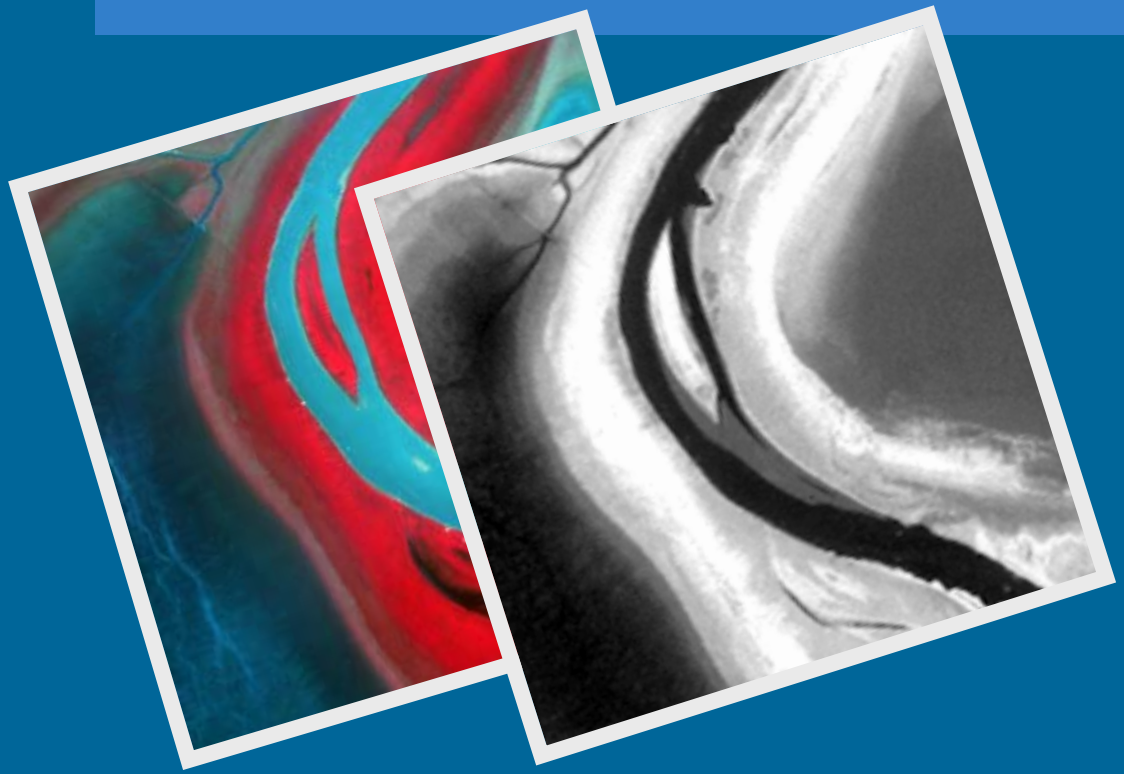
Assessment of extensional uncertainty modeled by random sets on segmented objects from remote sensing images



Xi Zhao [ab]
Xiaoling Chen [a]
Liqiao Tian [a]
Tiejun Wang [b]
Alfred Stein [b]

[a] Wuhan University, China
[b] ITC, University of Twente, the Netherlands

Outline

- Introduction
  - Uncertainty, Accuracy assessment, objective
- Method
  - Study area, data, field work, random set model
- Results
  - Field work measures, random set modeling results, accuracy report
- Conclusion and discussion

Introduction Uncertainty

- Modeling spatial objects with indeterminate boundaries

- uncertainty theories
  - probabilistic theory
  - fuzzy set theory
  - rough set theory
  - random set theory

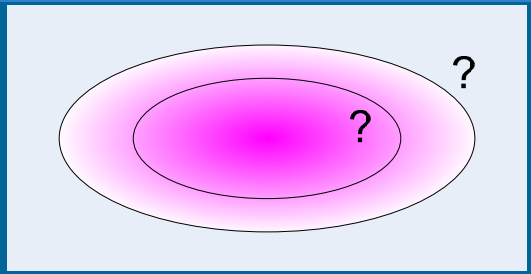
- application
  - beach
  - wetland
  - fire spread
  - flooding

Introduction Accuracy Assessment

- fuzzy confusion matrix
- even on the ground, the delineation of uncertain objects like a city may be impossible
- represented uncertainty does not always have corresponding objects in the field
- lack of detailed reference data

