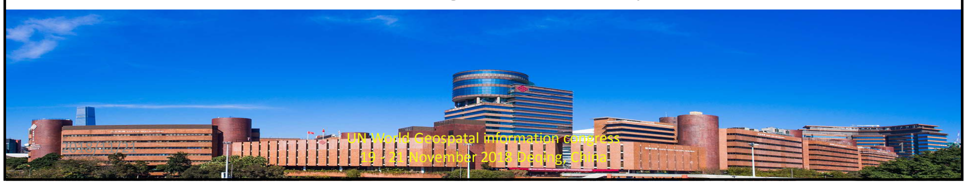
THE HONG KONG POLYTECHNIC UNIVERSITY
香港理工大學
Opening Minds • Shaping the Future

Methodology for Aggregation of Land Cover Data

LI, Zhilin (PhD, DSc)
The Hong Kong Polytechnic University

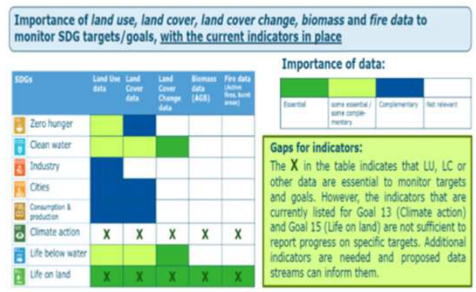
TAN Shiteng and XU Zhu
Southest Jiaotong University, China



Contents

Land Cover Data Required for SDGs

- SDGs: covering all aspects of the nature and society
  - Biosphere issues: 4
  - Society issues: 8
  - Economic issues: 4
  - Others: 1
- Relevance of Land cover data to SDGs
  - Essential X
  - Some essential / some complementary
  - Complementary
  - Not relevant

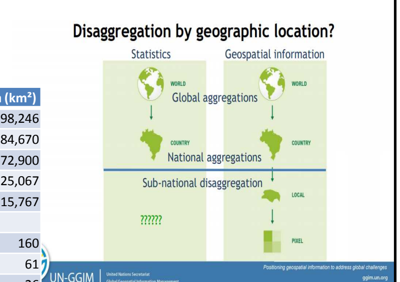


land cover data for measuring SDGs (Romijn et al 2008, *http://www.gofcgold.wur.nl/documents/newsletter/Sustainable_Development_Goals-infobrief.pdf*)

Need of land cover data at different resolutions/scales

- To assist disaggregation of statistical data
  - *"SDG indicators should be* ***disaggregated****,* where relevant, ***by*** income, sex, age, race, ethnicity, migratory status, disability and ***geographical location*** (UN IAEG-SDGs)
- National level➔ sub-national➔ …➔local➔ …➔ pixel
- Pixel size for a country/region appropriate to its physical size

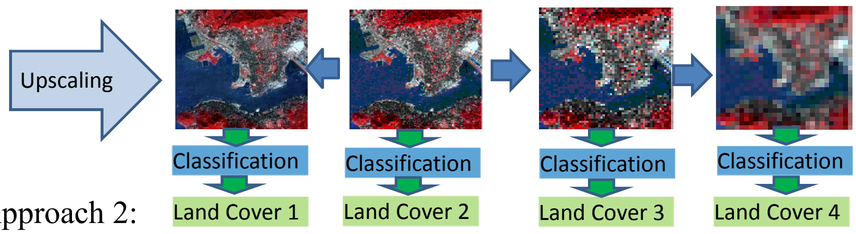
Pos	Country	Area (km²)
1	Russia	17,098,246
2	Canada	9,984,670
3	China	9,572,900
4	US	9,525,067
5	Brazil	8,515,767
191	Liechtenstein	160
192	San Marino	61
193	Tuvalu	26
194	Nauru	21
195	Monaco	2



