Spatial Prediction and Laws of Geography

A-Xing Zhu

Department of Geography
University of Wisconsin-Madison

Outline

1. SP and Laws of Geography

2. New Demands for SP

3. Third Law of Geography and SP

4. Case Studies

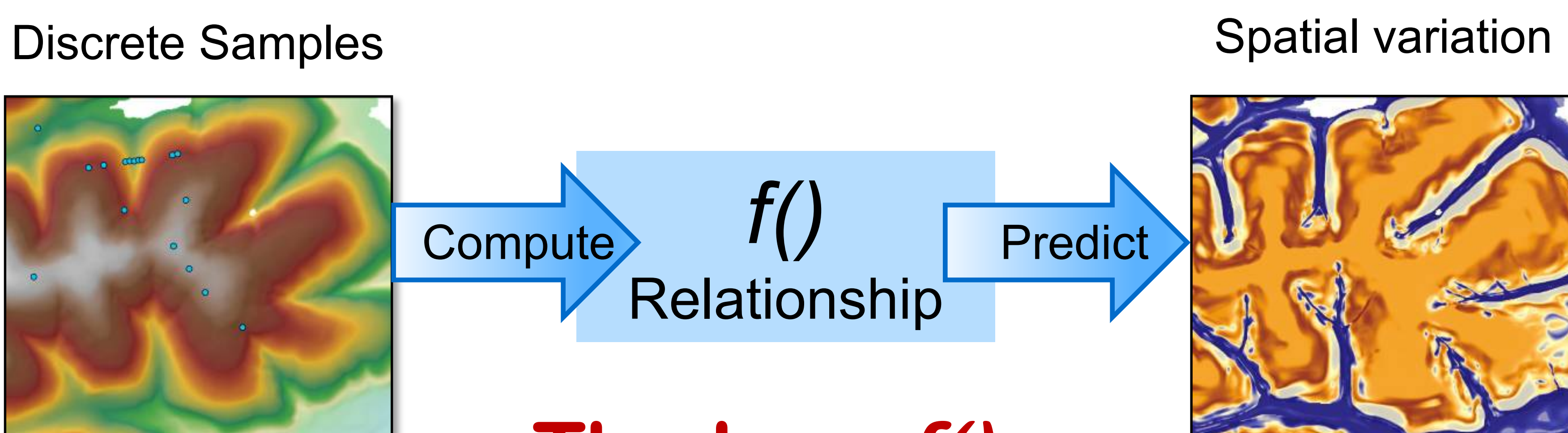
5. Implications

SP and Laws of Geography

Basic Idea

Spatial Prediction: Estimate what we do not know over space using what we have already known. Basic steps:

1) Compute the relationship between the target geographic variable and what we have already known;
2) Estimate the spatial variation of the target variable over study area



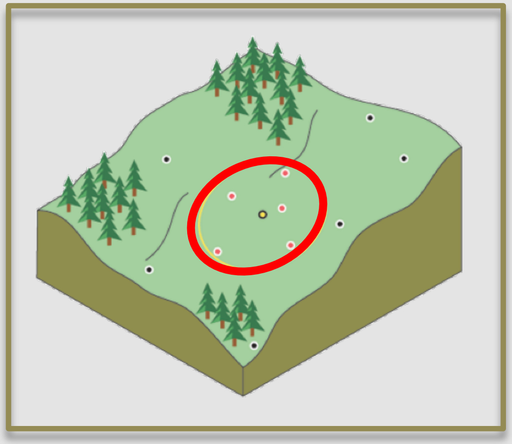
The key: *f()*

SP and Laws of Geography

Existing Theories

First Laws of Geography:
"Everything is related to everything else, but near things are more related to each other" (Tobler, 1970)

Spatial autocorrelation
of the target variable
(Spatial relationship)

