

Department of Geography, College of Science National Taiwan University Master Thesis

SAR影像的都市水體判釋: 考量都市表面形態造成的雷達二次反射效果

Detecting urban water bodies from SAR images: Measuring surface morphology contributing to radar double bounce effect

廖晧宇 Hao-Yu Liao

指導教授:溫在弘博士 Tzai-Hung Wen, Ph.D.

中華民國108年1月30日

Detecting water from space...

- Water bodies related topics:
 - Natural resources, ecological roles (Sakamoto et al., 2007; Schaffer-Smith et al., 2017)
 - Environmental impacts, human assets (Huang, Chen et al. 2018)

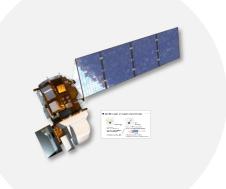
In the past...

Nowadays,



Field survey

Satellite remote sensing



The launch of Landsat-1 in the US. in 1972.

- ✓ Less labor investment
- ✓ At a global scale
- ✓ Regular and in-time surveying

Satellite remote sensing for water detection

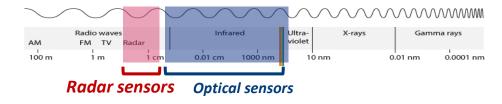


Optical images

- Passive sensors
- Reflectance at different bands
- Water index e.g. NDWI (McFleeters 1996), mNDWI (Xu 2006), ...
- Cons: Cloud coverage (Asner 2001)



- Active sensors → work day-and-night
- Radar band → Overcome weather conditions



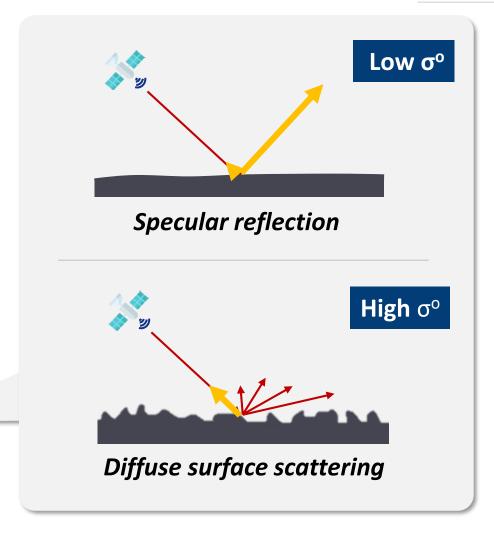
- Sentinel-1A, launched on 3 April 2014
- Globally open, various modes for different monitoring

About SAR images...



- Backscatter value (σ°)
- Primary factor:
 - Surface structure roughness

(Xia & Henderson 1997)

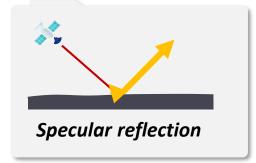


How to detect water?