Introduction Objectives

- Extensional uncertainty



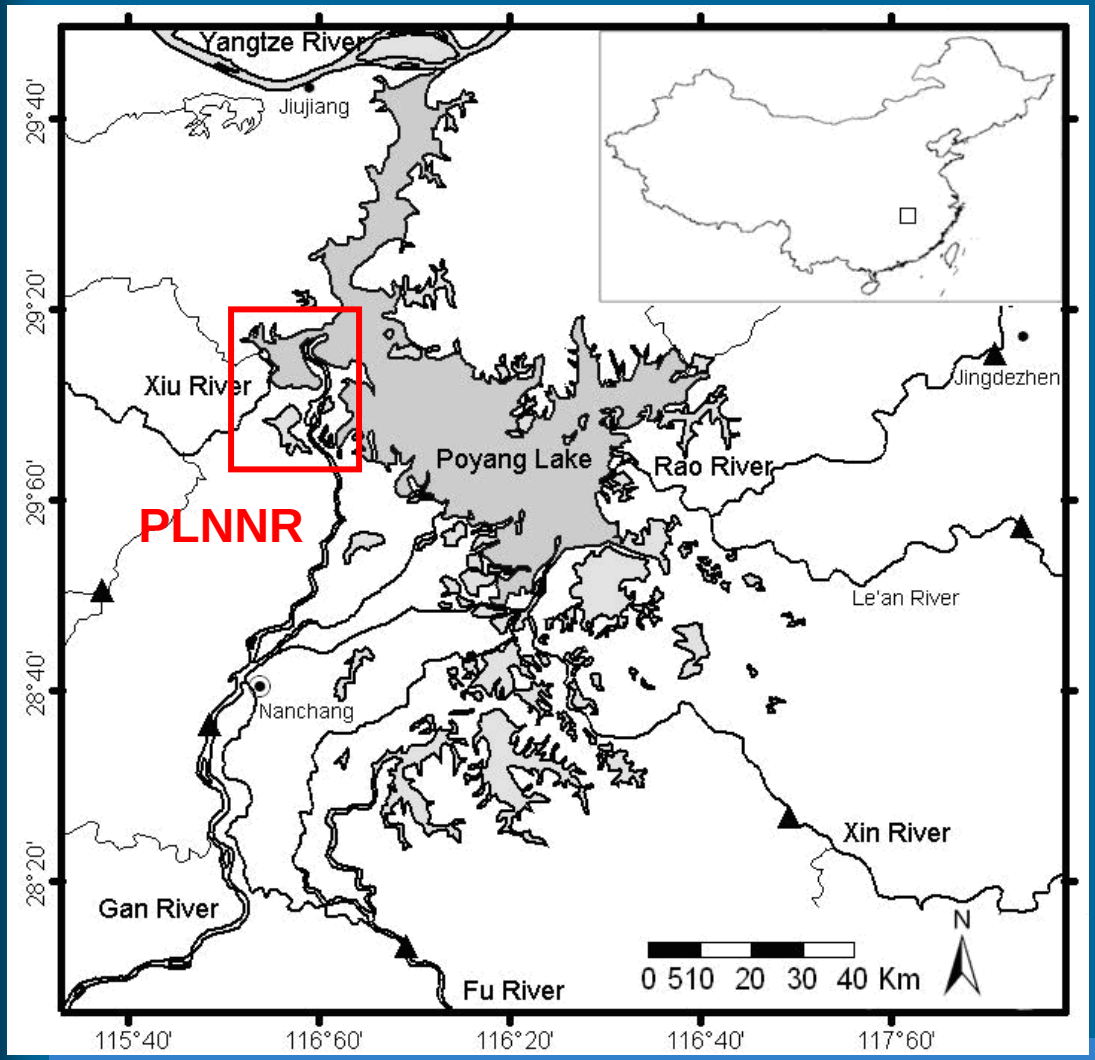
- (1) to explore the corresponding measurable variables collected on the ground for validating the uncertain image objects modeled by random sets

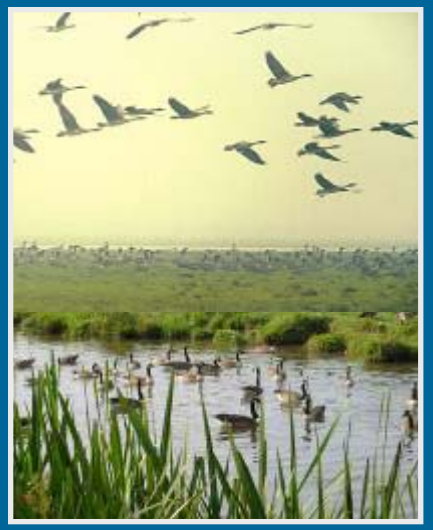
- (2) to quantify the quality of the random set modeling results.

Outline

- Introduction
  - Uncertainty, Accuracy assessment, objective
- Method
  - Study area, data, random set model, field work, accuracy assessment
- Results
  - Field work measures, random set modeling results, accuracy report
- Conclusion and discussion

Method Study Area_Poyang lake



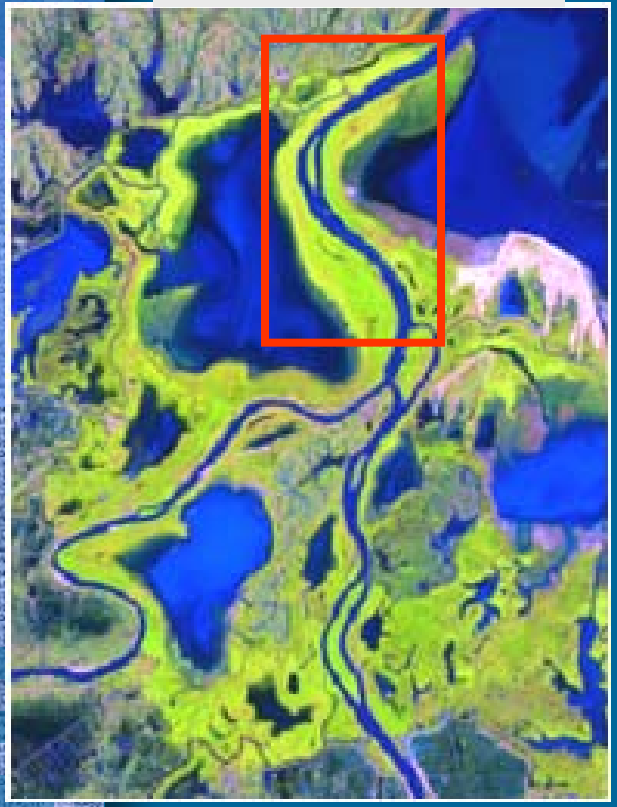


Method Data_HJ Satellite image

- HJ-1A/B satellites, launched on Sept 6, 2008 from China
- The constellation of the two satellites can form the multi-spectrum image with a 30m resolution of any location every 2 days
- HJ-1A image on November 24, 2009 downloaded from the China Centre for Resource Satellite Data and Applications (CRESDA)