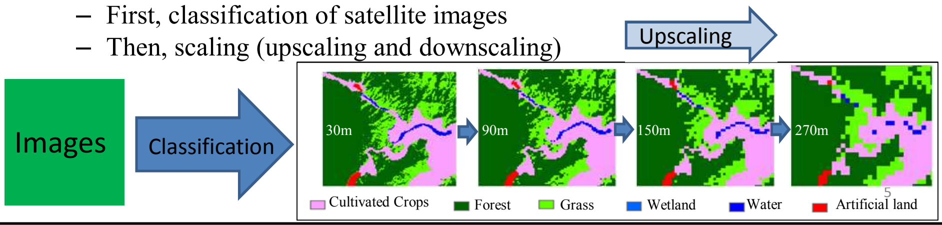
(UN-GGIM)

How to obtain land cover data at different resolutions/scales

- Approach 1:
  - First, scaling (upscaling and downscaling) of satellite images
  - Then, classification



- Approach 2:
  - First, classification of satellite images
  - Then, scaling (upscaling and downscaling)



Contents

- Why upscaling (aggregation) of land cover data: An introduction
- Upscaling of land cover data indirectly via upscaling of remote sensing image data
- Upscaling of land cover data directly via aggregation
- Conclusions and outlooks

Scaling of remote sensing images

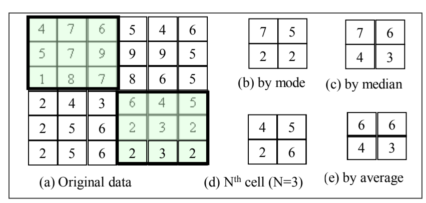
- Upscaling: fine ➔ coarse
  - e.g. 1m➔2m➔5m ➔10m➔25m
  - A lot of work done
- When needed
  - no images with required resolution available
  - Available but we don't want to spend more money

- Downscaling: coarser to fine
  - e.g. 300m➔200m➔100m ➔50m➔10m
  - Recent efforts
- Why needed?
  - because we have high resolution images (e.g. 0.5m) already?
  - Missing parts of higher images
    - e.g. Cloud

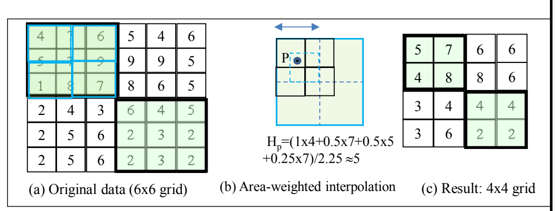
(images downloaded from Google Map)

Upscaling of remote sensing images

- Aggregation
  - Multiples of original resolution, e.g. 3x3 ➔1x1
  - No interpolation required
  - By "mode", "median", "average" and N^{th} cell
- Resampling
  - Not multiples of original resolution, e.g. 3x3➔2x2
  - Interpolation required
    - Nearest Neighbour
    - Bilinear interpolation
    - Bicubic interpolation



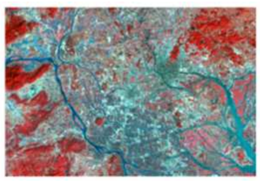
"3x3 to 1x1" aggregation of image data



"3x3 to 2x2" resampling of image data

Comparative study of image upscaling techniques

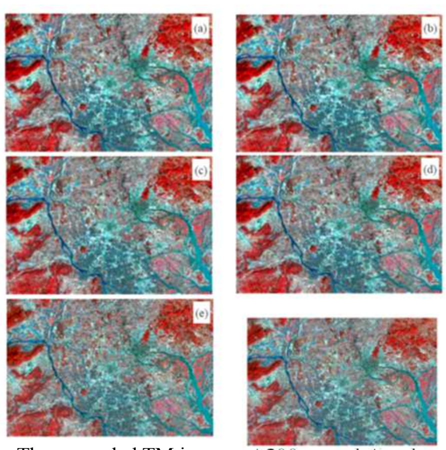
- Arithmetic Average Variability-Weighted (AAVW),
- Averaging (AV),
- Bilinear (BL),
- Bicubic (BC), and
- Nearest neighbor (NN)