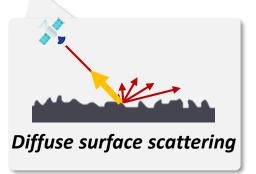


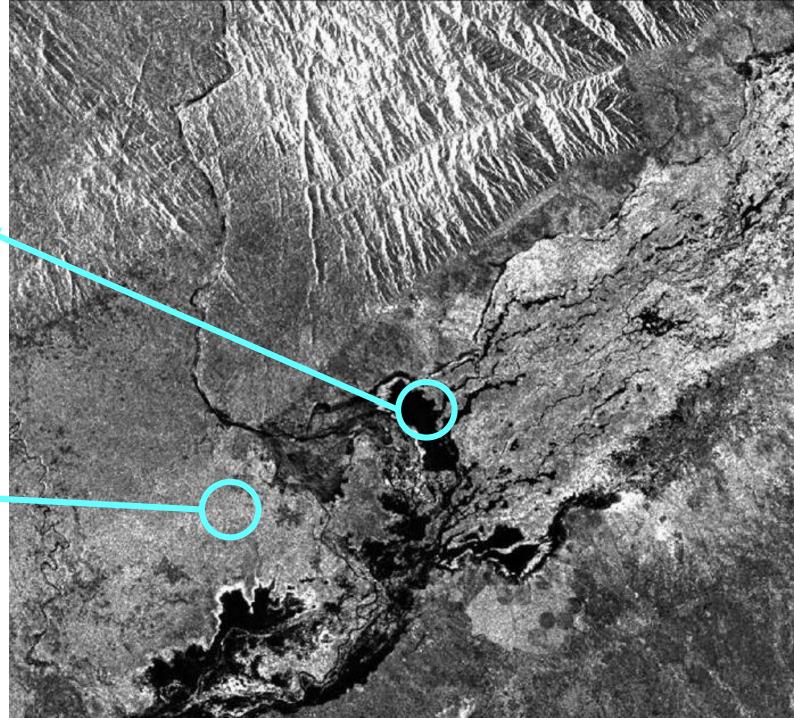
Detect water with SAR images

Water: Low backscatter value



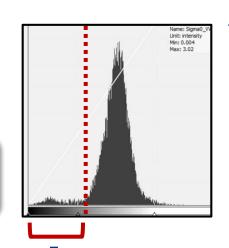
Land: High backscatter value

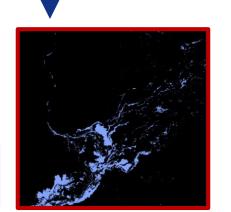




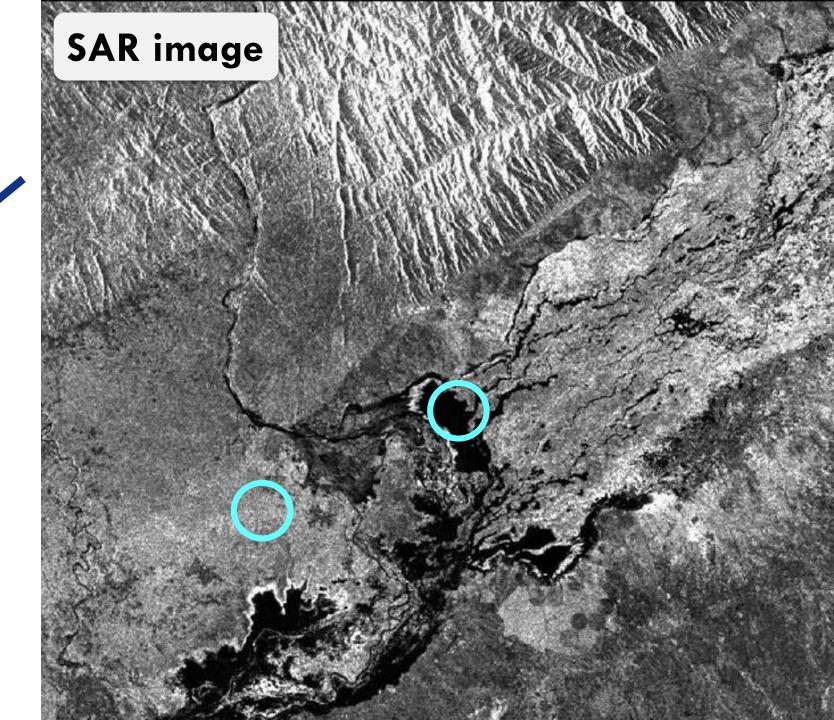
Detect water with SAR images

Backscatter distribution





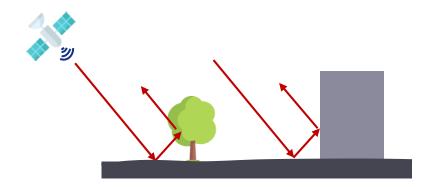
Water bodies



Double bounce effects

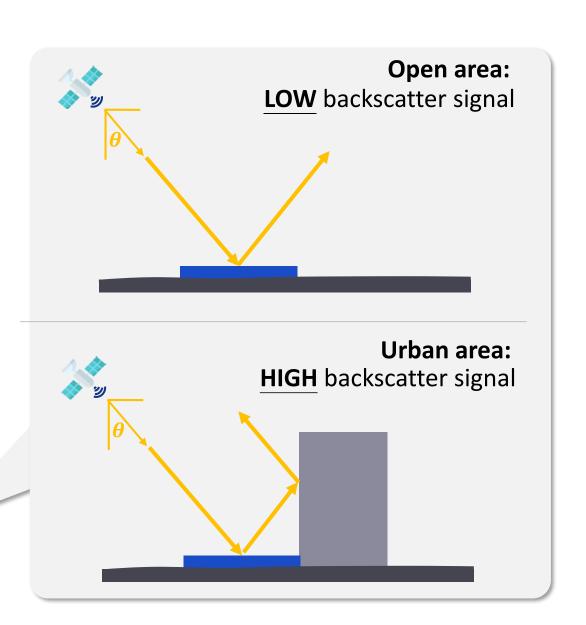
(Brunner et al. 2008)

- Vertical structures on the surface
 - e.g. Walls of building
- A strong scattering mechanism caused by a corner reflector



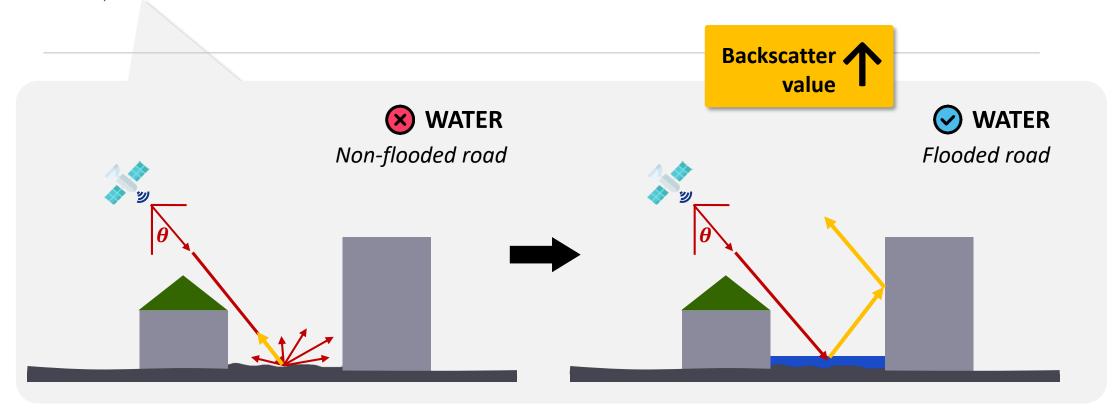
Double bounce effect

What happened to water's backscatter?



How does double bounce effect influence urban water detection?

- The accuracy of urban water detection significantly decreased. (Mason et al. 2010)
- Flooding roads adjacent to buildings had higher backscatter values. (Watanabe et al. 2010)



Detecting urban water in previous studies



- 1) Masking out interfered areas (Mason et al. 2010; Giustarini et al. 2013)
- 2) Detect urban areas with increasing backscatter response (Mason et al. 2014; Pulvirenti et al. 2016)

- ① Require **two** or more SAR images (need an image as reference)
- ② Finding "additional" water based on the referenced time
- 3 Double bounce effect **happens** in the entire urban area
- Double bounce effect is treated with the same rule in the entire urban area

Relationship between double bounce effect and adjacent structure

- Theoretical equations (Franceschetti, Antonio & Daniele 2002)
 - **Factors** of double bounce effect: Building height, texture of ground and buildings, building orientation, incident angle, ...
- Empirical studies (Brunner et al. 2008; Ferro et al. 2011)
 - Intensity of double bounce effect varies with different building characteristics and SAR system parameters

How do these related to urban water detection?

• With different nearby buildings or structures, **intensity** of water's double bounce effect is **different** \rightarrow **Different backscatter value** (σ °)

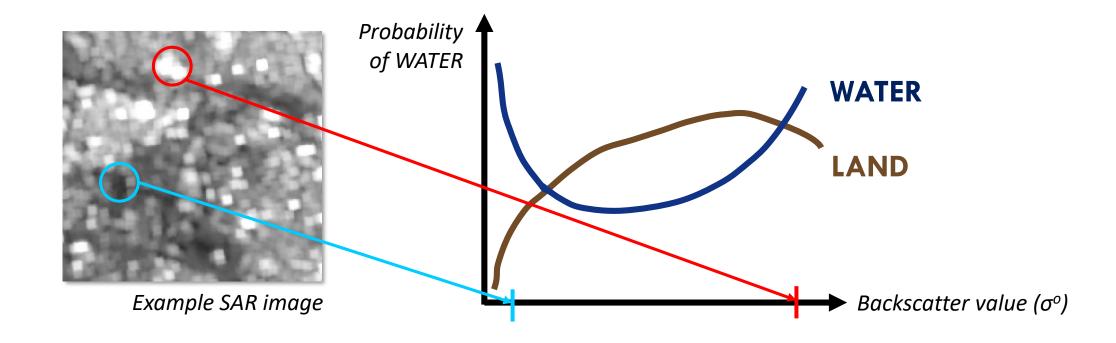
High $\sigma^{\circ} \rightarrow$ WATER

How do these related to urban water detection?