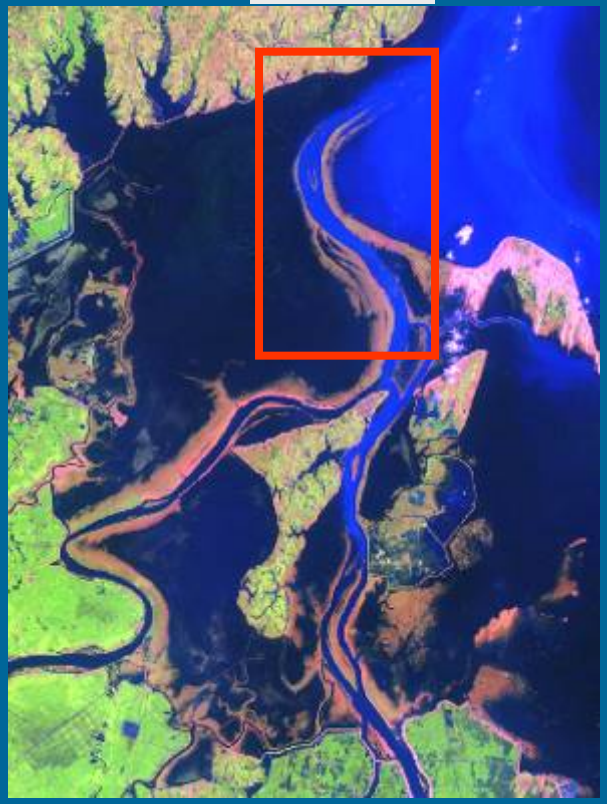
Method Grassland in PLNNR

Grass & sedge



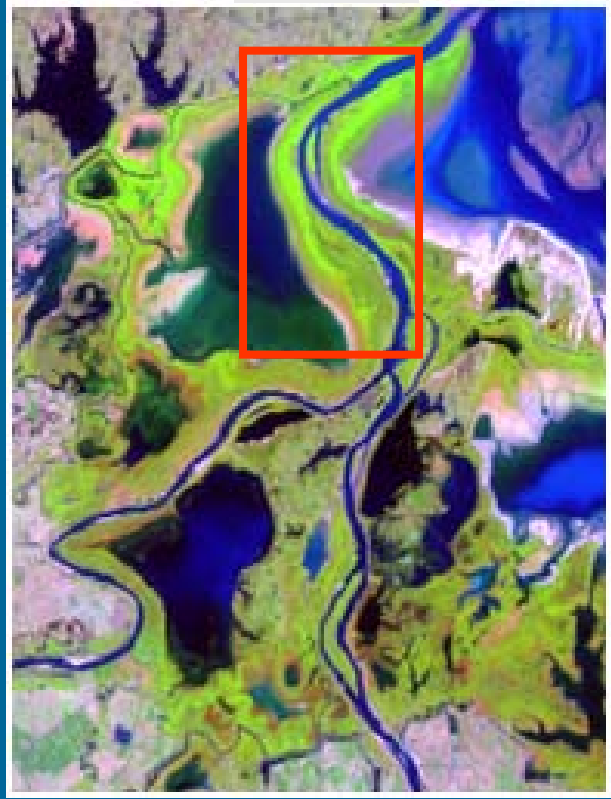
2004-05-05
Spring

Grass



2004-08-09
Flooding

sedge



2004-10-28
Autumn

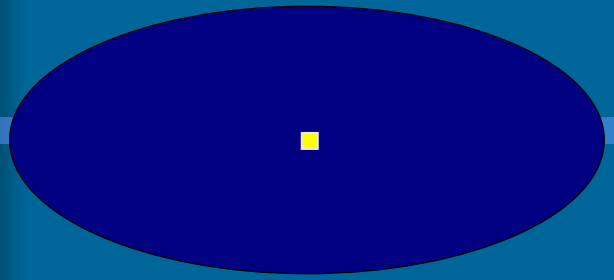
Method Random Set Theory

- Based on probability theory
- study randomly varying populations and geometrical shapes

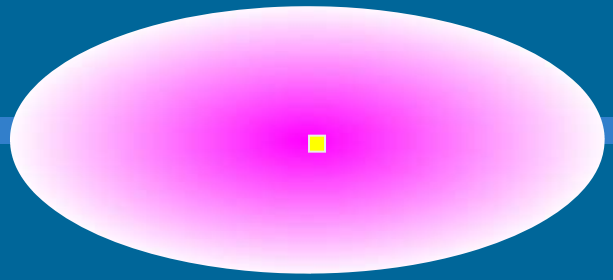
- A random set is a general random variable whose elements are sets
- provides a sound set-theoretic statistical exploratory of set-valued observations