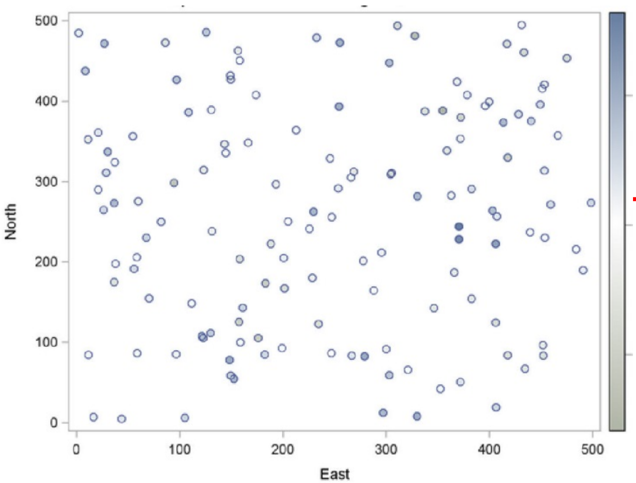


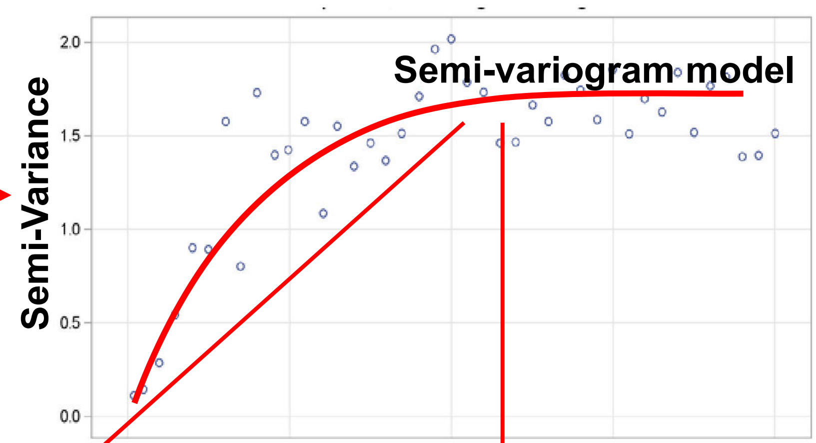
The idea: Extract spatial autocorrelation from a large set of well designed discrete samples, and then apply the relationship to the whole area (Krige, 1951; Collins,1999; Goovaerts, 1999).

Methods: The Kriging family of methods (ordinary kriging, universal kriging)

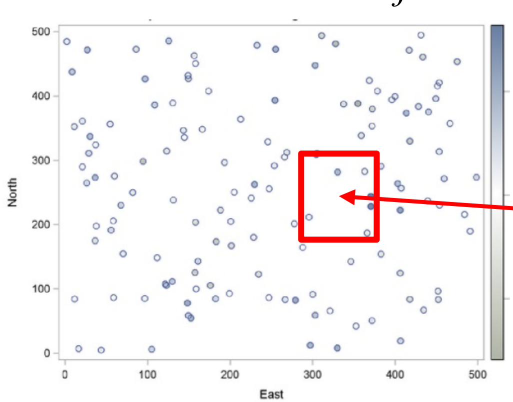
SP and Laws of Geography

Spatial Autocorrelation – Ordinary Kriging Method





$$\tilde{\sigma}_R^2 = \tilde{\sigma}^2 + \sum_{i=1}^{n_0}\sum_{j=1}^{n_0} w_i w_j (\tilde{\sigma}^2 - \tilde{\gamma}_{ij}) - 2\sum_{i=1}^{n_0} w_i (\tilde{\sigma}^2 - \tilde{\gamma}_{i0}) + 2\mu\left(\sum_{i=1}^{n_0} w_i - 1\right)$$



$$\hat{Z}_0 = \sum_{i=1}^{n} w_i \bullet z_i$$

Semi-variogram model:

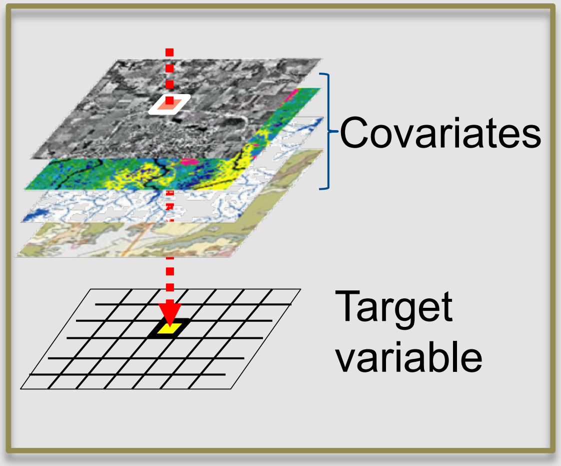
1) Able to represent the area (well distributed and sufficient samples)
2) Stable over space (Stationary)

SP and Laws of Geography

Existing Theories

Statistical Theories： If two variables exhibit relationship, this relationship can be statistically quantified and used.