• With different nearby buildings or structures, **intensity** of water's double bounce effect is **different** \rightarrow **Different backscatter value** (σ ^o)

Low $\sigma^{o} \rightarrow WATER$

High $\sigma^{\circ} \rightarrow$ WATER



RESEARCH GAP

Methods for detecting urban water bodies:

- The effect is treated uniformly in space
- Characteristics of nearby structures aren't considered
- Require multi-SAR images

How to control double bounce effect:

- Require assumptive, empirical parameters
- Require detailed data
- Difficult in large-scale applications
- Water is **not** the subject



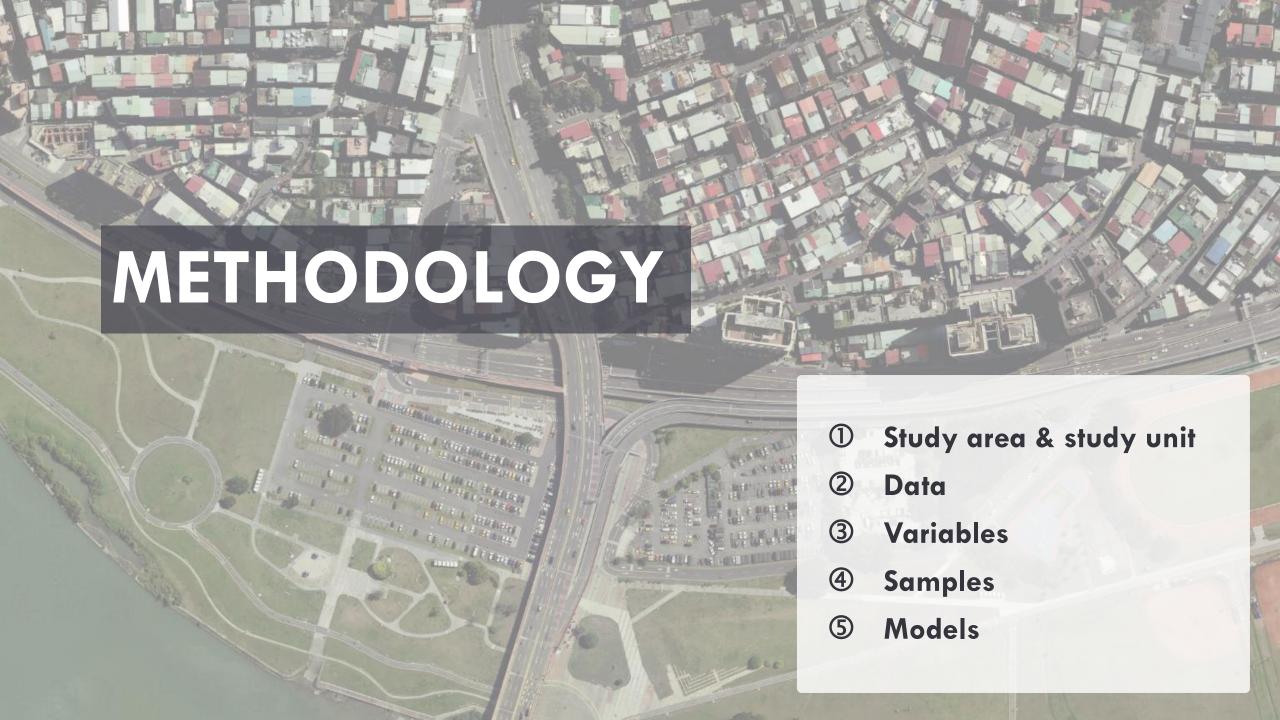
Highlights:

- Single SAR image
- Without detailed environmental data
- > Spatial difference of double bounce effect

Propose an urban water detection framework integrating urban surface morphological factors:

1. Mechanism & relationship

- How surface morphology contributes to double bounce effect
- 2. Data to quantify urban surface morphology
 - Digitized building data (vector) or DSMs and DEMs (raster)
- 3. Generic model / modeling framework
 - Feasible in other cities with varied urban landscape

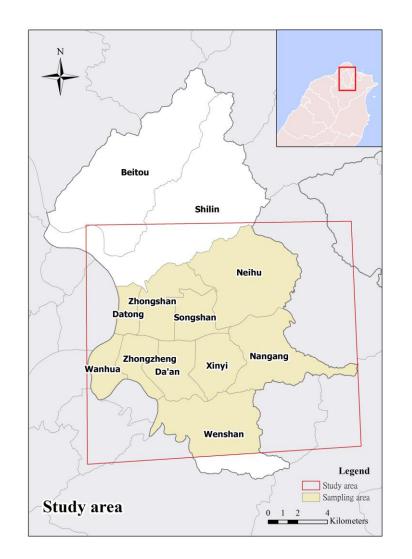




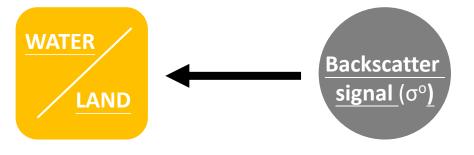
Study area & study unit

- O Study area, unit
- Data
- Variables
- Samples
- Models

- A part of Taipei city (10 towns)
 - 大同區、中山區、松山區、內湖區、萬華區、 中正區、大安區、信義區、南港區、文山區
- Study unit: 10m × 10m cell
 - The spatial resolution of SAR image
- Study interest:
 - WATER → Static water body
 - LAND → Impervious road









