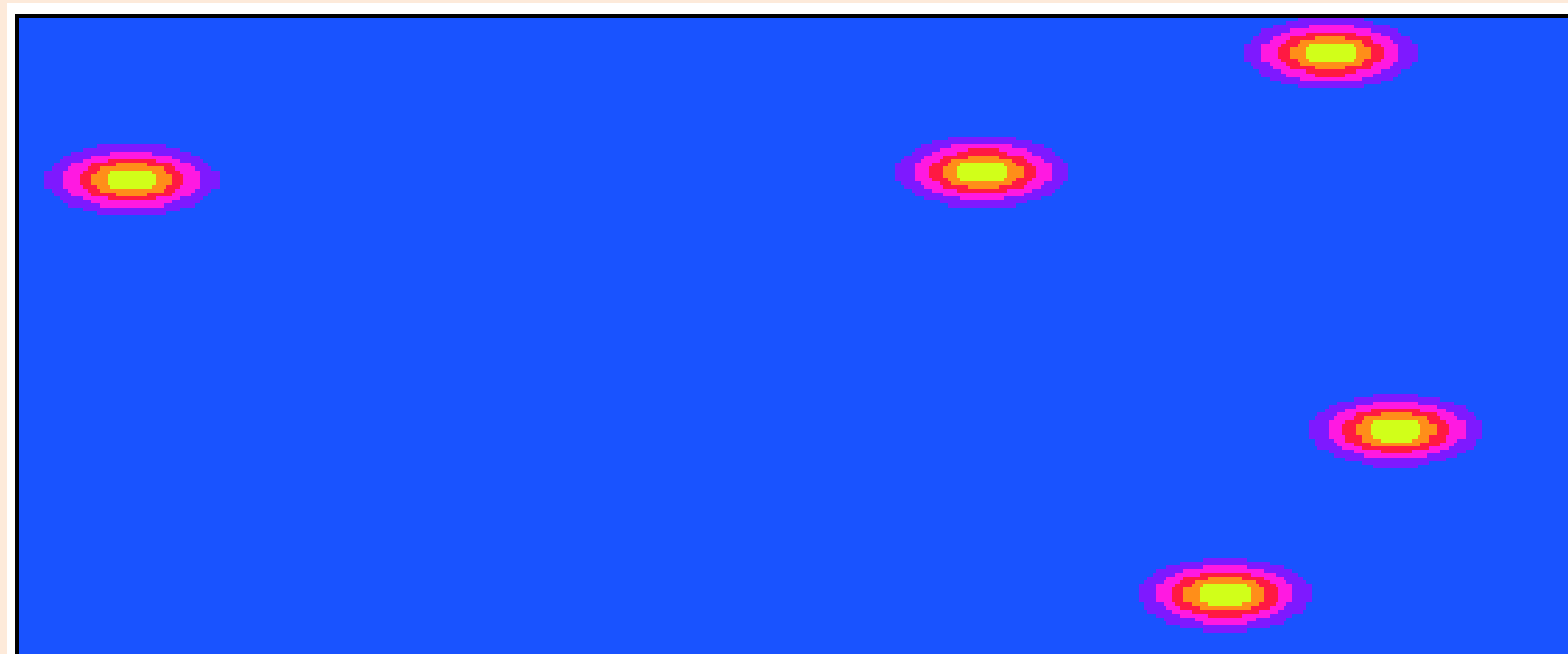
Robert T. Leaf, robert.t.leaf@usm.edu



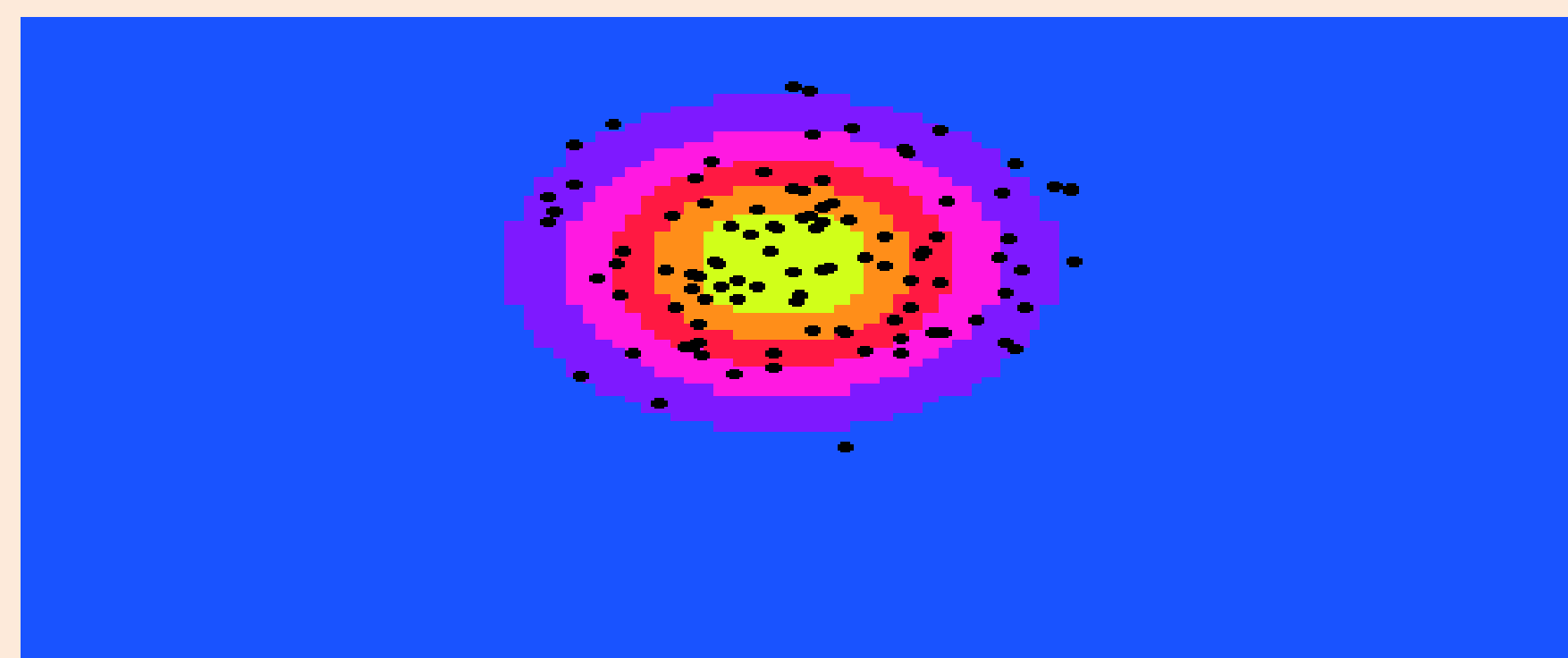
- The **objective** of this simulation analysis is to examine how sample allocation strategy influences the precision of total population estimates in a simulation approach and to use the properties of the estimator to refine the survey design of eastern oyster populations in the Chesapeake Bay.

- The simulation includes the following components:
 - Initialization of the two-dimensional heterogeneous **spatial landscape** at two scales: a landscape scale and a smaller patch scale.
 - Create particles that are the **target entities** of the sampling.
