UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Spatial Modeling of Wildlife Crime Events

Alta de Waal

Department of Statistics
University of Pretoria, South Africa

Overview

- Introduction
- Spatial Bayesian Networks
- Rhino Anti-Poaching
- Evaluation


ZIMBABWE
BOTSWANA
NAMIBIA
MOZAMBIQUE
Limpopo
Kruger National Park
Pretoria
Johannesburg
SWAZILAND
INDIAN OCEAN
ATLANTIC OCEAN
Orange
Bloemfontein
LESOTHO
Durban
SOUTH AFRICA
N
W
E
S
Cape Town
Port Elizabeth
© 2013 Encyclopædia Britannica, Inc.

Spatial Statistics - Status Quo

- Spatial statistics for wildlife conservation and protection
- Focus and objective: Counting and tracking
- Species: Elephants, red foxes, baboons and vesper sparrows
- Common spatial statistics approach: point patterns

Rhino

Facts about Rhino

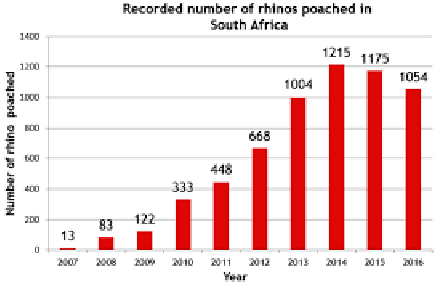
- Second largest land mammal (after elephants).
- White rhinos are grey
- Rhinoceros horns are made from keratin, the same substance that fingernails and hair are made of.
- Rhino have very poor eyesight.

Poaching Crisis

- Rhino horn trade value: $60,000 per kg (estimated)
- KNP is a transnational park (no borders)
- Poacher intrusion rate increase
- Methods more violent and cruel

Poaching Crisis

- Rhino horn trade value: $60,000 per kg (estimated)
- KNP is a transnational park (no borders)
- Poacher intrusion rate increase
- Methods more violent and cruel



Conservation Strategies

- Dehorn orphans
- Translocation
- Legalisation of trade (much debated)
- Education

Challenges

- Sparse datasets
  - Kruger National Park - $\pm 1900\text{km}^2$
  - Say 1200 poaching per year
- Operational challenges
  - Poachers enter by foot
  - Poacher stay in the park for days, surveilling
  - Corruption
  - KNP is a tourist destination - many civilians

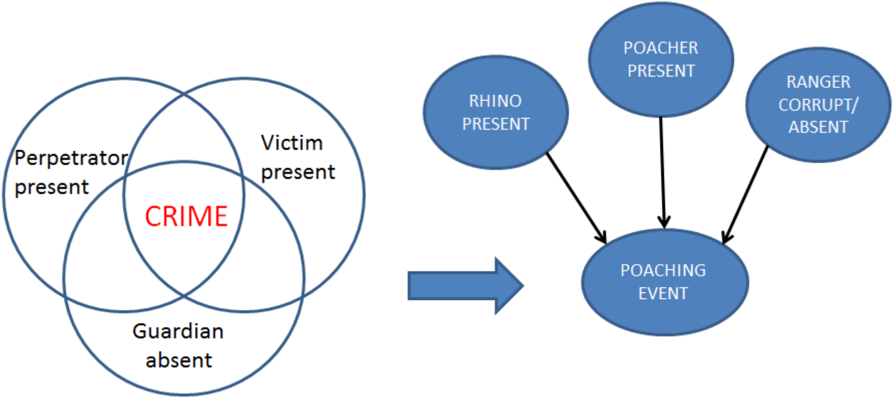
Bayesian networks in a geospatial space

- Include latent variables that experts understand
- Include context variables
- Train with expert knowledge and data
- Not only prediction, but also reasoning (what-if)
- Fuse prior knowledge, GIS info, and sensor data (soft and hard) at runtime