- Study area, unit
- Data
- Variables
- Samples
- Models

Ground truth

Aerial orthophotos



Spatial resolution: $0.1m \times 0.1m$

Backscatter value

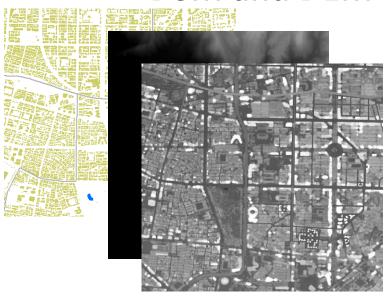
Sentinel-1 SAR images



Spatial resolution: $10m \times 10m$

Surface morphological variables

Digitized building layer; DSM and DEM













- Study area, unit
 - Data
- Variables
- Samples
- Models

Ground truth

Aerial orthophotos



Spatial resolution: $0.1m \times 0.1m$

Backscatter value

Sentinel-1 SAR images



Spatial resolution: $10m \times 10m$

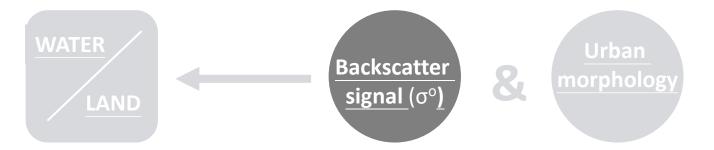
Surface morphological variables

Digitized building layer; DSM and DEM



Spatial resolution: DTM 20m, DSM 30m





Study area, unit Data

- Variables
- Samples
- Models

Ground truth

Aerial orthophotos



Spatial resolution: $0.1m \times 0.1m$

Backscatter value

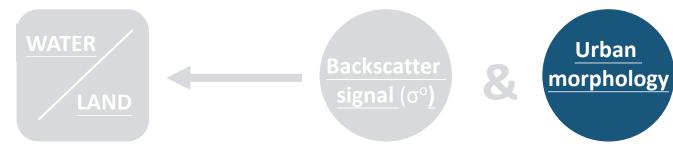
Sentinel-1 SAR images



Spatial resolution: $10m \times 10m$

- 15 images
- Mean backscatter value
- Dry dates
 - Compare with precipitation data
- Same SAR system-related parameters
 - C-band
 - Incidence angle: 32.9° ~ 43.1°
 - IW mode
 - VV polarization
 - Ascending images





- Study area, unit
 - Data
- Variables
- Samples
- Models

Ground truth

Aerial orthophotos



Spatial resolution: $0.1m \times 0.1m$

Backscatter value

Sentinel-1 SAR images



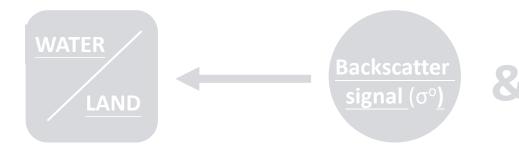
Spatial resolution: $10m \times 10m$

Surface morphological variables

Digitized building layer; DSM and DEM







- Study area, unit
- Data
- Variables
- Samples
- Models



1. Digitized building layer

- Vector data
- Detailed investigation from the city government

2. DSM and DEM



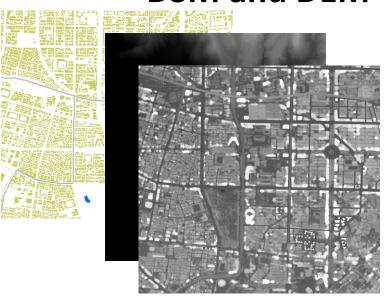
- Raster data
- Has been more easily acquired nowadays e.g. LiDAR, flying UAVs
- Free, open data

Surface morphological variables

Urban

morphology

Digitized building layer; DSM and DEM







- O Study area, unit
- Data
- Variables
- Samples
- Models



1. Digitized building layer

- Vector data
- Detailed investigation from the city government

2. DSM and DEM



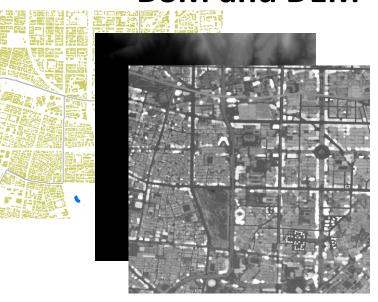
- Raster data
- Has been more easily acquired nowadays e.g. LiDAR, flying UAVs
- Free, open data

Surface morphological variables

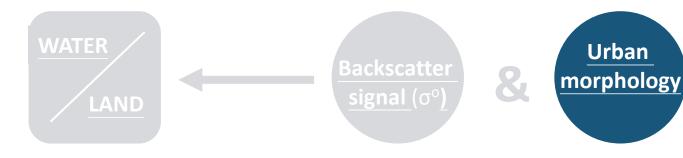
Urban

morphology

Digitized building layer; DSM and DEM







- Study area, unit
- Data
- Variables
- Samples
- Models



1. Digitized building layer

- Vector data
- Detailed investigation from the city government

- Each building's boundary → area and structural patterns
- Each building's attribute → building height
- → The characteristics and patterns of buildings

Surface morphological variables

Digitized building layer; DSM and DEM







- Study area, unit
- Data
- Variables
- Samples
- Models



- Vector data
- Detailed investigation from the city government

2. DSM and DEM



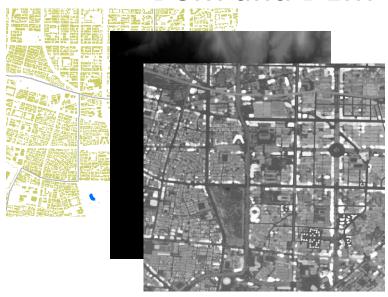
- Raster data
- Has been more easily acquired nowadays e.g. LiDAR, flying UAVs
- Free, open data

Surface morphological variables

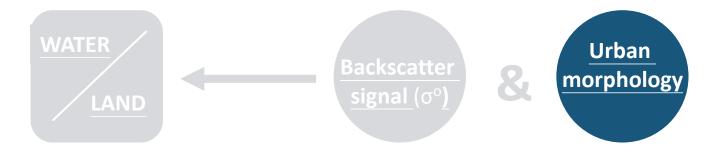
Urban

morphology

Digitized building layer; DSM and DEM







- Study area, unit
 - Data
- Variables
- Samples
- Models

- DSM DEM
 Næsset (1997); Popescu, Wynne, and Nelson (2002); Ural et al. (2011)
- → Surface of urban height