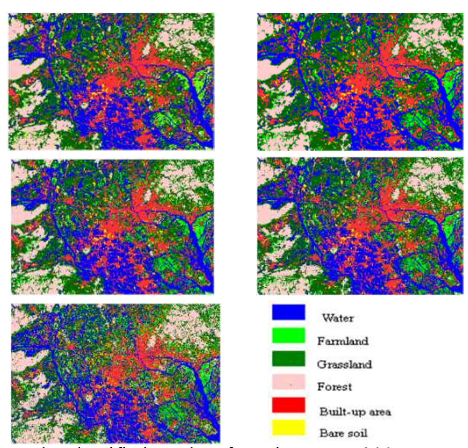
- TM image for test
- SPOT image
- Scaling: 2×2 , 3×3, 4×4 , 5×5, 6×6 , 7×7 , 8×8 , 9×9 , 10×10
- TM: 30→60m, 90m, …, 300m
- Classification and accuracy assessed
  - Overall accuracy
  - Class level accuracy

(Han et al., 2009, http://www.docin.com/p-1447444919.html)

Upscaled images: results and classification



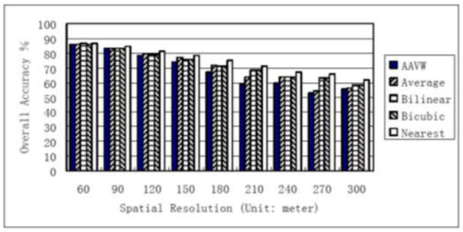
The up-scaled TM images at 300m resolution by different aggregation methods: (a) by AAVW, (b) by AV, (c) by BL, (d) by BC, (e) by NN



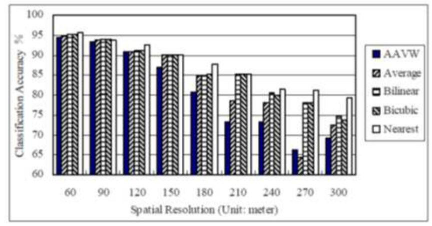
The classified results of TM images at 300m aggregated by 5 methods. (a) by AAVW, (b) by AV, (c) by BL, (d) by BC, (e) by NN

(Han et al., 2009, http://www.docin.com/p-1447444919.html)

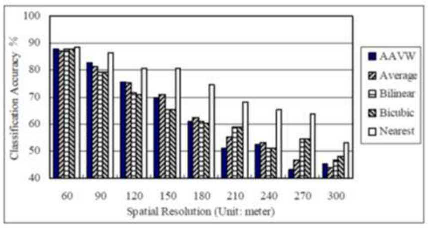
Effect of upscaling techniques on classification results



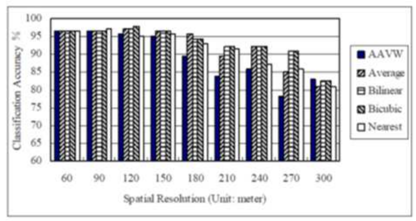
Effect of aggregation on **overall accuracy**



Effect of aggregation on water accuracy



Effect of aggregation on forest accuracy



Effect of aggregation on built-up-area accuracy

(Han et al., 2009, http://www.docin.com/p-1447444919.html)

Contents

Scaling of land cover data

- Upscaling:
  - fine ➔ coarse
  - e.g. 30m➔50m➔6m ➔100m➔200m ➔250m
- Downscaling:
  - coarser to fine
  - e.g. 300m➔200m ➔100m ➔50m

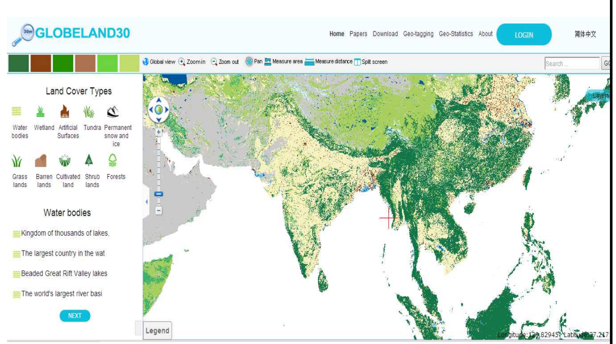
What about a resolution between 30m and 500m?

