第２０回 環境リモートセンシングシンポジウム

A New Unsupervised Classification of Collapsed Buildings Using TerraSAR-X Imagery, Hazard Distribution and Fragility Functions

Luis MOYA[1], Fumio YAMAZAKI[2], and Wen LIU[2]

[1] International Research Institute of Disaster Science, Tohoku University.
[2] Graduate School of Engineering, Chiba University.

February 15, 2018

Contents

- Remote sensing and machine learning.
- Introduction of the problem.
- A proposal solution.
- Case study.

2

Remote Sensing and Machine Learning

Extraction of features | Truth data

$\mathbf{z}_1 = (z_{1,1}, z_{1,2}, \dots, z_{1,M})$; $r_1 = 0$

$\mathbf{z}_2 = (z_{2,1}, z_{2,2}, \dots, z_{2,M})$; $r_2 = 1$

…

$\mathbf{z}_K = (z_{K,1}, z_{K,2}, \dots, z_{K,M})$; $r_K = 0$

$$Z = \{\mathbf{z}_i, r_i\}_{i=1}^{K}$$

⬇

Calibrate a discriminant function

Source of images: Prof. Fumio Yamazaki

3

A Machine Learning Method: Logistic Regression

Logistic function: $y = P(C_1|\mathbf{z}) = \dfrac{1}{1 + e^{-\boldsymbol{\theta}^T \mathbf{z}}}$ ➡ if $y \geq 0.5 \rightarrow \mathbf{z} \in C_1$; if $y < 0.5 \rightarrow \mathbf{z} \in C_2$

Likelihood of $\boldsymbol{\theta}$ given training set Z: $l(\boldsymbol{\theta}|Z) = \prod_{i=1}^{K} (y_i)^{r_i}(1-y_i)^{1-r_i}$

Cost function: $E(\boldsymbol{\theta}|Z) = -\sum_{i=1}^{K} r_i \ln y_i + (1-r_i)\ln(1-y_i)$

Key observation:

$$\boxed{\begin{array}{l} \text{if } r_i = 0 \ \rightarrow \ r_i \ln y_i = 0 \\ \text{if } r_i = 1 \ \rightarrow \ (1-r_i)\ln(1-y_i) = 0 \end{array}}$$

A training sample always cancel one term!

4

Introduction of the Problem

Some succeeded cases studies:

Research team	Method	Date of the event	Training data source	Date the training data was released
Bai et al. (2017)	Deep learning	March 11, 2011	MLIT	August 4, 2011
Wieland et al. (2016)	SVM	March 11, 2011	MLIT	August 4, 2011
Moya et al. (2018)	SVM	April 16, 2016	Field survey	Field survey: June 10-13, 2016

Can you see the common issue?

*After a large-scale disaster event, it is challenging to gather information to be used as **training data***

5

A Proposal Solution

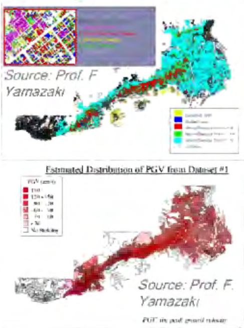
6

Lets Replace Training Data With Other Source of Information!

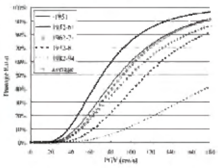
Researchers have been observing the relation between damage and hazard for several decades!

Building damage inventory (Source: BRI)

Peak ground velocity (Yamaguchi and Yamazaki, 2001)



Building damage function



7

Back to Logistic Regression

To following modification is proposed:

Cost function:

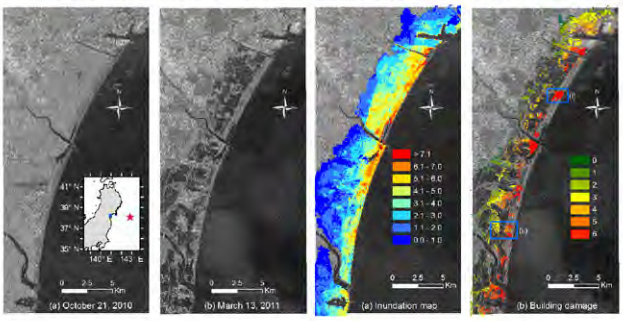
$$E(\boldsymbol{\theta}|Z) = -\sum_{i=1}^{K} \boldsymbol{p_i} \ln y_i + (1 - \boldsymbol{p_i}) \ln(1 - y_i)$$

p_i : probability that the sample i is collapsed given the demand of the hazard (estimated from the fragility curve)

Interpretation: Consider that every sample is duplicated. One as collapsed and the other as non-collapsed. However, their contribution to the cost function is weighted (p_i and 1-p_i).