Reference Replot from: https://3dmetrica.it/dtm-dsm-dem/

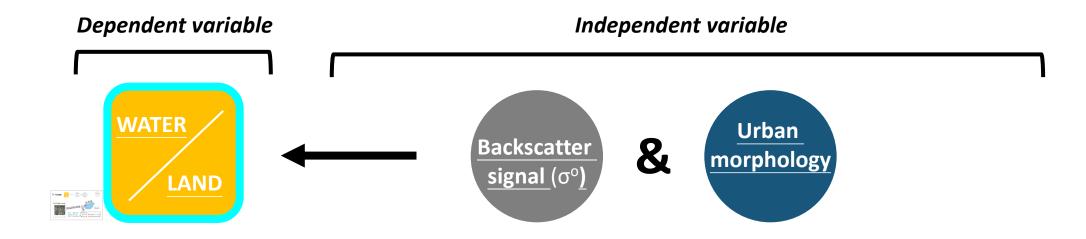
2. DSM and DEM



- Raster data
- Has been more easily acquired nowadays e.g. LiDAR, flying UAVs
- Free, open data

Variables

- Study area, unit
- O Data
- Variables
- Samples
- Models



Binary class

WATER:

34 100% of water in the cell14 1-99% of water in the cell

Backscatter value

 $\sigma^0 = 10 \times \log 10(amplitude)$

Unit: decibel

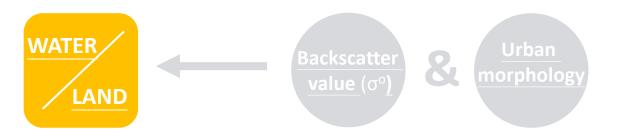
Urban morphological variables

- 1) Neighboring
- 2) Horizontal & Vertical aspects

LAND:

0% of water in the cell





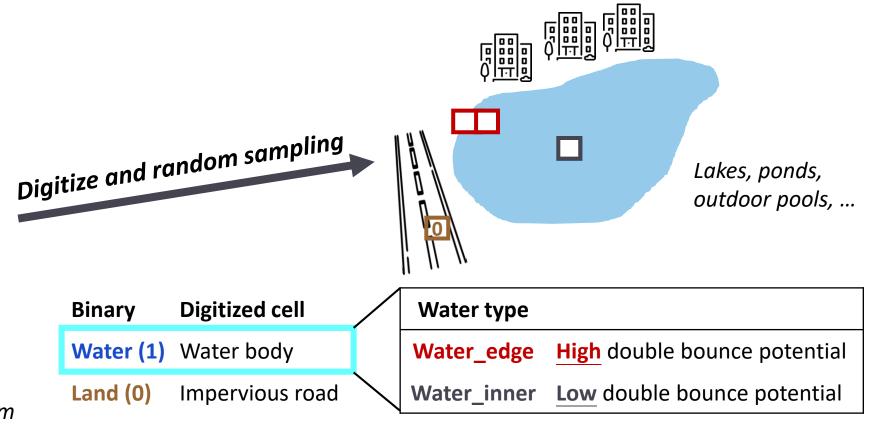
- Study area, unit
- O Data
- Variables
- Samples
- Models

Ground truth

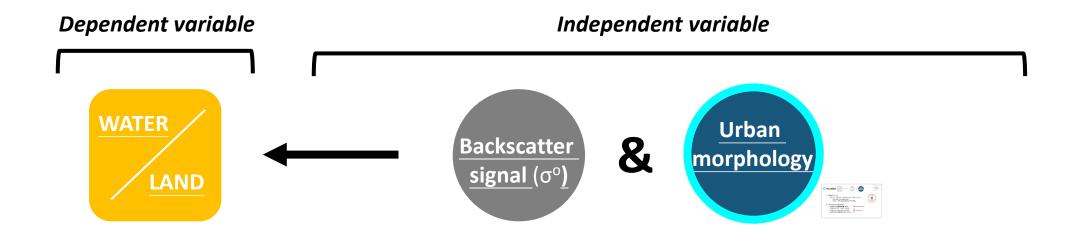
Aerial orthophotos



Spatial resolution: $0.1m \times 0.1m$



- Study area, unit
- O Data
- Variables
- Samples
- Models



Binary class

WATER:

¾ 100% of water in the cell¼ 1-99% of water in the cell

LAND:

0% of water in the cell

Backscatter value

 $\sigma^0 = 10 \times \log 10(amplitude)$

Unit: decibel

Urban morphological variables

- 1) Neighboring
- 2) Horizontal & Vertical aspects







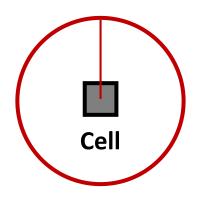




- Study area, unit
- O Data
- Variables
- Samples
- Models

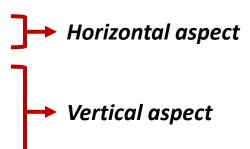
1) Neighboring

- The spatial extent of double bounce effects: 120m
 - Road width and building height (Mason et al. 2007; Soergel, Thoennessen & Stilla 2003)

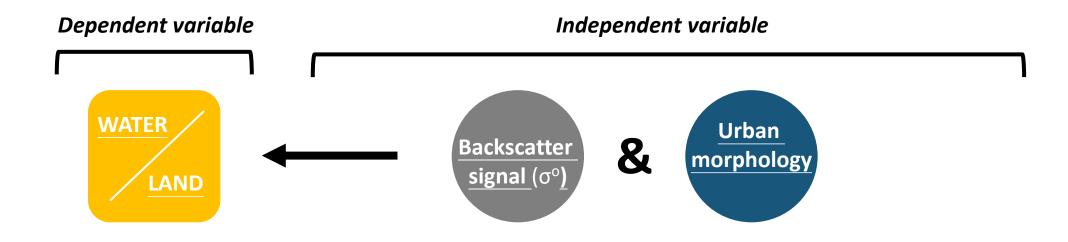


2) Calculating variables:

- Neighboring horizontal density
- Neighboring Max. vertical height
- Neighboring Mean vertical height
- Neighboring **vertical** height variation



- Study area, unit
- O Data
- Variables
- Samples
- Models



Binary class

WATER / LAND

Backscatter value

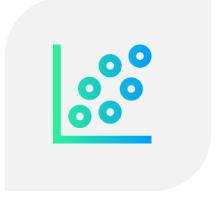
 σ^0

Urban morphological variables

Neighboring Max. vertical height Neighboring Mean vertical height Neighboring vertical height Neighboring vertical height variation



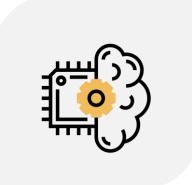
- Study area, unit
- O Data
- Variables
- Samples
- Models