- Resolutions of current land cover datasets
  - Globeland30: 30m
  - European GlobCover: 300m
  - MODIS12C1: 500m
  - UMD: 1km

	Land cover	Provider	Reso-lution	Classes
1	IGBP-DISCover	USGS	1km	vegetation
2	UMD	University of Maryland	1km	Multiple
3	MODIS 500m	University of Boston	500m	Multiple
4	GLC2000	European Joint Research Center	30m	Multiple
5	NLCD (US) 30m	USGS	30m	Multiple
6	**Globeland30**	**National Geomatics Center of China**	**30m**	**Multiple**

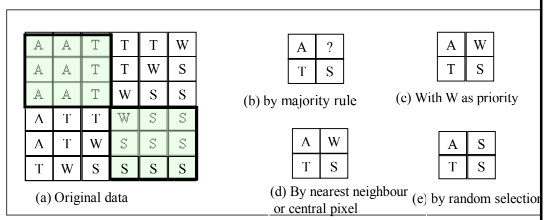
Globeland30: high-resolution land cover data

- 10 classes
- 30m resolution
- Global coverage
- Two epoches
  - 2000, 2010
- Accuracy:
  - Over 85%
  - By international assessment
- http://www.globallandcover.com/home/Enbackground.aspx



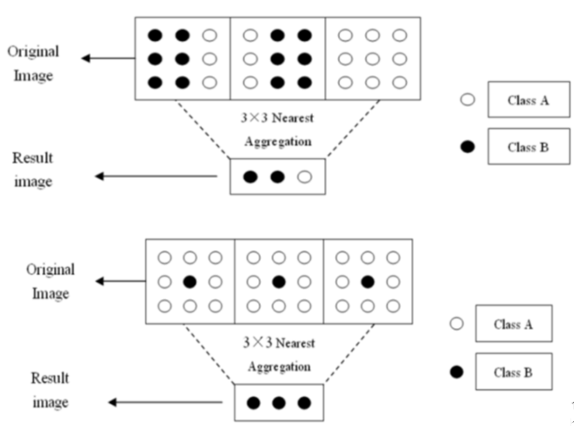
Upscaling of land cover data: Aggregation with classic techniques

- By majority rule
- By nearest neighbour (or central pixel)
- By random selection
- By some priority rules
  - based on global structural information
  - Based on local structural information

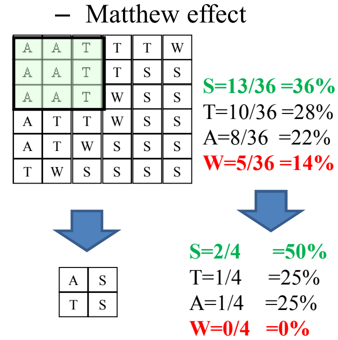


Methods for aggregation of land cover data - more

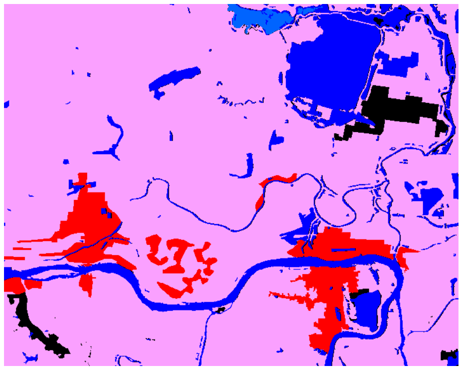
- Problems with simple solutions:
  - Nearest Neighbour (NN)



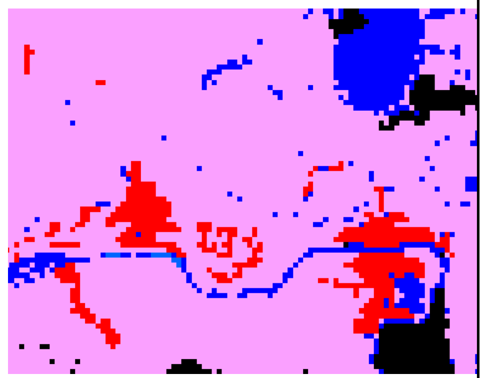
- Problems with simple solutions:
  - Mode (majority)
  - Matthew effect