Rhino Anti-poaching Model

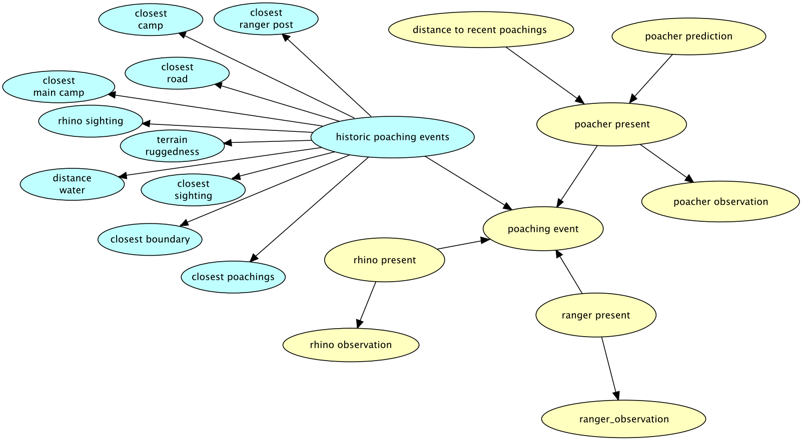
- Generate spatially discrete probability maps
- Application: Anti rhino poaching
- Maps should identify areas with a high poaching risk
- Output - probability heatmap
- Use predictions to optimise the use of available resources

Routine Activity Theory



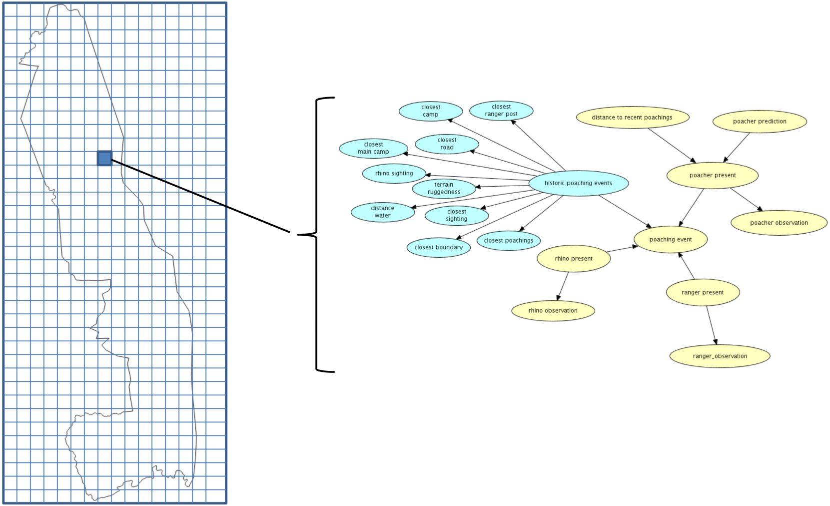
BN model consists of two portions:

- Causal portion
  - Combines the necessity to have a poacher and rhino present, and ranger absent.
  - Parameterised using domain experts.
- Classifier portion
  - Indication of historic vulnerability of an area
  - Parameterised using historic poaching data
  - Leaf nodes represents spatial attributes (covariates)
    - Water availability
    - Proximity to roads
    - Proximity to camps
  - Machine learning algorithm: Naive Bayes Classifier


closest camp
closest ranger post
distance to recent poachings
poacher prediction
closest road
closest main camp
rhino sighting
terrain ruggedness
historic poaching events
poacher present
distance water
closest sighting
poacher observation
poaching event
closest boundary
rhino present
closest poachings
ranger present
rhino observation
ranger_observation

Spatial Application of BN Model

- KNP are is subdivided in 1km^2 cells ($\pm$19000 square cells)
- For each cell
  - all relevant covariates are calculated
  - an instance of the model is created
  - a Bayesian inference process is executed
  - an output is generated - P(poaching)


closest camp
closest ranger post
distance to recent poachings
poacher prediction
closest main camp
closest road
rhino sighting
terrain ruggedness
historic poaching events
poacher present
distance water
closest sighting
poacher observation
closest boundary
closest poachings
poaching event
rhino present
ranger present
rhino observation
ranger_observation

Output: Probability of a poaching event

- Calculated for each cell
- Spatially discretised probability heatmap
- Probabilities are normalised over map

Output: Probability of a poaching event

- Calculated for each cell
- Spatially discretised probability heatmap
- Probabilities are normalised over map

Implementation

- Implemented in command and control system
- Currently operational
- Used for:
  - long term trend analysis
  - positioning of sensors

Evaluation Metrics

Challenges

- Typically $n << m$ - many more cells than samples
- Cells are sparsely populated with samples
  - Zero counts
  - Large degrees of freedom
  - Expected counts for all cells are typically $<< 1$, since $n << m$
- Comparative evaluation

Evaluation Preprocessing

Evaluation Preprocessing

- Spatially discrete maps can be 'unrolled' into a one dimensional probability mass function

Evaluation Preprocessing

- Spatially discrete maps can be ‘unrolled’ into a one dimensional probability mass function
- The **probability distribution** of model P_j is parameterised by probability values $P_j = [p_1, p_2, ..., p_m]$, where m is the number of cells contained in the map.