Statistics-based model

• Logistic regression





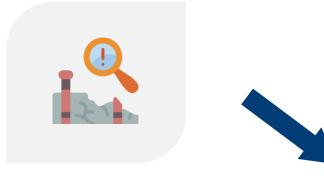
Machine learning-based model

- Support vector machine (SVM)
- Random forest (RF)
 - → Model performances



• Part 2. Model performances





1) Explore double bounce effect among different water types



2) The relationship between urban morphology and double bounce effect



3) Spatial distribution of double bounce effect

• Part 2. Model performances







1) Explore double bounce effect among different water types



2) The relationship between urban morphology and double bounce effect

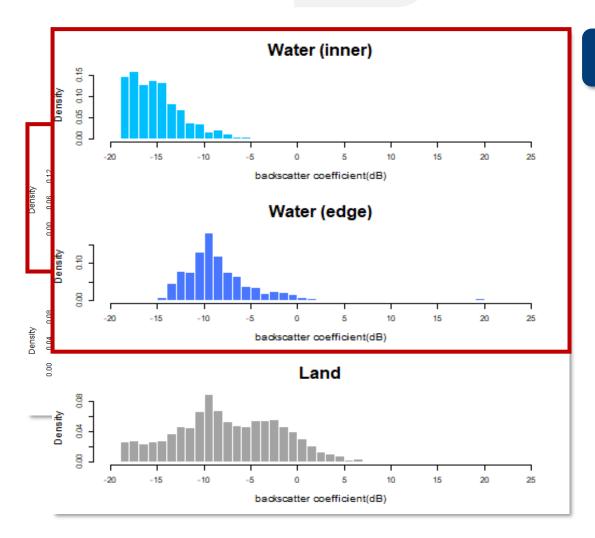


3) Spatial distribution of double bounce effect

• Part 2. Model performances



1) Explore double bounce effect among different water types



\triangleright Edge WATER's $\overline{\sigma^0}$ vs. Inner WATER's $\overline{\sigma^0}$

Two independent samples T-test					
Water cell types	N	Mean	StDev.	SE.	
Inner WATER cells	80	-14.41	2.59	0.289	
Edge WATER cells	80	-8.14	4.45	0.498	

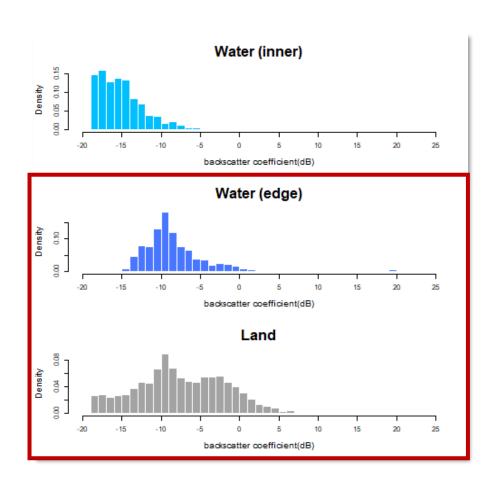
95% Confidence interval for $\mu_{inner} - \mu_{edge}$: $-7.409 \sim -5.131$ Alternative hypothesis: True difference in means is not equal to 0 t = -10.892; degree of freedom = 126.96; p-value $< 2.2 \times 10^{-16}$

- ${\cal O} \ \overline{\sigma^0}_{inner}$ is different from $\overline{\sigma^0}_{edge}$
- $\mathcal{Q} \ \overline{\sigma^0}_{\text{edge}} > \overline{\sigma^0}_{\text{inner}}$





1) Explore double bounce effect among different water types



\triangleright Edge WATER's $\overline{\sigma^0}$ vs. LAND's $\overline{\sigma^0}$

Two independent samples T-test

Types	N	Mean	StDev.	SE.
Edge WATER cells	80	-7.21	6.83	0.764
Impervious LAND cells	80	-8.09	6.18	0.691

95% Confidence interval for $\mu_{inner} - \mu_{edge}$: $-1.149 \sim 2.920$ Alternative hypothesis: True difference in means is not equal to 0 t = 0.8599; degree of freedom = 156.45; p-value = 0.3912

- ① It is hard to tell difference from $\overline{\sigma^0}_{WATER_edge}$ and $\overline{\sigma^0}_{LAND}$ (p-value = 0.3912)
- → Could be the reason of mis-classification in urban areas

Random Forest

0.878

• Part 2. Model performances



1) Explore double bounce effect among different water types

> Differentiating specific WATER type & LAND

- Water type:
 - Inner water
 - Edge water
- Models:
 - Null model
 - → NO morphological variables

* Recommended Practice for flood mapping. 2015
United Nations – Office of Outer Space Affairs
(UN-SPIDER).

0.956

Random Forest

•	Maii	n m		اما
•	ıvıaıı	1 11	IUU	e.

→ ADD morphological variables

Our proposed models

Model	Overall	Model	Overall		
Wiodei	Accuracy		Accuracy		
Inner water Edge water Null models Null models		Edge water			
		Null models			
Thresholding	0.929	Thresholding	0.786		
Logistic regression	0.930	Logistic regression	0.791		
Support Vector Machine	0.931	Support Vector Machine	0.799		
Random Forest	0.898	Random Forest	0.723		
Main models M		Main models			
(Characteristics and patterns of buildings)		(Characteristics and pattern	(Characteristics and patterns of buildings)		
I . I D	7 - 1	ogistic regression	0.820		
nended Practice for flood mapping. 2015 Nations – Office of Outer Space Affairs		support vector Machine	0.895		
		e Affairs Landom Forest	0.885		
IDER).		face of urban height)			
Logistic regression	0.928	Logistic regression	0.797		
Support Vector Machine 0.952		Support Vector Machine	0.859		





1) Explore double bounce effect among different water types

➤ Differentiating specific WATER type & LAND

Inner WATER vs. LAND

• Null models have overall accuracies **over 0.92**

Model	Overall Accuracy	Recall
Inner water		
Null models		
Thresholding	0.929	0.939
Logistic regression	0.930	0.945
Support Vector Machine	0.931	0.949
Random Forest	0.898	0.883
Main models		
(Characteristics and patterns of		
Logistic regression	0.932	0.939
Support Vector Machine	0.966	0.988
Random Forest	0.959	0.958
(Surface of urban height)		
Logistic regression	0.928	0.949
Support Vector Machine	0.952	0.976
Random Forest	0.956	0.941





