GENETIC ALGORITHMS TO OPTIMISE SPECIES DISTRIBUTION MODELS

useful for freshwater management

Sacha Gobeyn & Peter L.M. Goethals (09/05/2017, ISESS2017)
INTRODUCTION

– Freshwater management

(SOER 2015 — The European environment February 2015)
INTRODUCTION

– Freshwater management
SPECIES DISTRIBUTION MODELS

- Theoretical concepts
- Expert knowledge
- Data
  - Machine learning
  - Statistical approaches
  - Metaheuristics

Step 1: Model concept
Step 2: Data acquisition
Step 3: Model construction
Step 4: Model identification

Gobeyn et al. (2017)
SPECIES DISTRIBUTION MODELS

\[ HSI = f(SI_{X1}, SI_{X2}, \ldots, SI_{Xn}) \]

- \( Y = \text{species} \)
- \( X = \text{input variable} \)
- \( a_Y = \text{parameters of species response curve for input } X \)

Poff (1997)
MODEL OPTIMISATION

**GIVEN** data

**FIND**

- Input variables
- Species response curve parameters

**WITH**

- Optimisation algorithm

**CONDITIONS**

- Search a large unconstrained space
- Efficiently scan many possible distributions
- Preferably in an ensemble setting

**EVEN MORE**

- Automated
- Facilitate repeated analysis (on HPC)
- Open source
MODEL OPTIMISATION

Machine learning (SVM, Decision Trees, ..)
Metaheuristic (GARP)
Statistical (GLM, GAM, ..)
Other (MAXENT, Fuzzy Logic, ..)

Implement metaheuristic
Genetic Algorithm
GENETIC ALGORITHM

input variables

parameters

Chromosome

\[
\begin{align*}
0 & \quad 1 & \quad 0 & \quad 0 & \quad 0 & \quad 1 & \quad 0 & \quad 0 \\
1 & \quad 0 & \quad 1 & \quad 0 & \quad 1 & \quad 1 & \quad & \quad & \quad \\
0 & \quad 0 & \quad 0 & \quad 0 & \quad 0 & \quad 0 & \quad 1 & \quad & \quad \\
\end{align*}
\]

Mapper

\[
\begin{align*}
5 & \quad x_2 & \quad 6 \\
0 & \quad x_6 & \quad 4 \\
\end{align*}
\]

\[
\begin{align*}
a_{x_2,2} &= a_{x_2,1} + r \frac{5}{7} \\
a_{x_6,2} &= a_{x_6,1} + r \frac{0}{7} \\
a_{x_2,3} &= a_{x_2,1} + r \frac{6}{7} \\
a_{x_6,3} &= a_{x_6,1} + r \frac{4}{7} \\
with \quad r &= a_{x_2,4} - a_{x_2,1} \\
with \quad r &= a_{x_6,4} - a_{x_6,1}
\end{align*}
\]
GENETIC ALGORITHM

  100 chromosomes, mutation rate = 5%, crossover rate = 100%
GENETIC ALGORITHM

- Uncertainties!

*cloeon dipterum*

![Graphs showing pH vs. SI for 1% and 5% mutation rates.](image-url)
GENETIC ALGORITHM

- Hyper parameters
- Uncertainties!
- Encoding
  ➡️ Continuous versus binary encoding
- Objective function
  ➡️ Single or multi-objective optimisation
- How hard should a problem be for algorithm to fail?
  ➡️ Virtual approach
WHAT SOFTWARE IS AVAILABLE?

GitHub

https://sachagobeyn.github.io/SDMIT/

- Input variable selection
- Run on HPC
- Uncertainty analysis!
Sacha Gobeyn
Research Engineer - Model developer for environmental management // PhD Student

Department of Applied Ecology and Environmental Biology

E  sacha.gobeyn@ugent.be
T  +32 9 264 38 96
M  +32 4 244 44 10

www.ugent.be
REFERENCES