Correlation between target variable and auxiliary variables which co-vary (covariates)

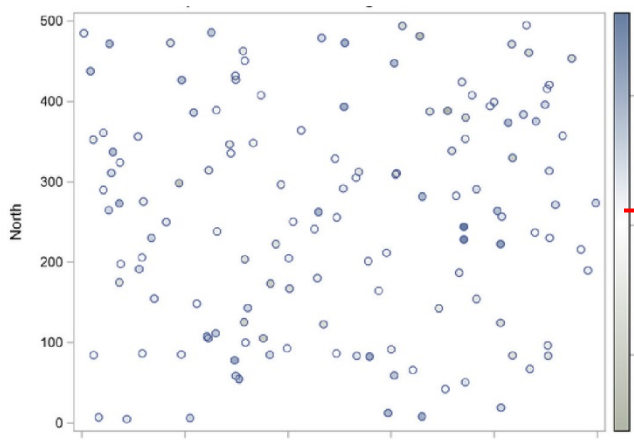


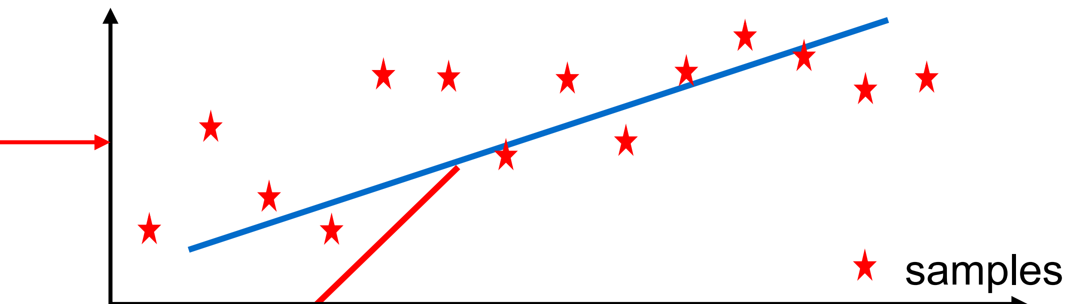
The idea: Extract correlation with covariates from a large set of well designed discrete samples, and then apply the relationship to the whole area (Croxton et al., 1968; Zhu et al., 2014)

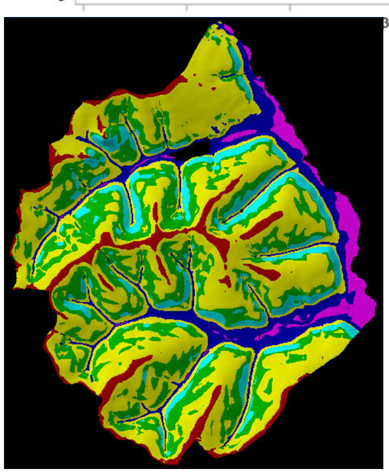
Methods: Regression, machine learning techniques, expert systems

SP and Laws of Geography

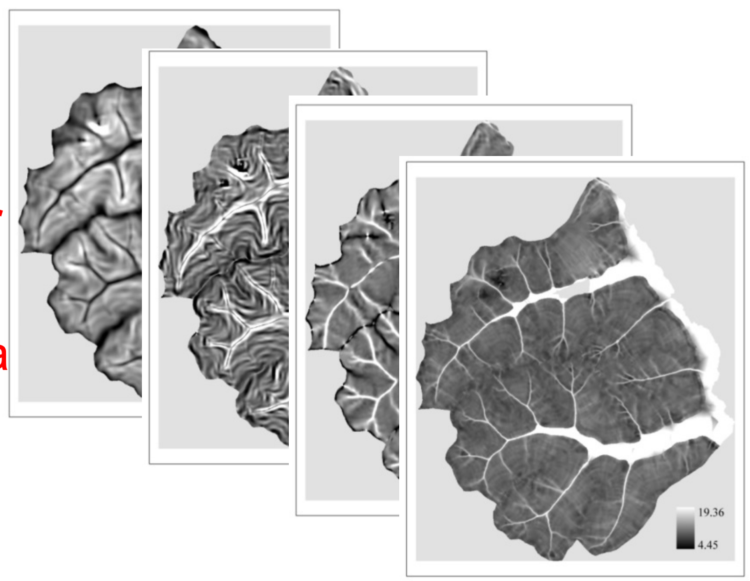
Statistical Correlation – Regression method







$$\hat{Z} = f(x)$$



Correlation:

1) Able to represent the area
(well distributed
sufficient samples)
2) Stable over space
(Stationary)

SP and Laws of Geography

Summaries on the above two theories

First Law of Geography
(Spatial autocorrelation)

Statistical Theories
(Correlation)

Require

The obtained relationships:

1) Able to represent the area
(well distributed
sufficient samples)

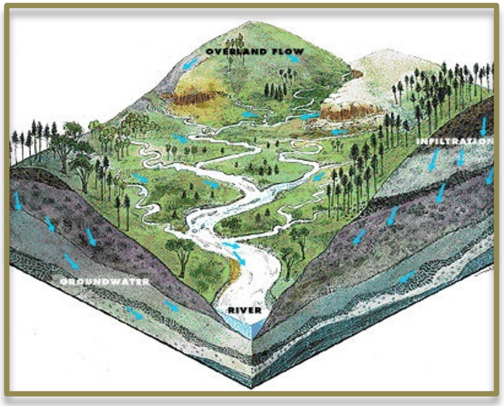
2) Stable over space
(Stationary)

SP and Laws of Geography

Second Law of Geography (Goodchild, 2004): "geographic variables exhibit uncontrolled variance", **Spatial Heterogeneity**

Therefore, the spatial autocorrelation and/or the correlation are also "uncontrolled" vary over space, particularly over large or complex geographic areas

Large areas



Complex areas

Thus, the requirements imposed by the First Law as well as the statistical theories would not be met, over large areas or complex areas.

SP and Laws of Geography

The Impacts of Second Law of Geography

To increase the accuracy of spatial prediction, scholars revised existing methods based on the Second Law (Haas, 1990; Cressie, 1993; Yamamoto, 2000; Fotheringham et al., 2002）。

Spatial Autocorrelation Methods
(First Law)

Correlation Methods
(Statistical Theories)

Spatial Heterogeneity
(Second Law)

Within class or local kriging

Geographic weighted regression

Limiting the spatial scope or requiring even more samples

SP and Laws of Geography

Summary

Spatial Autocorrelation (First Law)

Spatial Heterogeneity (Second Law)

Variable Correlation (Statistical Theories)

require

Large sample set, representative, stationary, limited areas

Would these characteristics of spatial prediction meet the emerging needs of geographic analysis and modeling?