8

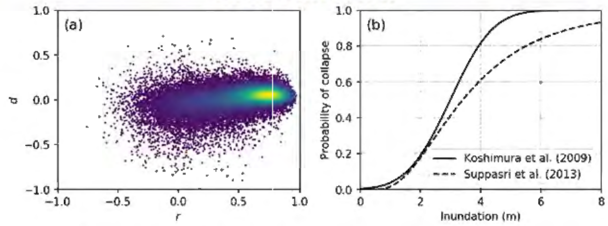
Case Study: The 2011 Tohoku Earthquake and Tsunami



October 21, 2010 Pre-event	March 13, 2011 Post-event	Inundation map	Building damage surveyed

9

Case Study: The 2011 Tohoku Earthquake and Tsunami



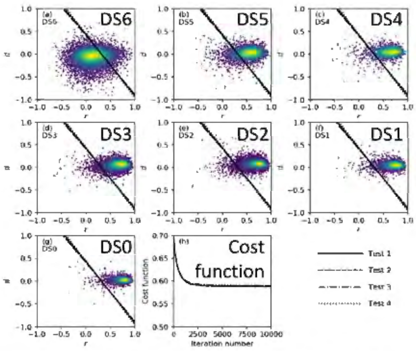
Two features: averaged difference and correlation coefficient (Bi-dimensional space)

Empirical fragility curve
Demand: inundation depth

10

Results: Linear Discriminant Function

DS: Damage state
C: collapsed bldg.
NC: non-collapsed bldg.
UA: User accuracy
PA: Producer accuracy

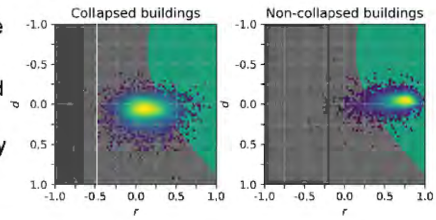


	DS0	DS1	DS2	DS3	DS4	DS5	DS0-DS5	DS6	Total	UA
NC	1793	2448	6264	5551	1475	2844	**20375**	**1861**	**22236**	**91.6**
C	64	130	385	469	220	788	**2056**	**6943**	**8999**	**77.2**
Total	1857	2578	6649	6020	1695	3632	**22431**	**8804**	**31235**	
PA	96.6	95.0	94.2	92.2	87.0	78.3	**90.8**	**78.9**		**87.5**

Results: Non-linear Discriminant Function

$$\boldsymbol{\theta z} = \theta_0 + \theta_1 z_1 + \theta_2 z_2 + \theta_3 z_1^2 + \theta_4 z_2^2$$

DS: Damage state
C: collapsed bldg.
NC: non-collapsed bldg.
UA: User accuracy
PA: Producer accuracy



	DS0	DS1	DS2	DS3	DS4	DS5	DS0-DS6	DS6	Total	UA
NC	1788	2419	6205	5492	1444	2738	**20086**	**1612**	**21698**	**92.6**
C	69	159	444	528	251	894	**2345**	**7192**	**9537**	**75.4**
Total	1857	2578	6649	6020	1695	3632	**22431**	**8804**	**31235**	
PA	96.3	93.8	93.3	91.2	85.2	75.4	**89.5**	**81.7**		**87.3**

12

CONCLUSIONS

- Avoiding training data for identification of building damage is crucial.
- A new unsupervised classification method for damage classification is proposed.
- Training data is replaced with aggregate information from fragility functions and hazard information.
- High accuracy was obtained in the detection of collapsed buildings due to the 2011 Tohoku Earthquake and Tsunami.

13

ありがとうございます

Thanks