A statistical methodology to downscale aggregated land use data

Nicolas Dendoncker [a]*, *Patrick Bogaert* [b], *Mark Rounsevell* [a]

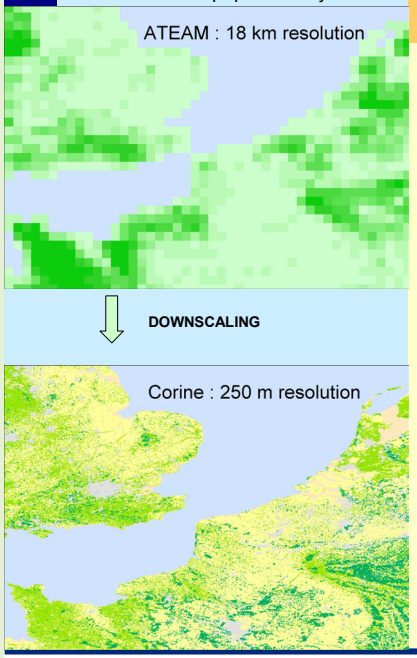
[a] Dept. of Geography, Place Pasteur 3, B-1348 Louvain-la-Neuve. Belgium. www.geo.ucl.ac.be

[b] Dept. of environmental sciences and land use planning - Environmetry and Geomatics, Croix du Sud 2, BP 16, B-1348 Louvain-la-Neuve. Belgium.

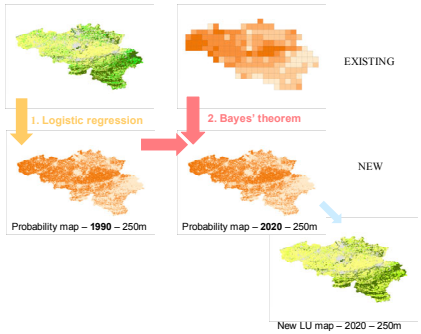
Downscaling – why?

Introduction

- Scenario and past Land Use (LU) data often at coarse resolution - reflecting global scale processes acting over large areas
- Fine resolution LU data is needed for ecological studies (e.g. population dynamics, biodiversity, carbon stocks…)
- In particular, there is a need for a fine resolution LU model over a large extent (the Svalbard migratory geese flyway) to model bird population dynamics

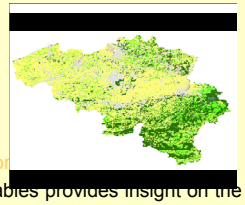


Downscaling – how?



1. Studying present LU patterns to derive probability maps of LU presence using statistical techniques

- Binomial and multinomial logistic regressions studying the presence/absence of each LU within each cell based on environmental variables **and** neighbourhood variables



Main conclusion from logistic regressio

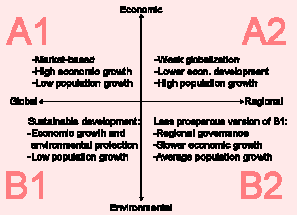
While the study of environmental variables provides insight on the factors that have led to the current LU patterns in Belgium, the best statistical fit is achieved with a purely autoregressive model (i.e. including only neighbourhood variables), for all LU studied, at the resolution of CORINE.

Consequently:

- Probability maps of land use presence can be derived using one LU dataset only (i.e. no ancillary data needed)
- A multinomial logit model can be applied
- This procedure is replicable over a wide area
- The resulting probability maps can serve as a baseline to downscale aggregated LU data

2. A method to update these probability maps based on Bayes' theorem and existing coarse scale LU scenario data

The ATEAM LUC scenario dataset: 4 scenarios based on the IPCC-SRES Framework – 18km resolution – 3 time steps (2020-2050-2080)



The multinomial logit model giving initial probability maps:

$p(c_i|N_v) = \exp(u_i N_v) / \sum \exp(u_i N_v) \quad i = 1,\ldots,k$

Where:

- $p(c_i|N_v)$ is the conditional probability that a cell v would take LU i among the set of k possible LU classes and given N_v
- N_v is the neighbourhood variable

From Bayes' theorem – the updating rule:

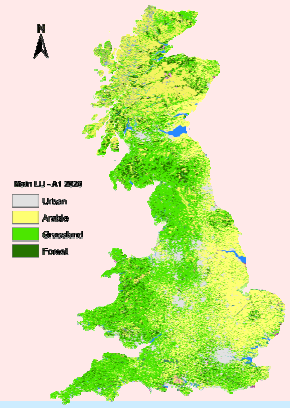
$p'(c_i|N_v) \propto p(c_i|N_v)\, p'(c_i)/p(c_i)$

Updated conditional probabilities at time t' | Initial (time t) marginal probabilities i.e. given by the CORINE dataset

Marginal probabilities at time t' i.e. given by the ATEAM LU frequencies

Results:

- For each cell LU with highest conditional probability is represented
- No border effects (between ATEAM cells) can be seen
- LU frequencies given by ATEAM are respected
- LU patterns are correctly represented despite increased clustering as a consequence of using neighbourhood variables



Conclusion

The statistical downscaling procedure is efficient and simple to implement (routine programmed in MATLAB):

- Only one baseline LU dataset is needed (i.e. Corine)
- Set to downscale the ATEAM scenarios over the whole of Europe15
- It could also downscale other aggregated LU data e.g. past statistics at the NUTS level

Research project funded by the European commission Directorate-General for Research
Contract No: EVK2-2001-00235

UCL Université catholique de Louvain
Université membre de l'Académie universitaire 'Louvain'