Outline

1. SP and Laws of Geography

2.New Demands for SP

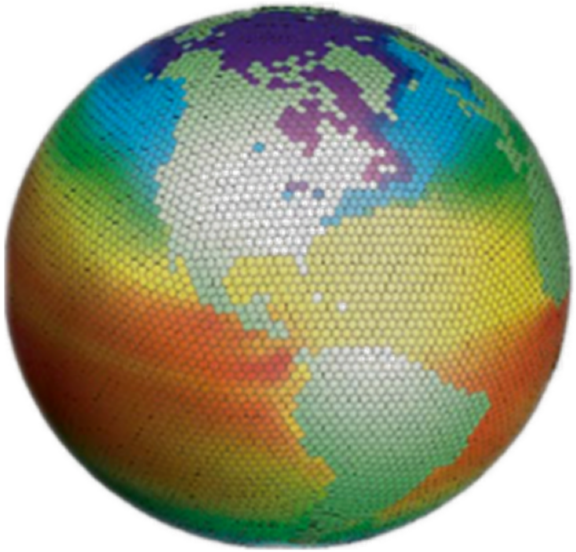
3.Third Law of Geography and SP

4. Case Studies

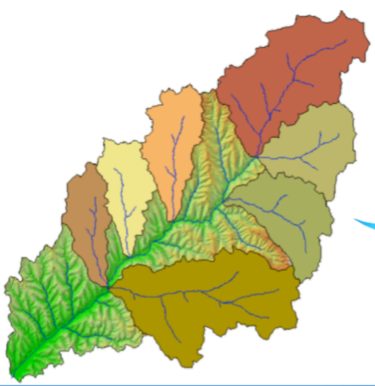
5. Implications

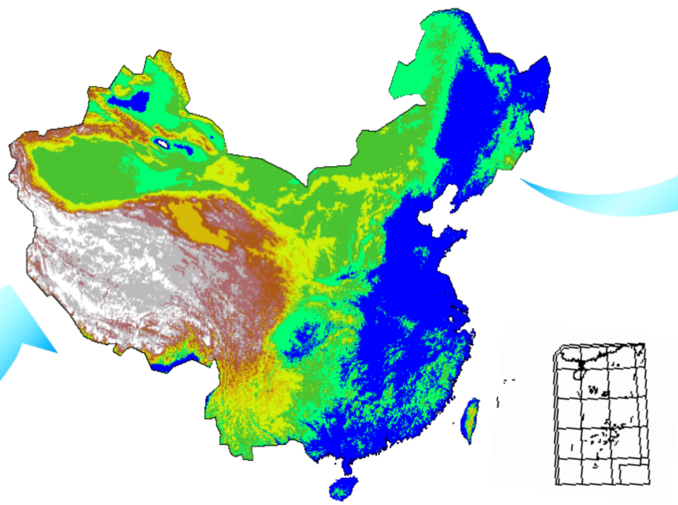
New Demands for Spatial Prediction

Changes in demand



The need for information on spatial variation of geographic variables expands from local watersheds, to continental, even to global scales.



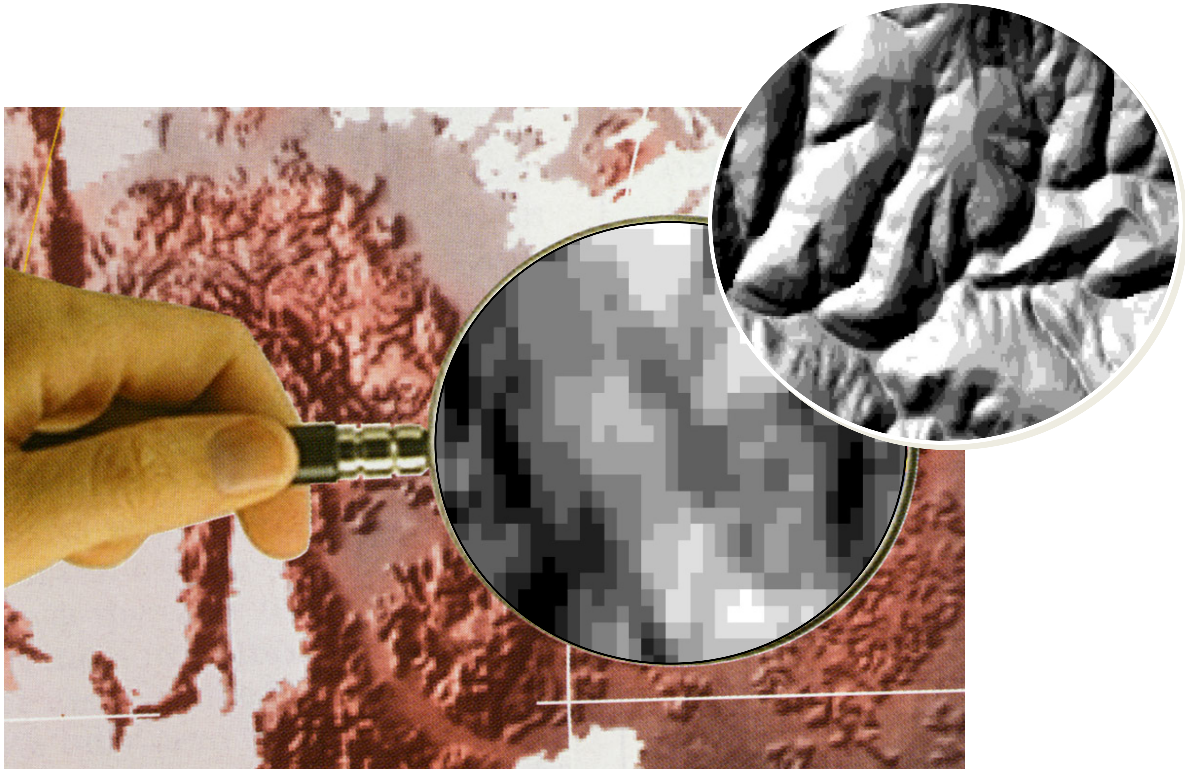


New Demands for Spatial Prediction

Changes in demand

Expansion in geographic scope

In addition to expansion of spatial scope, demand for accuracy and spatial details are increasing.



