Impact of Uncertainty on the Digital Observatory of Protected Areas, including eHabitat

Jon Olav Skøien, Gregoire Dubois, Michael Schulz
Outline of the presentation

- Objective
- Sources of data
- Models
- Uncertainty of data
- An example application (eHabitat Web Processing Service)
Objective

Protected areas (PA)
- 130,000 and increasing
- eHabitat part of larger web portal (DOPA)
- Web services should be generic
- Uncertainty propagation
- What are the consequences of climate change?
  - Tolerance
  - Migration
  - Protection or prison
- Mahalanobis distance – for similarity between parks and surroundings
- Species density maps (biodiversity value of a park)
- Ranking of biodiversity value
- Pressure (population, agriculture)
- Alerts (deviations from seasonal averages of precipitation, small water bodies, vegetation index)
Mahalanobis distance

- Similarity measure
- Estimates of habitat boundaries - presence only data
- Layers of continuous variables as indicators
- Probability of finding similar habitats today or in the future
- Variance substitutes habitat tolerance

\[ D^2 = (X-\mu)\Sigma^{-1} (X-\mu) \]

- Tree coverage
- Elevation
- Population density
- Climatic data
- Others
eHabitat - web client

Protected Area

Nowcasting

bioT, annual precip., PET/Rainfall
eHabitat - web client

Protected Area

Forecasting

bioT, annual precip., PET/Rainfall
Sources of uncertainty

- Park boundaries
- Species observations
- Species distributions
- Input layers (satellite images, climate interpolations, climate forecast)

Others

Uncertainties?

- % tree cover
- % herbaceous cover
- % barren cover
- Elevation in metres
- Slope in degrees
- NDVI
- NDWI
- Aridity index
- % water body presence
Park boundaries

- How to quantify uncertainty?
- Fence or line on map
- Based on geographic feature
- Uncertainty based on trust
- Paper parks
Species observations

- Typically available from GBIF (Global Biodiversity Information Facility)
- Which coordinates are used?
- Significant digits of 16.166667?
- Classification of observations (dead/alive, free/zoo etc)
- Is the species still where it was observed?

2000  2008
Species ranges

- Based on observations
- Boundaries from habitat models (such as Mahalanobis)
- Modified by experts

⇒ Extremely difficult to assess uncertainty

⇒ Development of tool to help park managers report on species being observed/not observed
Products from satellite images
Uncertainty from

- Representativity of observed property (NDVI is “measured” without errors, but which species care about NDVI?)

- Classification / conversion (how to identify lakes, forests etc)
- Interpolated maps – such as current climate (worldclim.org)
- Interpolated with splines – no kriging error
- Using derived products, such as biotemperature and aridity
Uncertainty rasters

Kriging errors
- + takes nearness to stations into account
- - climate variables not stationary on a global scale

Averaged cross-validation errors
- + takes non-stationarity better into account than kriging errors
- - need large averaging blocks for some regions
- - Will not reflect nearness to measurement stations
Uncertainty of climate data

Have to guess uncertainty – ideally related to:

- station density
- elevation homogeneity
- distance to coast
- Resolution
- method for deriving other climate products
Need:

- Generic and automatic solution
- Uncertainty - based on fixed CV
  \[ \text{stdev}(x) = 0.05 \times \text{obs}(x) + \min(\text{obs} > 0)/10 \]
- Gives spatially non-homogeneous errors
- No correlation with observation locations
- Does not take spatial patterns into account
- Cross-correlogram from sample of raster data
- Individual uncertainty is not enough
- Interoperable communication with u-NetCDF and uncertML
- Realisations generated from multivariable geostatistical simulations

Demo:
http://ehabitat.jrc.ec.europa.eu/
Forecast climate

Based on large scale GCM
No uncertainty information but
- QUMP-runs from one model
- Different GCMs
- Different scenarios

What about:
- Downscaling techniques?
- Changes in patterns
For which parks do the climate models do a good job today?

- Is scaling a problem?
- Is the seasonal pattern represented?
- Are observed changes also predicted?
Demo
uncertainties in eHabitat WPS

Simulation No. 1/20

Request for 20 simulations
Demo
uncertainties in eHabitat WPS

Simulation No. 2/20
Demo

uncertainties in eHabitat WPS

Mean of Probabilities
Demo uncertainties in eHabitat WPS

Std. Dev. of Probabilities
Probability > 50% similarity
Conclusions

- Data quality mainly missing
- Easier to create for some data than for others
- Simple variability or variograms of errors not always enough, also cross-correlations and/or variograms

http://www.uncertweb.org/

jon.skoien@jrc.ec.europa.eu

Funded by the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreements n° [226487 and 248488]