DEMO

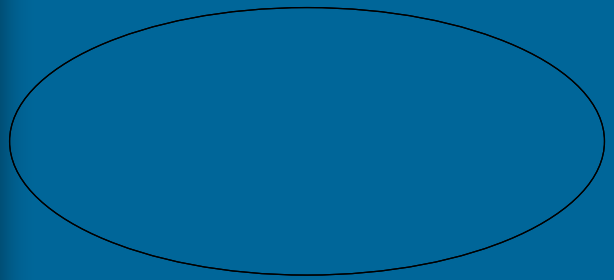
Object A



Object B

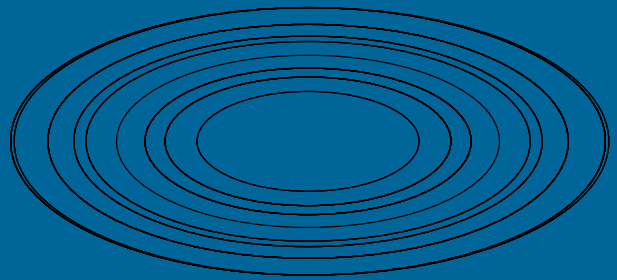


Random set A

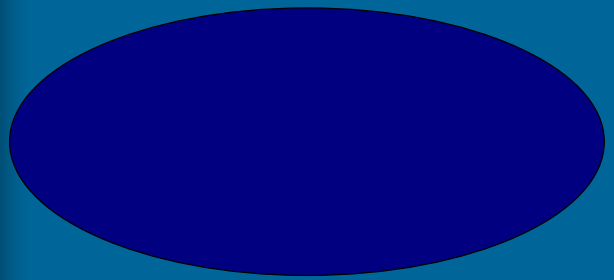


n = 8

Random set B

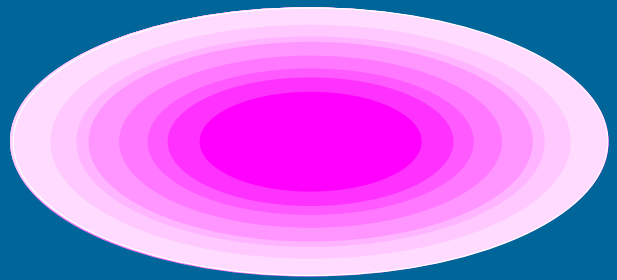


Covering function A



n = 8

Covering function B



Cov. function A stable at n = 2

Cov. function B stable at n = 8

Method randomized region growing

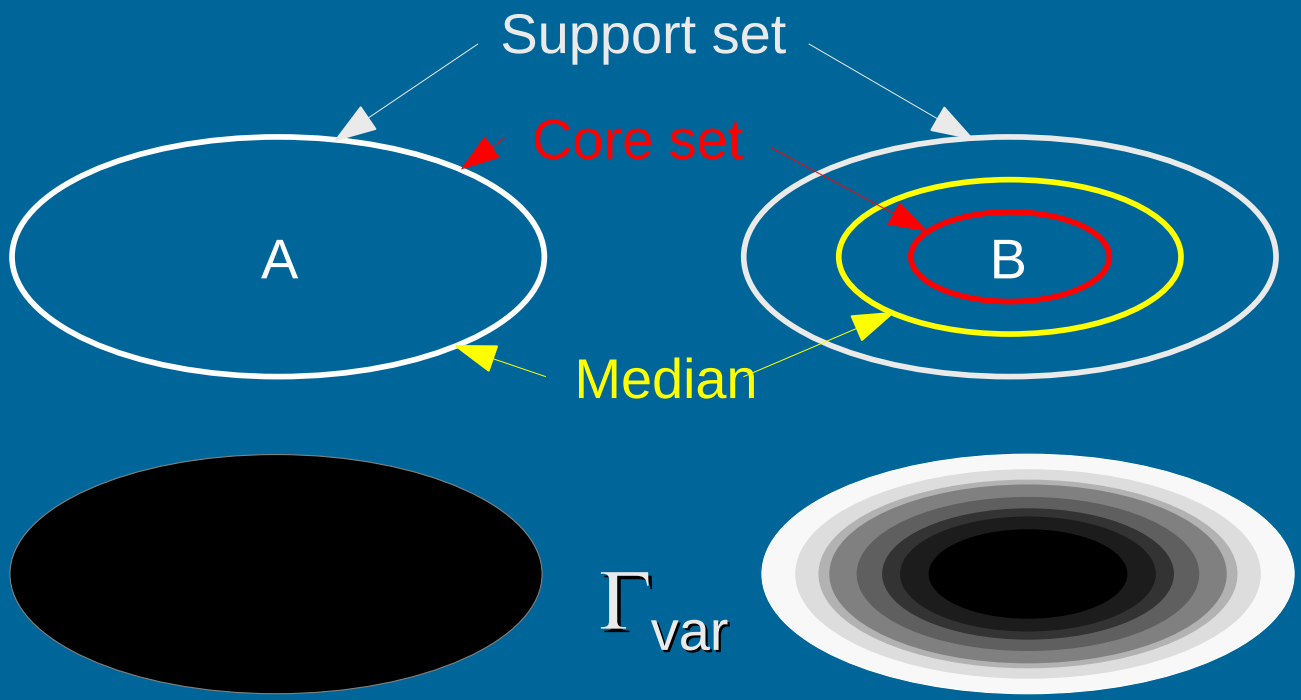
- Select growing seeds interactively
- Randomize parameters in algorithm
- Obtain random set $\{O_0,\ldots,O_n\}$ and its coverage function

$$f^{(n)} = p_\Gamma(x)dx = \frac{1}{n}\sum_{i=1}^{n} I_{O_i}(x)$$

- Stop the algorithm at step *n* when coverage function changes slightly $| f^{(n)} - f^{(n-1)} | < \varepsilon$