High spatial resolution
High attribute accuracy

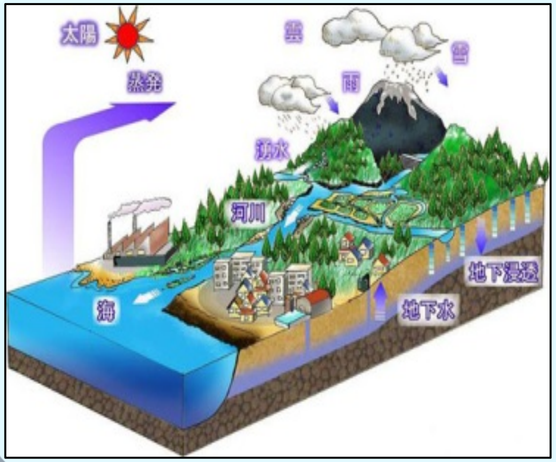
Detail spatial information

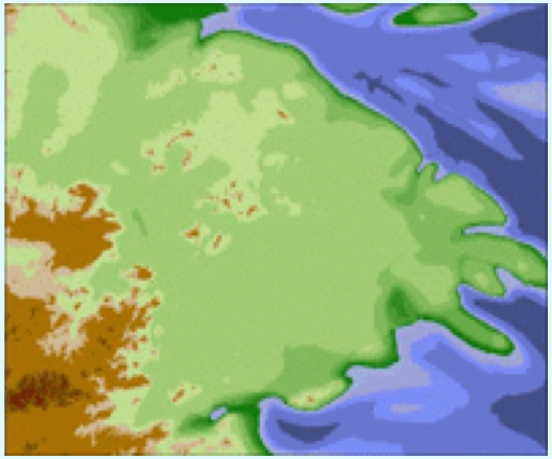
New Demands for Spatial Prediction

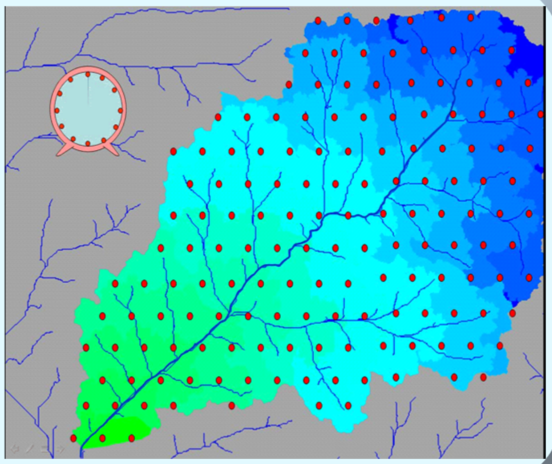
Changes in demand

- Expansion in geographic scope
- Increase in accuracy and details

Many geographic analysis and modeling now require the assessment of reliability or sensitivity of results which requires the provision of uncertainty of input data