Results from simple aggregation



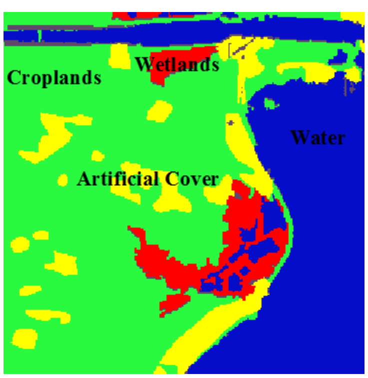
30m

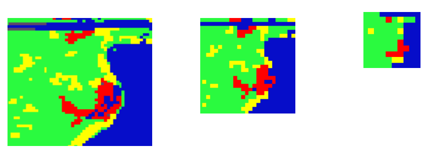


250m

Some rivers are broken

Aggregation of Globeland30 by Majority





- 1x1➔5x5➔7x7➔15*15
- By majority
- **Matthew effect created**
  - Artificial cover reduced significantly
  - Cropland increased significantly

2018/11/21

UN World Geospatal information congress
19 - 21 November 2018 Deqing, China

18

Measures for aggregation effects

Landscape pattern indices

- PLAND: % of total landscape area
- PAFRAC: The difference of perimeter area fractal dimension
- AI: Aggregation index
- LSI: landscape shape index

(1). PLAND is expressed as follows:

$$PLAND = p_i = \frac{\sum_{j=1}^{n} a_{ij}}{A}(100) \quad (1)$$

While, a_{ij} is area of patch ij, A is total area of landscape.

(2). PAFRAC is expressed as follows:

$$PAFRAC = \frac{2}{\left[\frac{n_i \sum_{j=1}^{n}(\ln p_{ij} \cdot \ln a_{ij}) - \sum_{j=1}^{n}\ln p_{ij}\sum_{j=1}^{n}\ln a_{ij}}{(n_i \sum_{j=1}^{n}\ln p_{ij}^2) - (\sum_{j=1}^{n}\ln p_{ij})^2}\right]} \quad (2)$$

While, a_{ij} is area of patch ij, p_{ij} is perimeter of patch ij, n_i is number of patch.

(3). AI is represented as follow:

$$AI = \left[\frac{g_{ii}}{\max \rightarrow g_{ii}}\right](100) \quad (3)$$

While, g_{ii} equals the number of like adjacencies between pixels of patch type i based on the single-count method. $\max \rightarrow g_{ii}$ is the maximum of g_{ii}. If A_i is the area of class i and n is the side of a largest integer square smaller than A_i, and m = A_i -n2, then the largest number of shared edges for class i, $\max \rightarrow g_{ii}$ will take one of the three forms:

$\max \rightarrow g_{ii}$ = 2n (n-1), when m = 0, or

$\max \rightarrow g_{ii}$ = 2n (n-1) + 2m -1, when m = n, or

$\max \rightarrow g_{ii}$ = 2n (n-1) + 2m -2, when m > n.

(4). LSI is expressed as follows:

$$LSI = \frac{e_i}{\min e_i} \quad (4)$$

e_i equals the total length of edge of class i in terms of number of cell surfaces. It includes all landscape boundary and background edge segments involving class i.

Effect of aggregation on land cover distribution: experiment

- Globland30
- 30m (1x1) ➔ 990 (33x33)
- By majority and random rules
- Aggregation index as measure

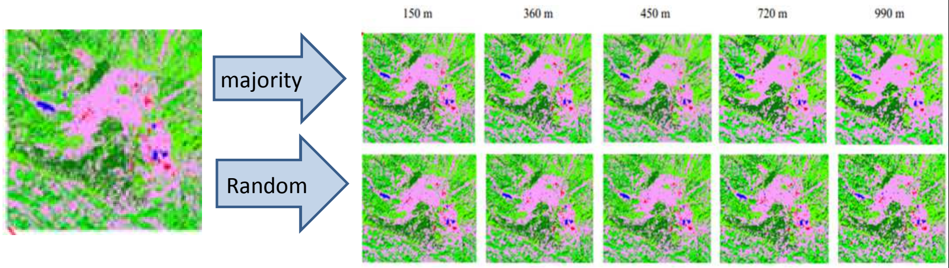
$$AI = \left[\frac{g_{ii}}{\max \rightarrow g_{ii}}\right](100)$$