1) Explore double bounce effect among different water types

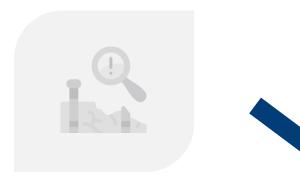
Differentiating specific WATER type & LAND

Edge WATER vs. LAND

- ① Poor in null models
- Adding urban morphological variable,
- Overall accuracy \uparrow (0.79 \rightarrow 0.90)
- 94% of edge water is detected (Null 81% → 94%)

Model	Overall Accuracy	Recall
Edge water		
Null models		
Thresholding	0.786	0.805
Logistic regression	0.791	0.841
Support Vector Machine	0.799	0.820
Random Forest	0.723	0.714
Main models		
(Characteristics and patterns of buildings		
Logistic regression	0.820	0.817
Support Vector Machine	0.895	0.941
Random Forest	0.885	0.844
(Surface of urban height)		
Logistic regression	0.797	0.786
Support Vector Machine	0.859	0.890
Random Forest	0.878	0.848

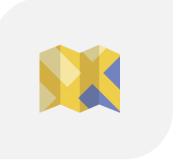








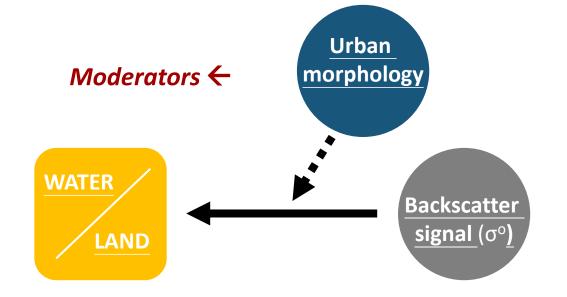
2) The relationship between urban morphology and double bounce effect



3) Spatial distribution of double bounce effect



2) The relationship between urban morphology and double bounce effect



Logistic regression:

• $\sigma^{o} \times$ neighboring surface morphological variables

$$\ln\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1\sigma^0 + \beta_2\sigma^0 \times horizontal_density + \beta_3\sigma^0 \times vertical_mean_height + \beta_4\sigma^0 \times vertical_height_variation$$

WATER / LAND

Backscatter value (σ°)

Interaction terms

 $\sigma^{o} \times urban surface morphology variables$





2) The relationship between urban morphology and double bounce effect

Backscatter value (σ^0)

Backscatter value(σ^0) × surface morphology variables

$$\ln\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1\sigma^0 + \beta_2\sigma^0 \times \text{Neighnoring horizontal density} + \beta_3\sigma^0 \text{Neighboring mean vertical height}$$

Variable	Coefficient	Standard error	z score	P(> z)		VIF
Intercept	-1.908	0.086	-22.074	$< 2 \times 10^{-16}$	***	
Backscatter	-0.357	0.012	-28.652	$<2\times10^{-16}$	***	1.41
Backscatter × Neighboring horizontal density	0.847	0.061	13.819	$< 2 \times 10^{-16}$	***	1.11
Backscatter × Neighboring mean vertical height HIGH	0.106	0.020	5.406	6.4×10^{-8}	***	1.34

Significance codes:

- ***: Significant at $\alpha = 0.001$
- **: Significant at $\alpha = 0.01$
 - *: Significant at $\alpha = 0.05$
 - . : Significant at $\alpha = 0.1$

- ① Main effect of σ^0
- ② Interaction terms
- **③** Total effect of backscatter value (σ^0)



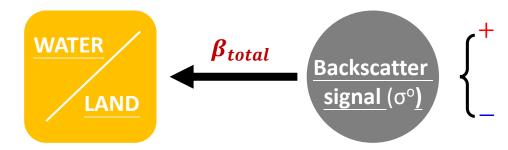
3) Spatial distribution of double bounce effect

> Spatial distribution of σ^0 's total effect

$$\ln\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1\sigma^0 + \beta_2\sigma^0 \times horizontal_density + \beta_3\sigma^0 \times vertical_mean_height + \beta_4\sigma^0 \times vertical_height_variation$$

$$= \beta_0 + (\beta_1 + \beta_2 \times horizontal_density + \beta_3 \times vertical_mean_height + \beta_4 \times vertical_height_variation) \times \sigma^0$$

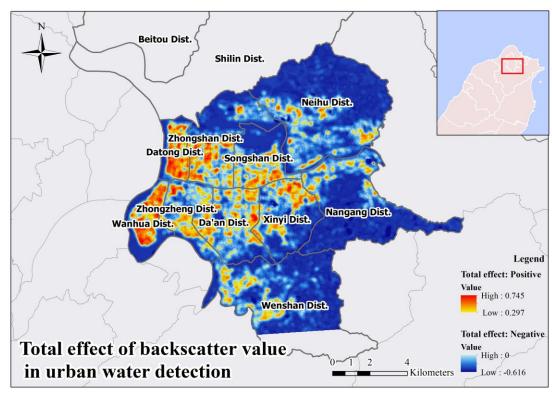
$$= \beta_0 + \beta_{total} \times \sigma^0$$



Different neighboring urban morphology,

Different β_{total} in space

 \rightarrow Spatial heterogeneity of σ°



- Part 1. Spatial difference of double bounce effect
- Part 2. Model performances

Brief summary of Part 1. Spatial difference of double bounce effect

Similar part:

- ① Low σ^0 implies water (Brivio et al., 2002; Henry et al., 2006)
- ② Double bounce effect (Franceschetti, Antonio & Daniele 2002; Ferro et al. 2011; Brunner et al. 2008)
 - Building width
 Road width

 Neighboring horizontal density
 - Building height vertical height
- → Match with small-scale studies using theoretical equations & empirical studies)

What we've further explored:

Double bounce effect is treated uniformly in urban areas... (Mason et al. 2014; Pulvirenti et al. 2016)

- Difference of water's backscatter and double bounce effect in space
- ② The dominance of neighboring horizontal density

Part2: Model performances



Inner validation:

- Model: Taipei City
- Data: Taipei City

- > Highest overall accuracy model
- > Characteristics of building or urban height



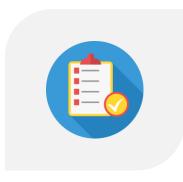
Cross-city validation:

Model: Taipei City

Data: Taichung City

> Generic models?

Part2: Model performances



Inner validation:

- Model: Taipei City
- Data: Taipei City

- > Highest overall accuracy model
- > Characteristics of building or urban height



Cross-city validation:

- Model: Taipei City
- Data: Taichung City

> Generic models?