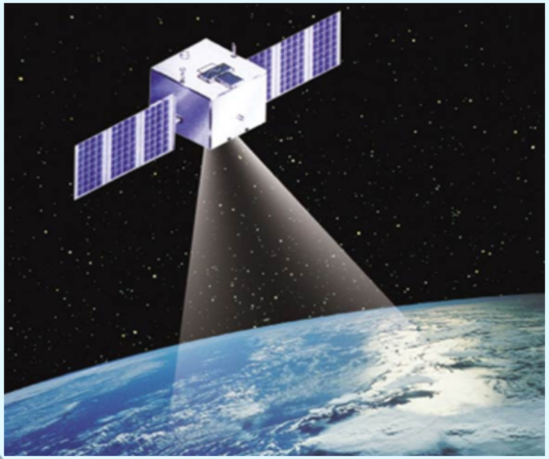


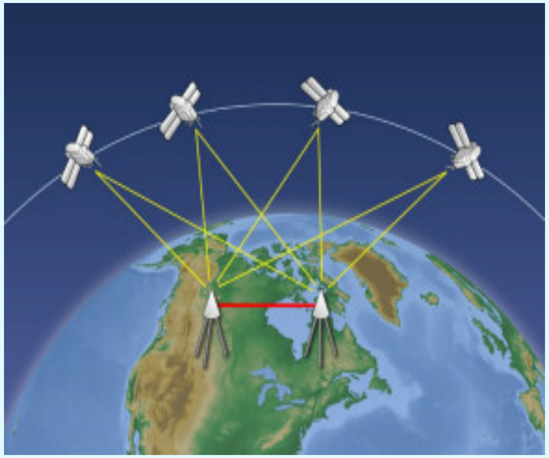
New Demands for Spatial Prediction

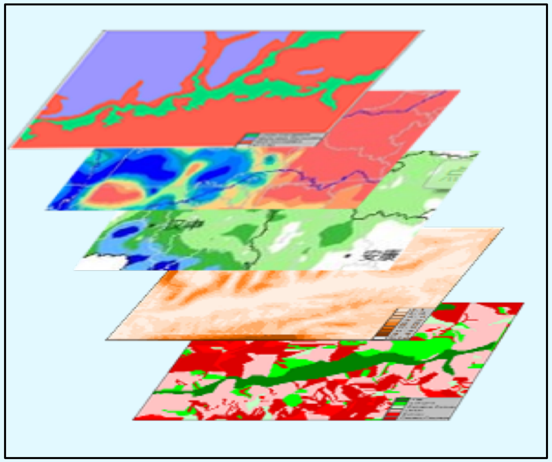
Provision of new data

Rapid development of Earth observation systems and other sensor networks provide a rich set of spatial data which can be used as covariates but not available before.

Vast amount of **auxiliary** data







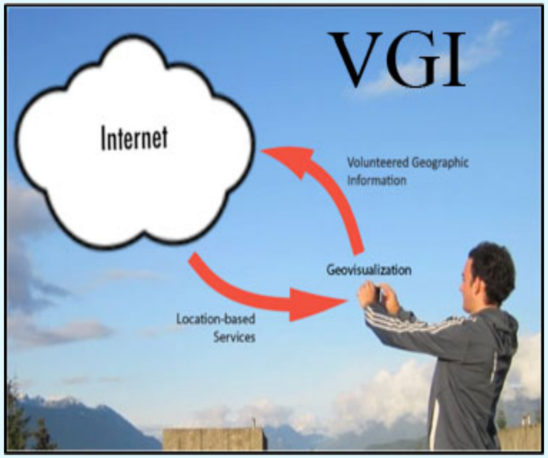
New Demands for Spatial Prediction

Provision of new data

Location based services dramatically increase the availability of data (samples) from citizens (non-specialists). This provides inexpensive samples for geographic analysis but these samples are ad-hoc in nature.

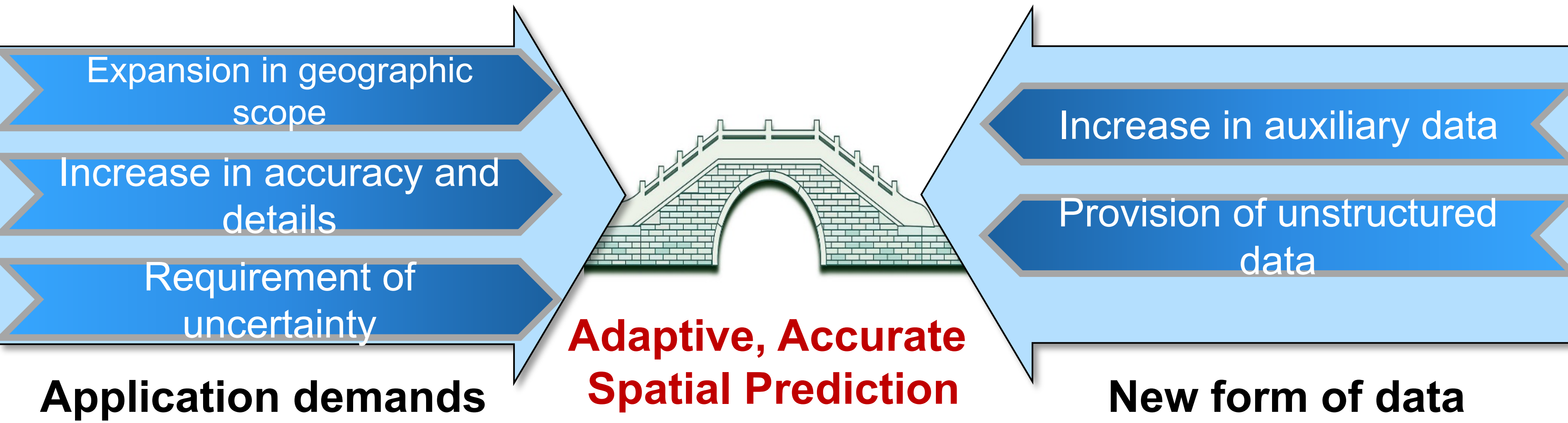
Vast amount of **auxiliary** data

Provision of unstructured data



New Demands for Spatial Prediction

A new bridge needed



A bridge which is able to make use of the rapid increasing inexpensive unstructured data on one hand and meet the needs of application demands on the other hand.