 - Construct **sampling units** used to estimate the oyster density.
 - Compare the results of the simulated sampling regime using cluster and random sampling algorithms

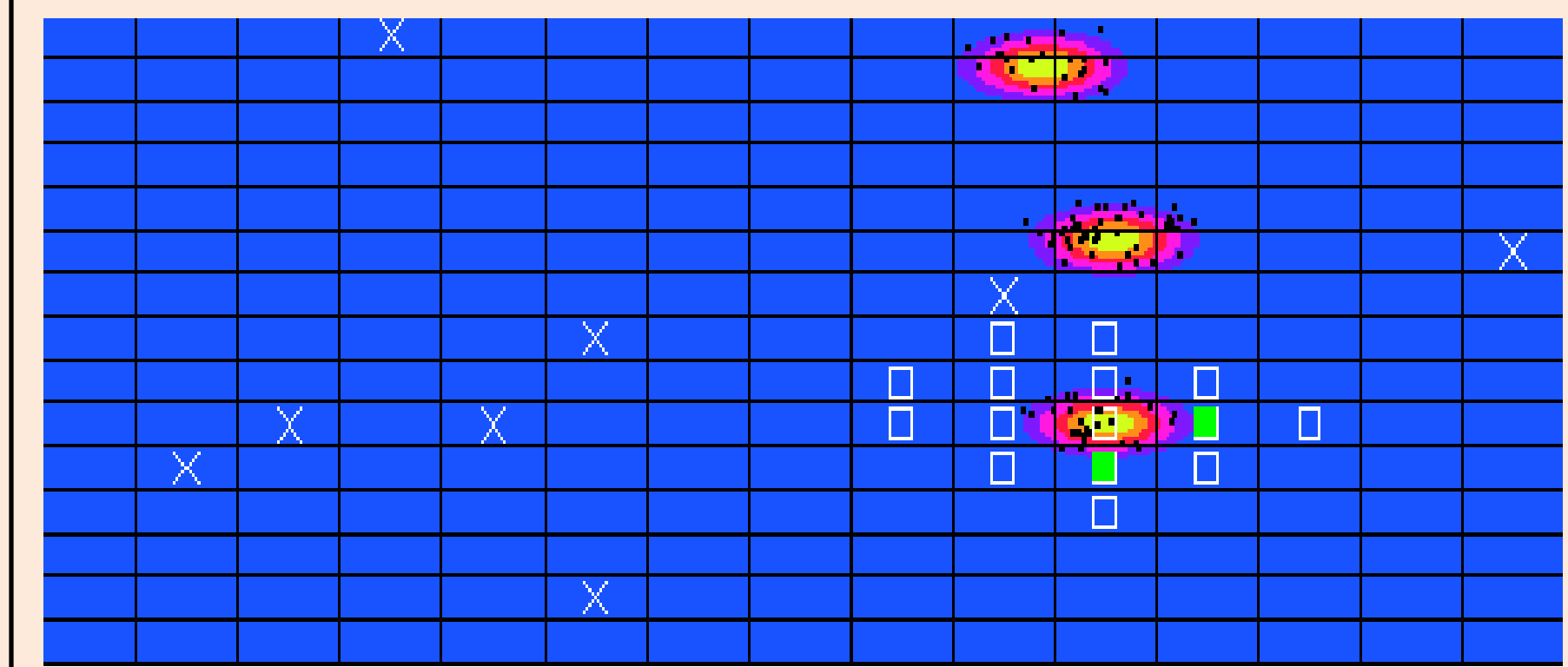
- The dimensions of the landscape were established using vertical and horizontal dimensions of 500 by 500 cells with cell-specific of arbitrary resolution (0.02 by 0.02).
- Random parameter, bivariate skewed-normal distribution within patches were used to determine the probability of occurrence of a particle at that location.
- The parameters that define the variance of the distribution control patchiness.