Evaluation Preprocessing

- Spatially discrete maps can be ‘unrolled’ into a one dimensional probability mass function
- The **probability distribution** of model P_j is parameterised by probability values $P_j = [p_1, p_2, ..., p_m]$, where m is the number of cells contained in the map.
- The **data vector** is given by $X = [x_1, x_2, ..., x_m]$, where x_i corresponds to the sample count in cell i, which in turn is parameterised by probability p_i.

Evaluation Preprocessing

- Spatially discrete maps can be ‘unrolled’ into a one dimensional probability mass function
- The **probability distribution** of model P_j is parameterised by probability values $P_j = [p_1, p_2, ..., p_m]$, where m is the number of cells contained in the map.
- The **data vector** is given by $X = [x_1, x_2, ..., x_m]$, where x_i corresponds to the sample count in cell i, which in turn is parameterised by probability p_i.
- The total number of samples is given by $n = \sum_{i=1}^{m} x_i$.

Evaluation Preprocessing

- Spatially discrete maps can be 'unrolled' into a one dimensional probability mass function
- The **probability distribution** of model P_j is parameterised by probability values $P_j = [p_1, p_2, ..., p_m]$, where m is the number of cells contained in the map.
- The **data vector** is given by $X = [x_1, x_2, ..., x_m]$, where x_i corresponds to the sample count in cell i, which in turn is parameterised by probability p_i.
- The total number of samples is given by $n = \sum_{i=1}^{m} x_i$.
- Multinomial distribution

Log-likelihood

- Compare log-likelihoods of the data given some model M_j parameterised by P_j
- First two terms are constant for evaluating different models (independent of P_j)
- The term $\sum_{i=1}^{m} x_i \log(p_i)$ causes a problem (next slide).

Definition

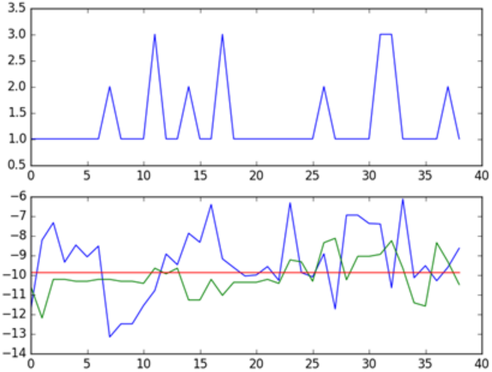
The log-likelihood of a multinomial distribution is given by:

$$ll(X|P_j) = \log(n!) - \sum_{i=1}^{m} \log(x_i!) + \sum_{i=1}^{m} x_i \log(p_i)$$

The problem with $\sum_{i=1}^{m} x_i \log(p_i)$

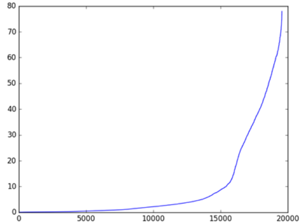
- Cells with low probabilities containing samples are penalised significantly, owing to the $\log(p_i)$ term.
- Cells with significant probabilities which do not have any samples are not penalised at all. They are omitted from the sum, owing to the cell count x_i being zero.
- Analog of *false alarm* (saying that a cell has high probability, but which doesn't reveice a sample) are not penalised

Log-likelihood (40 samples)

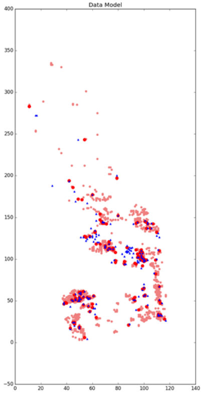


Distance metric

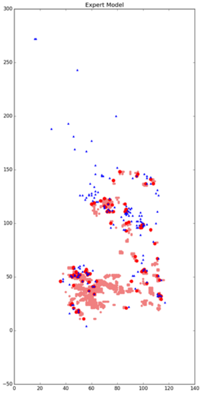
- For all cells, sort P(poaching)
- Select highest 5%
- Calculate distance between highest 5% probabilities and poaching events (in test set)
- For each poaching, choose shortest distance between highest 5% probabilities and poaching event



Distance metric - comparing two models



Average distance: 1.95



Average distance: 8.04

Future Work

- Historical poaching data
  - Alternative classification models (Logistic regression, Kriging, GMM)
  - Point process models
  - Feature selection
  - Optimal time frames to take into account changing patterns
  - Smaller areas (sub-sections)
  - Non-uniform cells
- Evaluation
  - Kolmagorov-Smirnoff Test
  - Evaluation of complete BN, not only classification portion