Method moments of random sets

- Support set, median set and core set
- variance Γ_{var}



Method Ground Survey

- October 26th ~ November 6th, 2009
- 4 transects, 73 sample plots

30*30m, visual
GPS, V, VC

1*1m, measure
H, V, VC

1*1m, measure
spectroradiometer

Method Accuracy Assessment

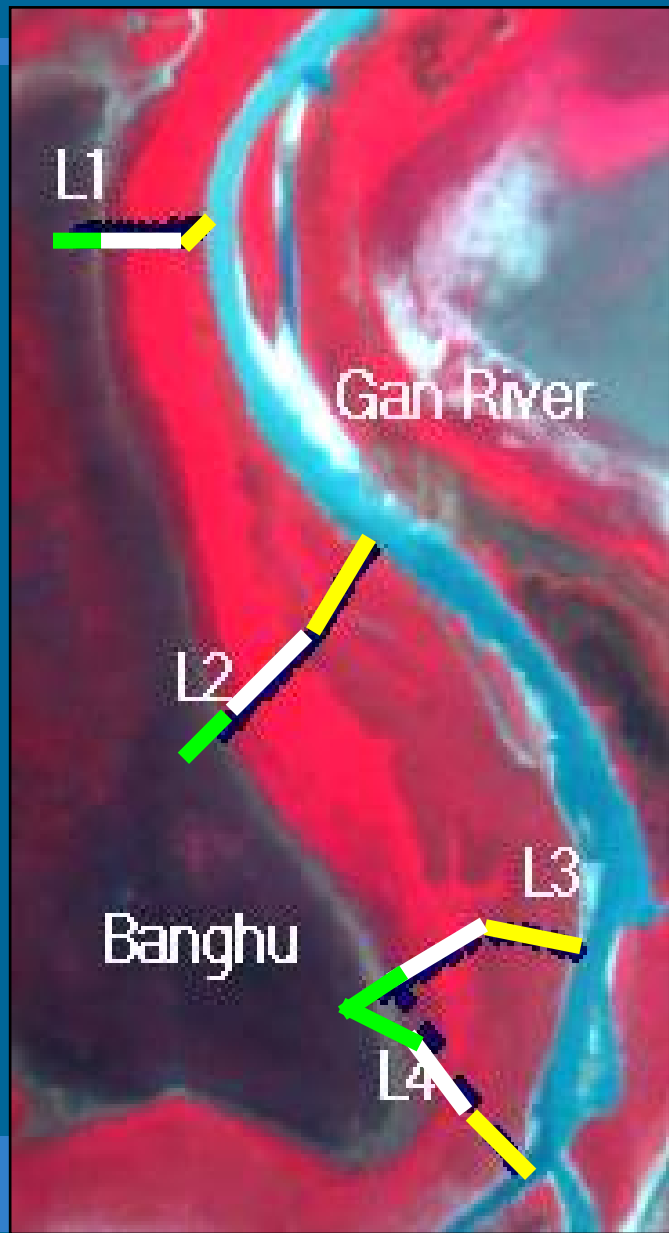
- relationship between *Carex* coverage, height, density and covering function value derived from random set model
- Independent samples t-tests: mean value of the *Carex* coverage is different for sample plots which are included and excluded by the median set
-
- The overall accuracy (OA), producer accuracy (PA), user accuracy (UA) and kappa coefficient are derived from error matrix

Outline

- Introduction
  - Uncertainty, Accuracy assessment, objective
- Method
  - Study area, data, random set model, field work, accuracy assessment
  - Results
  - Field work measures, random set modeling results, accuracy report
- Conclusion and discussion

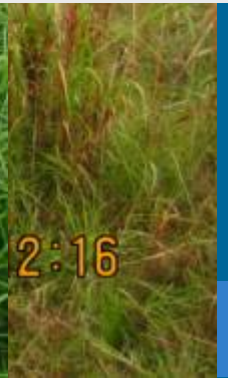
Results different zones along transects

- near the river bank
  - flowered *Miscanthus* of 1-2 m
  - mixed with *Cynodon*, *Carex*, *Polygonum*, *Artemisia* and human planted poplar
- *Carex* dominant zones
  - 500m belt with gradual changes in boundary
  - few mixed with *Artemisia*
- on the lake bank
  - low density young short *Carex* on wet soil
  - shallow water and dead *Potamogeton* and *Vallisneria* beneath