Inner validation:

Model: Taipei City

Data: Taipei City

Pooled WATER

(25% inner WATER; 25% edge WATER; 50% LAND)



	Characteristics and patterns of buildings	Surface of urban height
Main models	(Overall accuracy; Recall)	(Overall accuracy; Recall)
Logistic regression model	0.852; 0.914	0.853; 0.909
SVM model	0.881; 0.923	0.890; 0.944
RF model	0.900; 0.965	0.905; 0.966

- Models: **RF model** has highest accuracy
- **Characteristics of buildings or Surface of urban height?**
 - Similar performances



Inner validation:

Model: Taipei City

Data: Taipei City

Testing data: Edge WATER

(50% edge WATER; 50% LAND)

From Null model to Main model:

• Overall accuracy: $0.78 \rightarrow 0.88$

● Detected edge water: 71% → 88%

	Characteristics and patterns of buildings	Surface of urban height
Null models	(Overall accuracy; Recall)	(Overall accuracy; Recall)
Thresholding	0.777; 0.707	0.777; 0.707
Main models		
Logistic regression model	0.775; 0692	0.792; 0.718
SVM model	0.840: 0.851	0.878, 0.881
RF model	0.867; 0.881	0.878 0.882

Part2: Model performances



Inner validation:

- Model: Taipei City
- Data: Taipei City

- > Highest overall accuracy model
- > Characteristics of building or urban height



Cross-city validation:

- Model: Taipei City
- Data: Taichung City

> Generic models?



Cross-city validation:

• Part 2. Model performances

• Part 1. Spatial difference of double bounce effect

Model: Taipei City

Data: Taichung City

Model	Overall Accuracy	Recall
Training data: Samples of Taipei city		
Null model		
Thresholding	0.868	0.919
Main models		False alarm: 0
(Characteristics and patterns of buildings)		
Logistic regression model	0.897	0.970
SVM model	0.853	0.980
RF model	0.878	0.842
(Surface of urban height)		`
Logistic regression model	0.902	0.973
SVM model	0.735	0.694
RF model	0.814	0.813

Testing data: Pooled WATER

(25% inner WATER; 25% edge WATER; 50% LAND)



Pooled WATER vs. LAND

(25% inner WATER; 25% edge WATER; 50% LAND)

From Null model to Main model:

- Overall accuracy: $0.868 \rightarrow 0.902$
- 98% of water can be detected

False alarm: 0.17; 0.28





Cross-city validation:

Model: Taipei City

Data: Taichung City

> Performance of specific WATER:

Testing data: Edge WATER

(50% edge WATER; 50% LAND)

Edge WATER vs. LAND

(50% edge WATER; 50% edge LAND)

From Null model to Main model:

• Overall accuracy: $0.82 \rightarrow 0.88$

● Detected edge water: 85% → 95%

Model	Overall Accuracy	Recall
Training data: Samples of Taipei city		
Null model		
Thresholding	0.815	0.846
Main models		False alarm: 0.30
(Characteristics and patterns of buildings)		
Logistic regression model	0.874	0.937
SVM model	0.832	0.951
RF model	0.804	0.800
(Surface of urban height)		
Logistic regression model	0.878	0.944
SVM-model	0.706	0.650
RF model	0.745	0.773

False alarm: 0.19; 0.29

Brief summary of Part2. Model performances



Data

to quantify urban morphology

- ① Recommend "DSM-DEM" for urban height
 - Calculation loading
 - Detailed building layer not available



Models

for water detection

- 2 Machine learning model has highest accuracy
 - Edge water's overall accuracy = 0.88 (*Null: 0.78*)
 - 93% of water is detected (*Null: 84%*)
 - 88% of edge water is detected (*Null: 71%*)
- 3 Generic models for cross-city applications
 - Edge water's overall accuracy = **0.88** (*Null: 0.82*)
 - **98%** of water is detected (*Null: 92%*)
 - 95% of edge water is detected (Null: 85%)

To sum up, when detecting urban water...

• From Part 1. Spatial difference of double bounce effect

Property of double bounce effects:

1. Intensity difference of water's double bounce effect in space

Mechanism of double bounce effects:

2. Importance of neighboring horizontal density

• From Part 2. Model performances

Data:

3. Advantage of "DSM-DEM" to summarize contribution from neighboring structures

Generic model / modeling framework:

- 4. Increased overall accuracy and percentage of detected water
- 5. Single-city data to capture radar's physical process of double bouncing

LIMITATIONS

- 1. The direction of SAR system and orientation of buildings
- 2. The **spatial extent** of double bounce effect
- 3. The **proportion** between WATER & LAND;

inner WATER & edge WATER

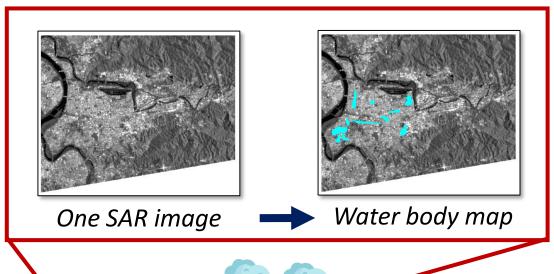


Remote sensing data and images + **GIS** techniques,

- > Detecting framework related to urban water/flood mapping
- A general approach to consider radar contributions from nearby structures
 - Without detailed data
- > Required data is all free and open
- ➤ Generic model for cross-city water detection
 - From one-city data

CONCLUSIONS (cont.)

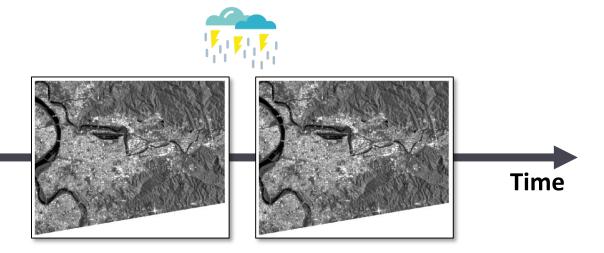
From this study:



Future applications:

Temporal scale of urban water detection

- Weather effects
- Additional water bodies
- Event-induced water bodies

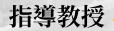




Thank you!

廖晧宇 Hao-Yu Liao





溫在弘 博士 Tzai-Hung Wen Ph. D.