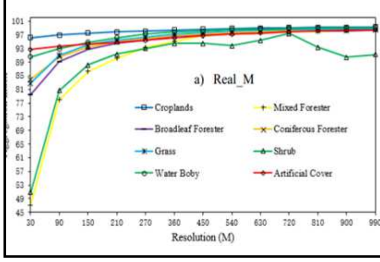
- *AI=1* ➔ highest level of aggregation, i.e. comprised of pixels sharing the most possible edges.
- *AI* =0 ➔completely disaggregated (lowest aggregation)

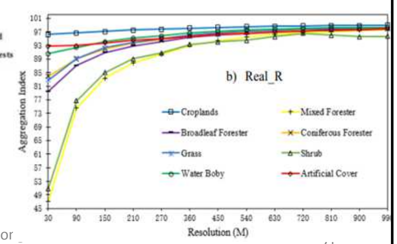
- g_{ii} = number of like adjacencies (joins) between pixels of patch type (class) i based on the *single-count* method.
- max-g_{ii} = maximum number of like adjacencies (joins) between pixels of patch type (class) i (see below) based on the *single-count* method.

(http://www.umass.edu/landeco/research/fragstats/documents/Metrics/Contagion%20-%20Interspersion%20Metrics/Metrics/C116%20-%20AI.htm)

Effect of aggregation on land cover patterns: experiment with Globeland30







orld Geospatial information cor
21 November 2018 Deqing, China

21

Effect of aggregation on land cover patterns: experiment with Globeland30-AI

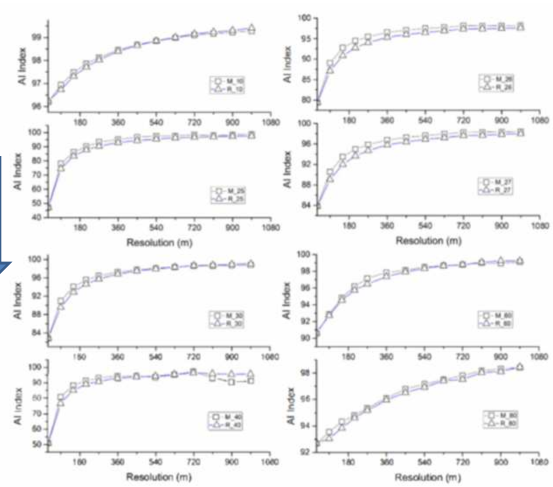


Figure 2. AI among majority rule (black line) and random rule (blue line) of varies resolution (30-m to 990-m). 8 subfigures represent the measure values for a land cover class with two aggregation methods separately.

- Artificial cover
- Broadleaf forest
- Coniferous forest
- Cropland
- Grass land
- Mixed forest
- Shrub
- water

22

Effect of aggregation on land cover patterns: experiment with Globeland30-PLAND

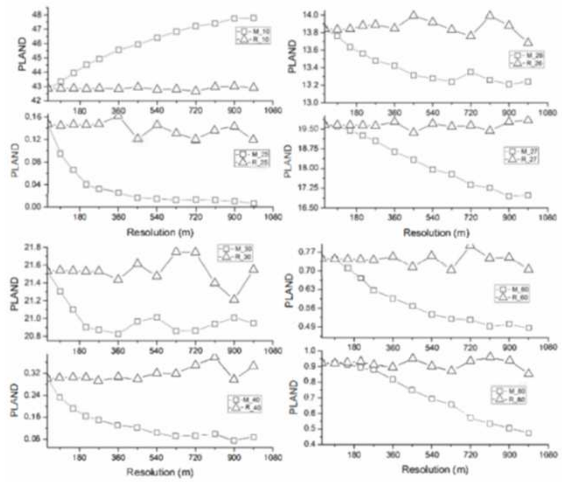


Figure 5. PLAND among majority rule (black line) and random rule (blue line) of varies resolution (30-m to 990-m). 8 subfigures represent the measure values for a land cover class with two aggregation methods separately.