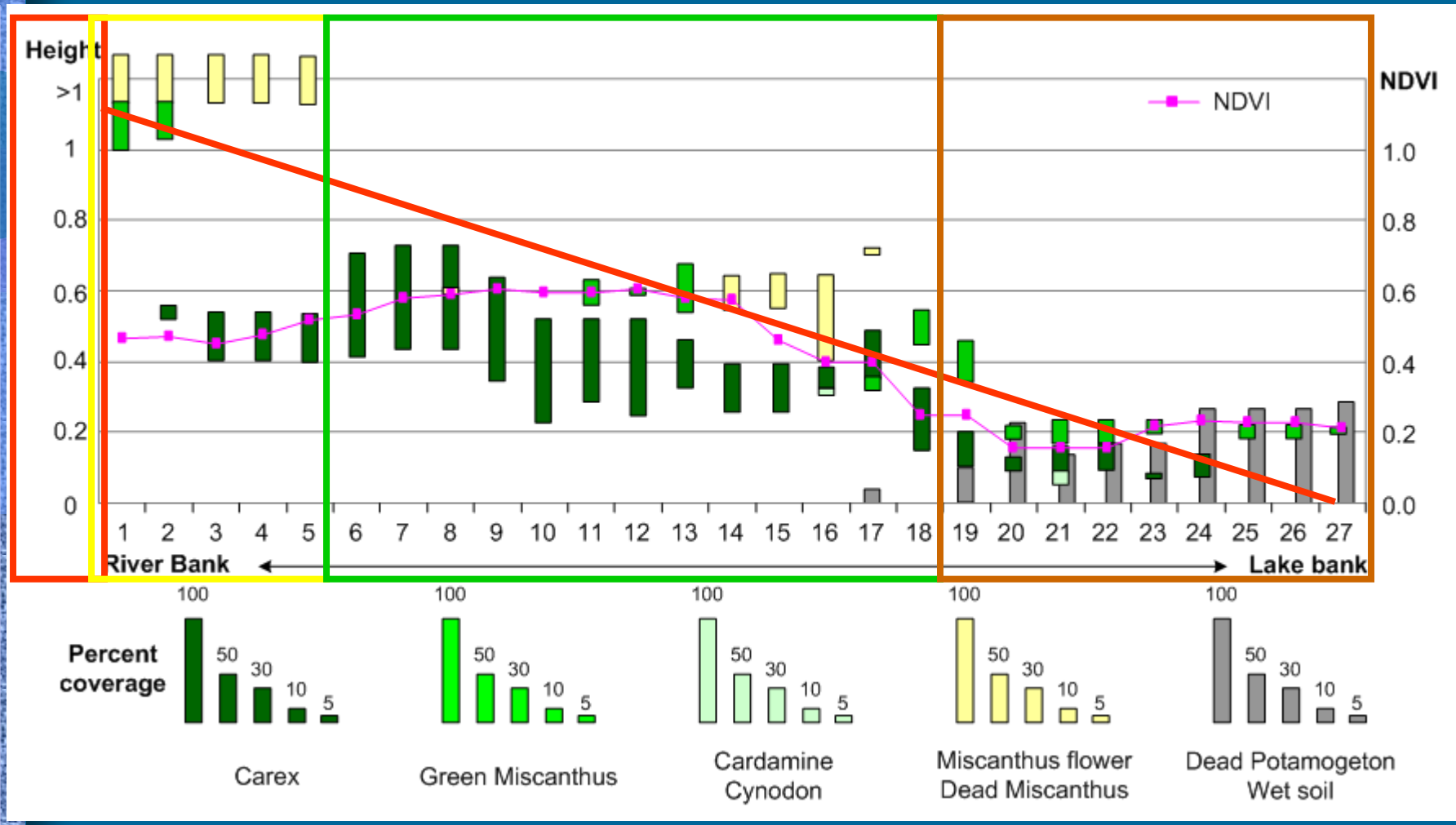


Results

On the lake bank
shallow water, dead *Potamogeton*
and *Vallisneria* beneath

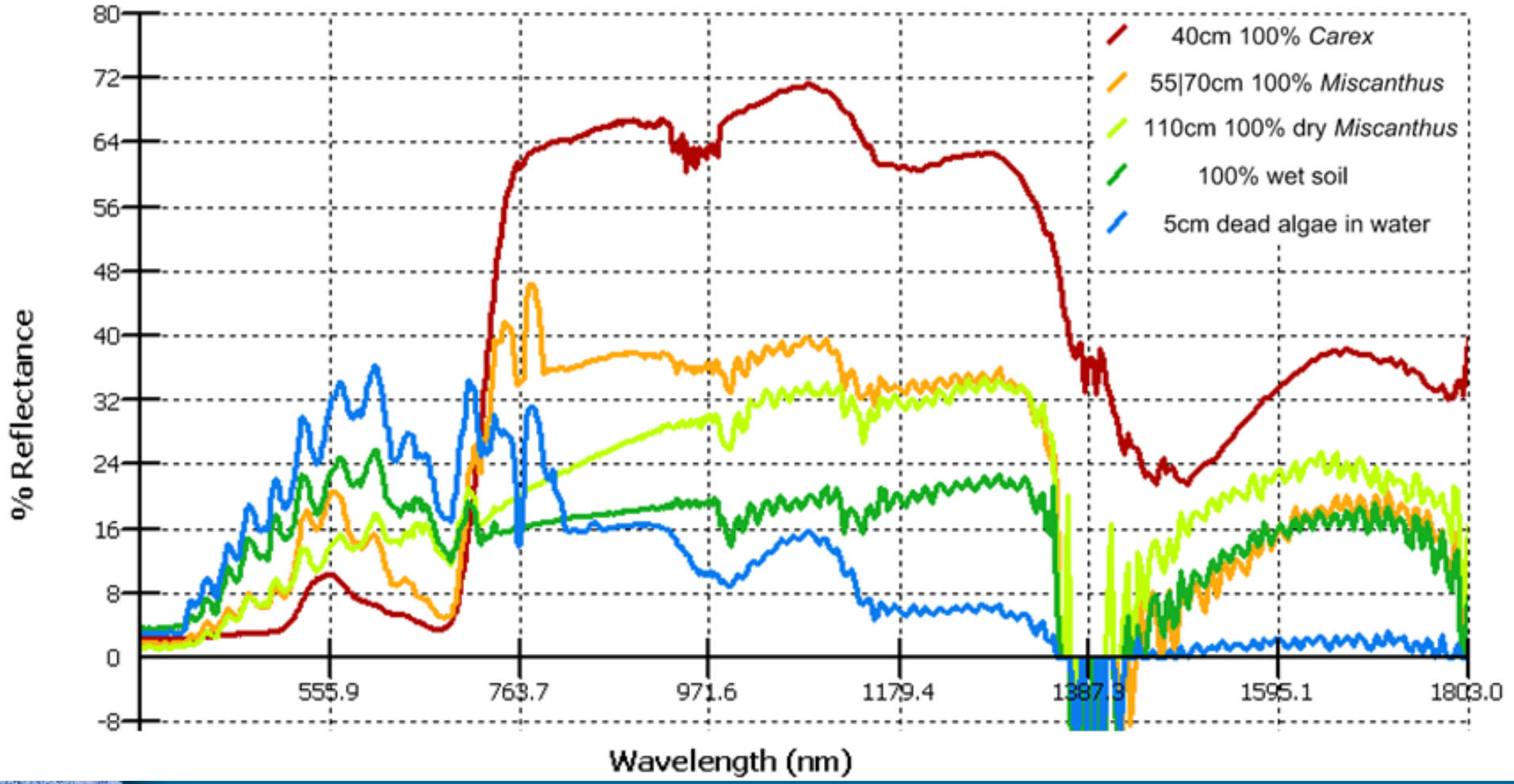


Results



Types of dominant vegetation, their percent coverage and heights along transect L1 are compared with NDVI extracted from corresponding pixels at field samples

Results Reflectance curves



Results V & C & H & field NDVI

	Carex	*Miscanthus* flower|green	*Artemisia*
C (%)	100	100	100
H(m)	0.4-0.6	1.2|1.0	0.4
NDVI	0.89	0.73	0.79