- A grid function, was used to allocate the landscape into evenly sized quadrilateral cells (N x N).
- The area of one cell represents the area that can be searched with one sample.
- The cell-specific resolution can be altered by the user.
- Particles are assigned to the landscape based on probability set within a patch.
- The number of particles assigned to the landscape, n_p , is a variable determined by the user.



- Initial cells to sample from the landscape are selected using a random number generator.
- When a cell contains at or above a critical value C number of particles, the cells in that neighborhood, or adjacent cells, are searched as well.
- This iterative process continues until no new cells are adaptively searched.
- The initial number of cells to search, the critical value C, and also the definition of a neighborhood are all variables that can be altered by the user.



Variables, Parameters and Equations of ACS and SRS

Variables

N = total number of cells in the landscape.

n = number of initial samples taken.

n^* = number of samples taken during SRS (the total amount searched during ACS)

y_i = The number of particles in cell i .

Ψ_i = The indices of the cells that belong to the network that the initial cell i belongs to.

m_i = The number of plots in network Ψ_i .

$\tau_i = \sum_{j \in \Psi_i} y_j$ = The total number of objects per plot in network Ψ_i .

$w_i = \frac{\tau_i}{m_i}$ = The mean number of objects per plot in network Ψ_i .

ACS Mean and Variance Estimators

Hansen-Hurwitz mean per plot estimate- $\hat{\mu}_{acs} = \frac{1}{n} \sum_{i=1}^n w_i$

Hansen-Hurwitz variance estimate- $\widehat{Var}(\hat{\mu}_{HH}) = \left(\frac{N-n}{N}\right) \left(\frac{1}{n(n-1)}\right) \sum_{i=1}^n (w_i - \hat{\mu}_{acs})^2$

SRS Mean and Variance Estimators

Simple random sample mean per plot estimate- $\hat{\mu}_{srs} = \frac{1}{n^*} \sum_{i=1}^{n^*} y_i$

Simple random sample variance estimate- $\widehat{Var}(\hat{\mu}_{srs}) = \left(\frac{N-n^*}{N}\right) \left(\frac{1}{n^*(n^*-1)}\right) \sum_{i=1}^{n^*} (y_i - \hat{\mu}_{srs})^2$

Confronting the model with data...

- The observed distribution and abundance of an organism is a conditional probability that is determined by its abundance, scale of aggregation, and availability to the sampling gear.
- To evaluate the adequacy of alternative sampling designs we will:
 - 1.) Compare the observed distribution to that in the model base run.
 - 2.) Alter the parameters in the base model that control the distribution and abundance of particles to match that observed in the field survey. These parameters include:
 - Number of clusters of organism density in the study
 - Number of trawls (samples)
 - Parameters to control the kurtosis and skew of the frequency of occurrence distributions
 - Number of particles
 - Trawl dimension as a fraction of the current grid extent what is the trawl length as a fraction of grid size
 - 3.) Derive a set of alternative sampling strategies that can be employed in future surveys.
 - 4.) Evaluate the ability of each of the alternative sampling strategies provide precision estimates of organismal density.