- Artificial cover
- Broadleaf forest
- Coniferous forest
- Cropland
- Grass land
- Mixed forest
- Shrub
- water

Analysis of experimental results

- **Both aggregation approaches**
  - cause distortions of cover type proportions and spatial patterns.
- **Major-rule (M_rule)** :
  - filters out minor patches so as to obtain more clumped landscapes
  - **Maintains spatial pattern better**
- **Random-rule (R_rule)**:
  - **maintains cover type proportions better**, but
  - tends to make spatial patterns change toward disaggregation.

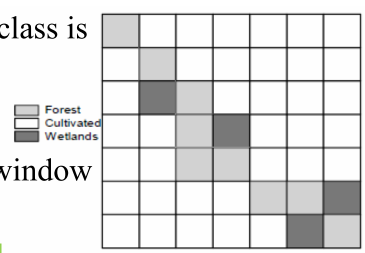
Ideas arising from experimental results

- Take care of spatial structure
  - Local structure
  - Global structure
- Two corresponding techniques
  - Markov random field
  - Spatial scan statistic

- Markov random field
  - keep the pattern similarity between two scale
  - preserve the spatial continuity
- Spatial scan statistic
  - preserving heterogeneity and information from rare classes
  - Consideration of global percentage of each class

Spatial scan statistic

- 7x7 ➔ 1x1
- What to be assigned to the new pixel?
  - Nearest=wetland
  - Majority=cultivated
- Which is most likely according to the known global percentages
- Suppose the global percentage for each class is
  - Cultivated = 72.13%
  - Wetlands =5.93%
  - Forested =7.93%



- Considering both the percentage in the window and global, the likelihood ratio
  - Cultivated = 0.022
  - Wetlands =0.197
  - **Forested =0.836.**

See Coulston 2004 for mathematical models

Coulston, 2004, The spatial scan statistic: A new method for spatial aggregation of categorical raster maps
https://arxiv.org/ftp/arxiv/papers/1408/1408.0164.pdf

Markov random field (MRF)

- MRF is a graphical model of probability distribution over random variables.
- It provides a convenient and consistent way to represent spatial dependency among random variables
- With RMF, two aspects can be taken into consideration
  - Similarity before aggregation
  - Spatial correlation during aggregation
- Procedure
  - Represent the Globeland30 by a 2-D MarkovRandom Field;
  - Built an energy function over the pixel class proportions and the neighbor pixels' contributions;
  - Determine the final pixel class at coarse resolution through the comparison of energy value for each class.

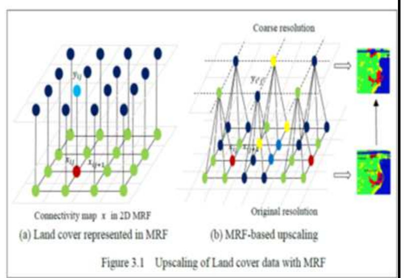
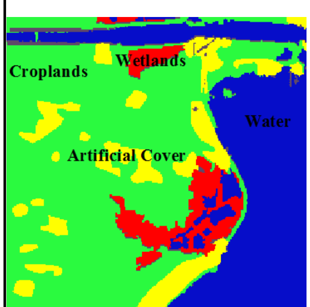
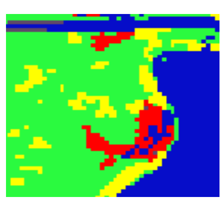
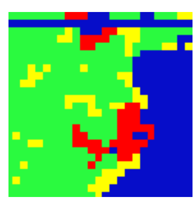
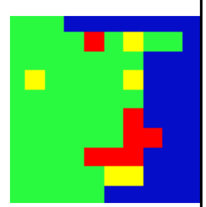


Figure 3.1 Upscaling of Land cover data with MRF

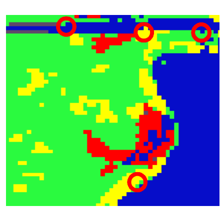
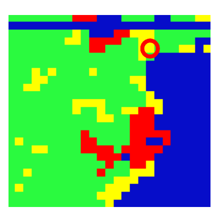
Aggregation of Globeland30: Majority vs MRF