(a) Carex

(b) Artemisia

Results

Carex patch presented by Random set

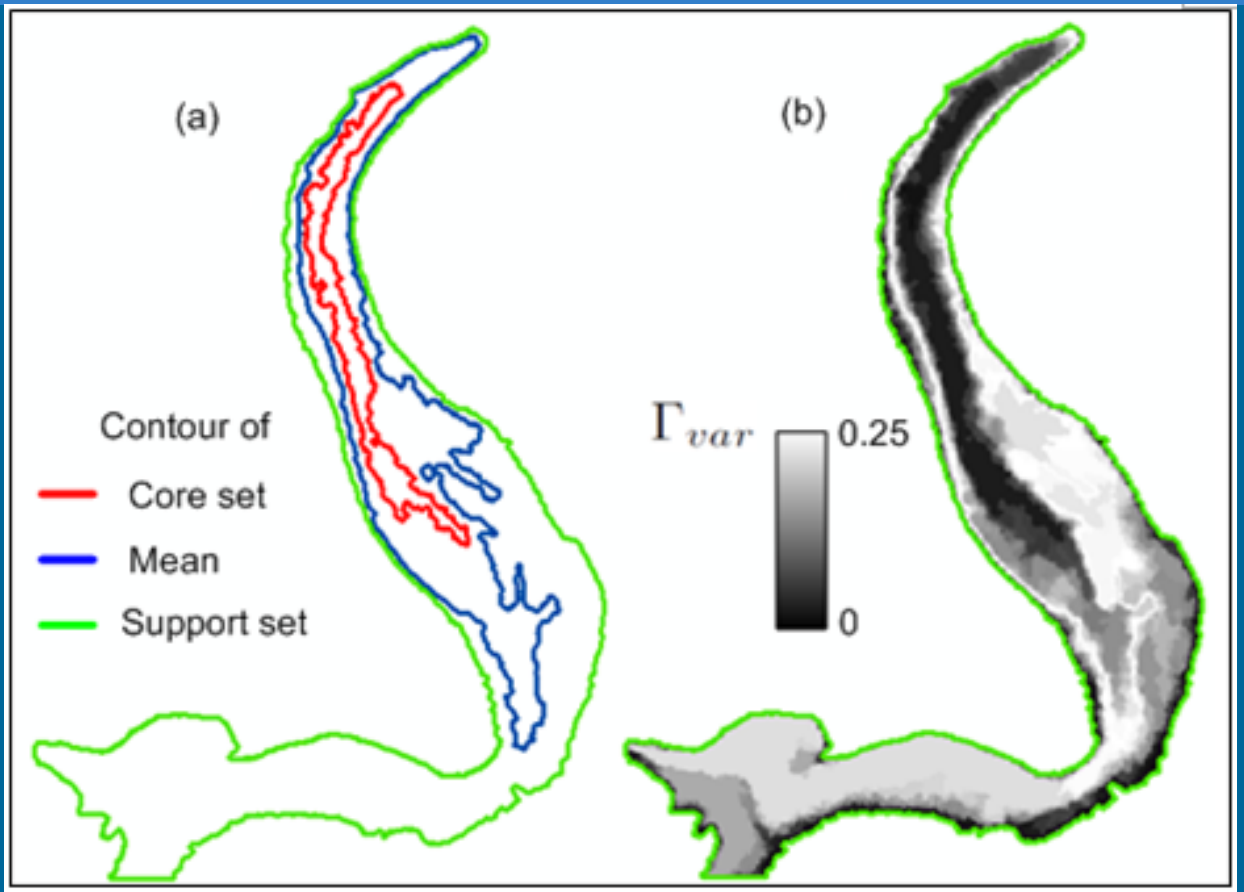
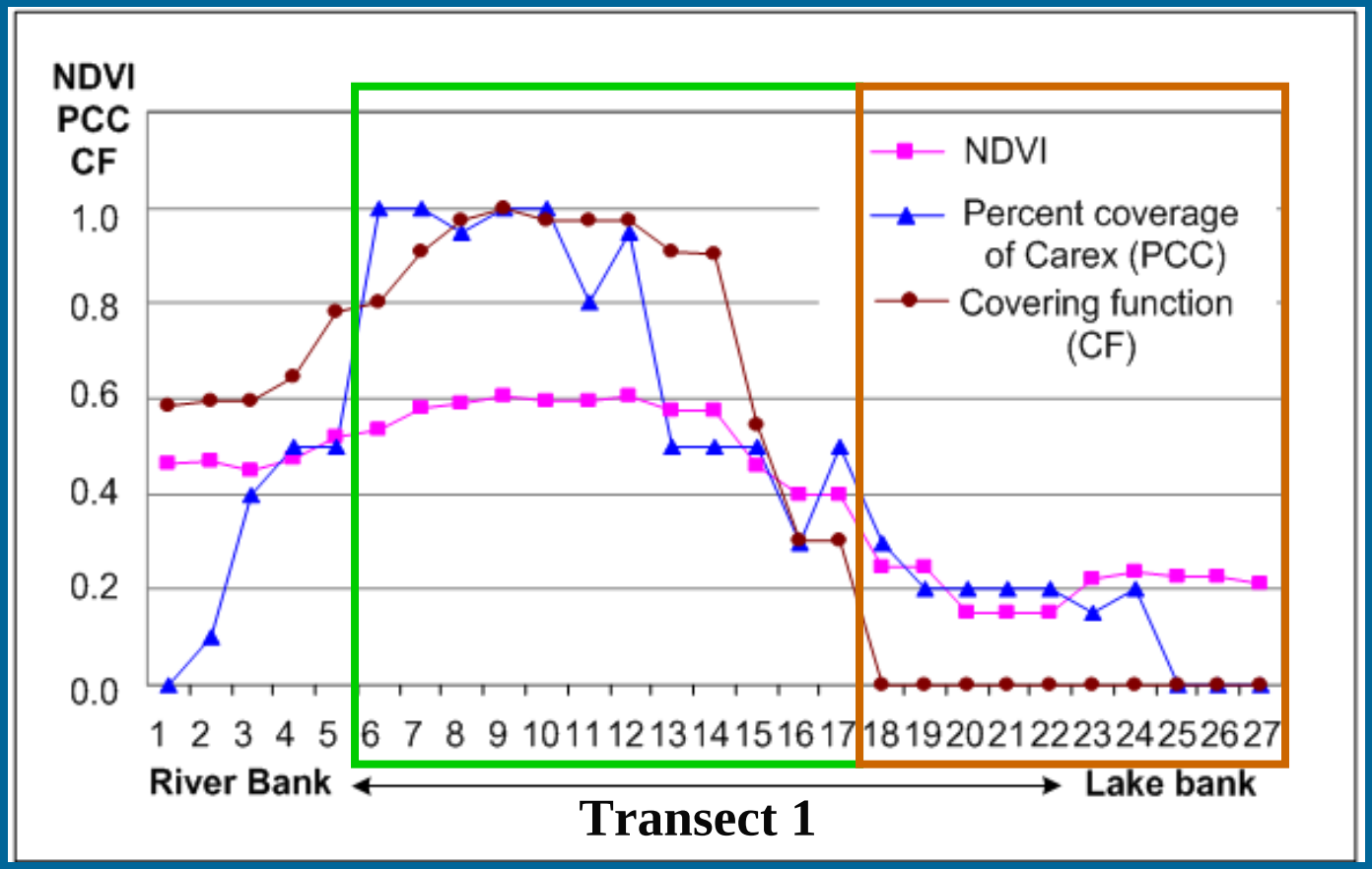