1x1

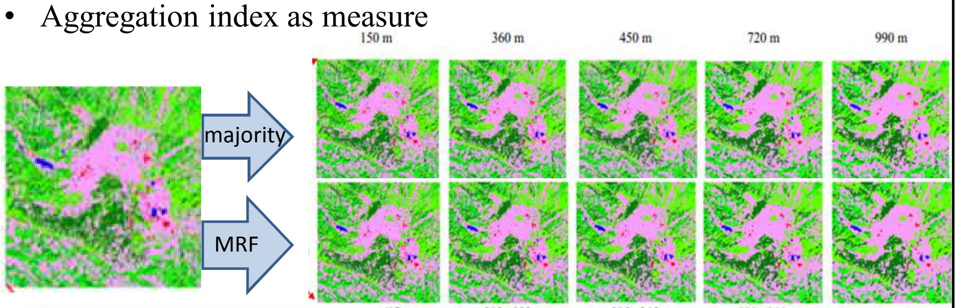




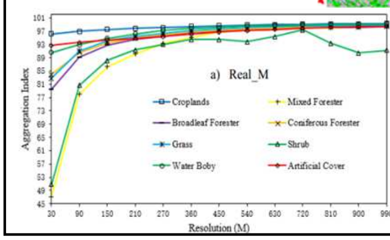
5x5 7x7 15*15

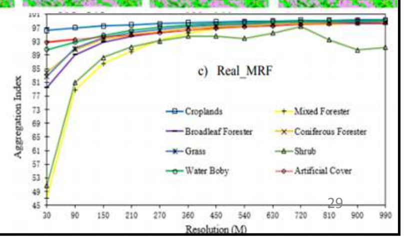

Evaluation of aggregation of Globeland30: Majority vs MRF

- 30m (1x1) → 990 (33x33)
- Aggregation index as measure

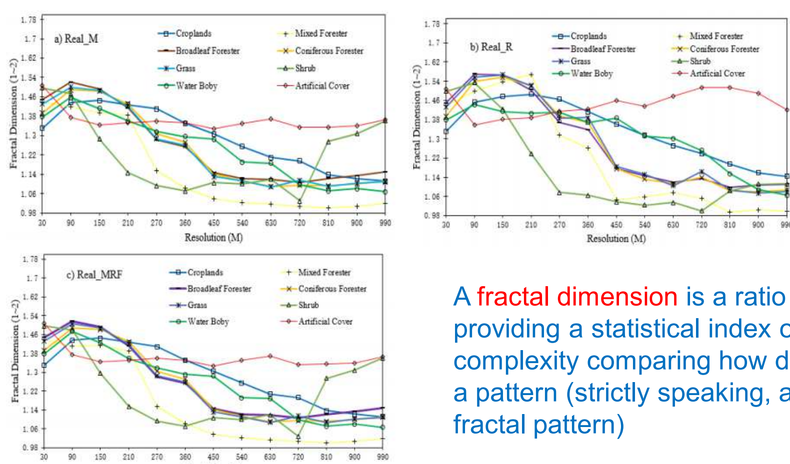




- Left:
  - majority
- Right:
  - MRF



Evaluation of aggregation of Globeland30: Random, Majority & MRF



A fractal dimension is a ratio providing a statistical index of complexity comparing how detail in a pattern (strictly speaking, a fractal pattern)

Figure 3. Fractal dimensions (FD) measured using corresponding map resolutions for the Land Cover image aggregated from 30 to 990m under the scenarios of (a) the real landscape with the majority aggregation (Real_M), (b) the real landscape with the random aggregation (Real_R), and (c) the real landscape with MRF aggregation.

Contents

Conclusions

- All aggregation techniques caused distortions
  - cover type proportions
  - spatial patterns
  - Continuity
- Which performs better?
  - Spatial pattern: M-rule better than R_rule;
  - Type proportion: R_rule better than other two;
  - Pattern and spatial continuity: MRF better than other two

Outlooks

- Downscaling of remote sensing images
  - Block-to-point Kriging interpolation
  - Super-resolution mapping
  - Downscaling via spar
- Downscaling of land cover data
  - Still need of preserving different cover type and different properties
  - Why not making use of all existing land cover data at different resolutions
  - se representation with double dictionaries
- Quality of aggregation and disaggregation vs reliability of SDG indicators computed

End of presentation

Supported by NSFC and HK Polytechnic University