Fig. 4. Extracted object and its extensional uncertainty described by concepts from random set theory: (a) support set, median and core set (b) variance

Results

Relationship between NDVI, PCC and CF



Results

Relationship between PCC and median set

- T-test to explore the relationship between median set and *Carex* coverage

- $H_{0:}$ the mean value of the *Carex* coverage of samples which included by the median set = the mean value of the *Carex* coverage of samples which excluded by the median set.

Results

Relationship between PCC and median set

Independent Samples Test				
	Levene's Test for Equality of Variances			
	F	Sig.	df	Sig. (2-tailed)
Carex Coverage	2.755	0.101	71	0.000

- reject the null hypothesis
- there is sufficient evidence to conclude that samples included and excluded by the median set have different *Carex* coverage

Results Reference variable

Table 3. R^2-values of the correlation relationships between covering function (CF) and percent coverage of *Carex* (PCC) and NDVI for four transects separately and in total

Transect	L1	L2	L3	L4	L1-L4
CF-PCC	0.67	0.51	0.71	0.46	0.54
CF-NDVI	0.97	0.91	0.82	0.91	0.93
PCC-NDVI	0.63	0.62	0.63	0.48	0.56

- use *Carex* coverage percent as the reference variable on the ground for accuracy assessment

Results Accuracy report

Core set	PA (%)	UA (%)	OA (%)	Kappa
presence	47	90	85	0.54
absence	98	84		
Median set				
presence	76	85	77	0.52
absence	78	66		
Support set				
presence	82	93	78	0.22
absence	50	25		

- presence and absence of *Carex* has high PA and UA for support set and core set respectively, which indicate that these two classes are reliable in support set and core set respectively

Results Accuracy report

Core set	PA (%)	UA (%)	OA (%)	Kappa
presence	47	90	85	0.54
absence	98	84		
Median set				
presence	76	85	77	0.52
absence	78	66		
Support set				
presence	82	93	78	0.22
absence	50	25		

- Presence of *Carex* has high UA and low PA in core set, showing that there is more area of *Carex* in the field than is indicated by the core set

Results Accuracy report

Core set	PA (%)	UA (%)	OA (%)	Kappa
presence	47	90	85	0.54
absence	98	84		
Median set				
presence	76	85	77	0.52
absence	78	66		
Support set				
presence	82	93	78	0.22
absence	50	25		

- the grouping criteria of making testing samples for the support set is not appropriate
- the support set is not sensitive to the *Carex* coverage lower than 20 percent

Results Accuracy report

Core set	PA (%)	UA (%)	OA (%)	Kappa
presence	47	90	85	0.54
absence	98	84		
Median set				
presence	76	85	77	0.52
absence	78	66		
Support set				
presence	82	93	78	0.22
absence	50	25		

- A kappa z-test for pair-wise comparison in accuracy shows:
  - significant difference between the support set and other sets
  - no significant difference between the core set and the median set

Outline

- Introduction
  - Uncertainty, Accuracy assessment, objective
- Method
  - Study area, data, field work, random set model
- Results
  - Field work measures, random set modeling results, accuracy report
- Conclusion and discussion

Conclusion and Discussion

- Covering function of the random set can be quantified and interpreted by NDVI derived from image and *Carex* coverage measured in the field
  - other variables: vegetation types, height and density
  - belong to different scales: nominal (e.g. vegetation type), ordinal (e.g. big or small density) and ratio scale (e.g. height and coverage)
  - difficult to integrate

Conclusion and Discussion

- The accuracy of core set is better than that of the median set and much better than that of the support set
  - has a better performance on the high coverage area and criteria for validating the support set should be determined not only based on the coverage
  - it supports that *Carex* coverage cannot be the only variable fully explaining the covering function and other variables such as height should be considered especially when the coverage is low

Conclusion and Discussion

- the accuracy of random set model applied in this study is just moderate according to the assessment report
  - the parameters in the region growing segmentation algorithm need further adjustments.
  - More efficient procedure for selecting parameters in the random set generation should be explored

Thanks for your attention