New Demands for Spatial Prediction

Challenge

Spatial Autocorrelation (First Law)

Spatial Heterogeneity (Second Law)

Variable Correlation (Statistical Theories)

require

Large sample set, representative, stationarity, limited areas

Thus, methods based on existing theories cannot make a good use of the unstructured data and do not meet the need of modern geographical modeling/analysis over large geographic areas.

Expansion in geographic scope

Increase in accuracy and details

Requirement of uncertainty

Application demands

Existing Methods

Increase in auxiliary data

Provision of unstructured data

New form of data

The grand challenge

Scientific Question：

How do we mitigate the deficiencies of the existing theories, overcome the rigid requirements of existing methods, make effective use of new form of data, so that spatial prediction at high accuracy and high spatial detail over large and complex geographic areas is possible?

Outline

1. SP and Laws of Geography

2. New Demands for SP

3. Third Law of Geography and SP

4. Case Studies

5. Implications

Third Law and Spatial Prediction

Third Law of Geography

In addition to spatial autocorrelation (First Law) and spatial heterogeneity (Second Law), there is another important and widely used geographic principle (**Third Law of Geography**):

> **The more similar of geographic configuration the more similar the attribute values and the processes. Geographic Similarity**

Geographic configuration: the **collection and the structure** of covariates of the target variable at a location.

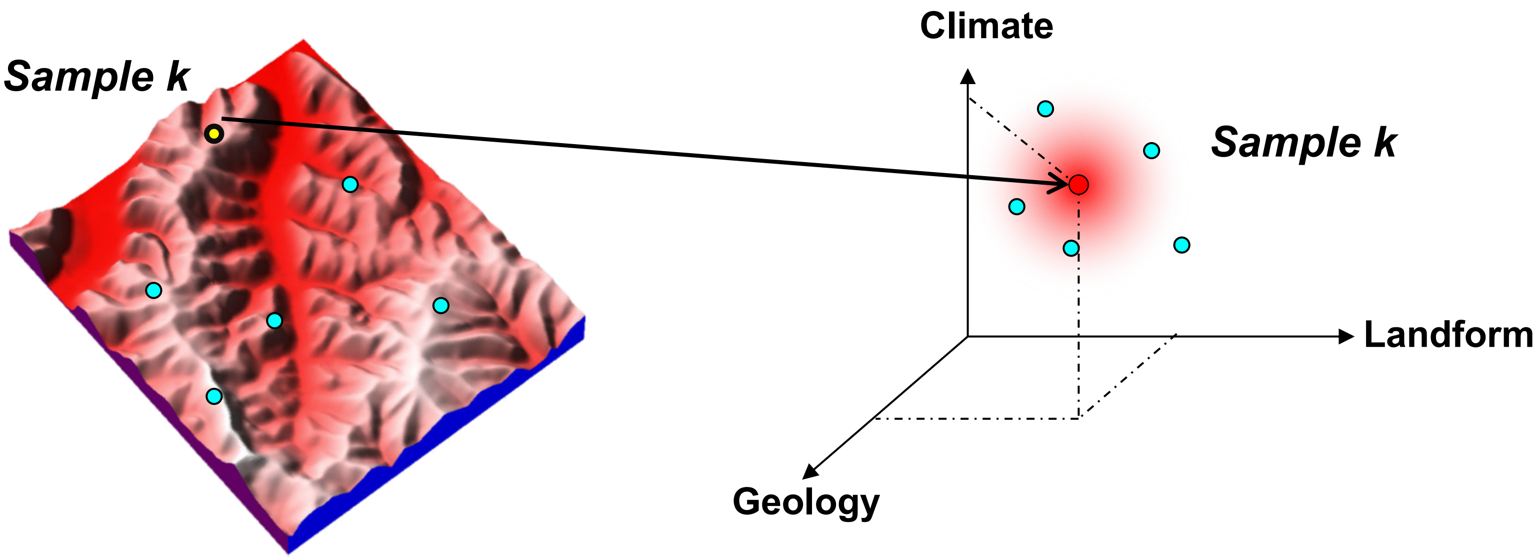
Geographic similarity: the similarity in the geographic configuration between two locations.

Third Law and Spatial Prediction

How can it be used?

Existing approaches quantify the relationship from the entire sample set. The Third Law allows us to explore the use of the representativeness of **a single sample**. A sample can represent locations which are "**similar in geographic configuration**".

A soil sample example:



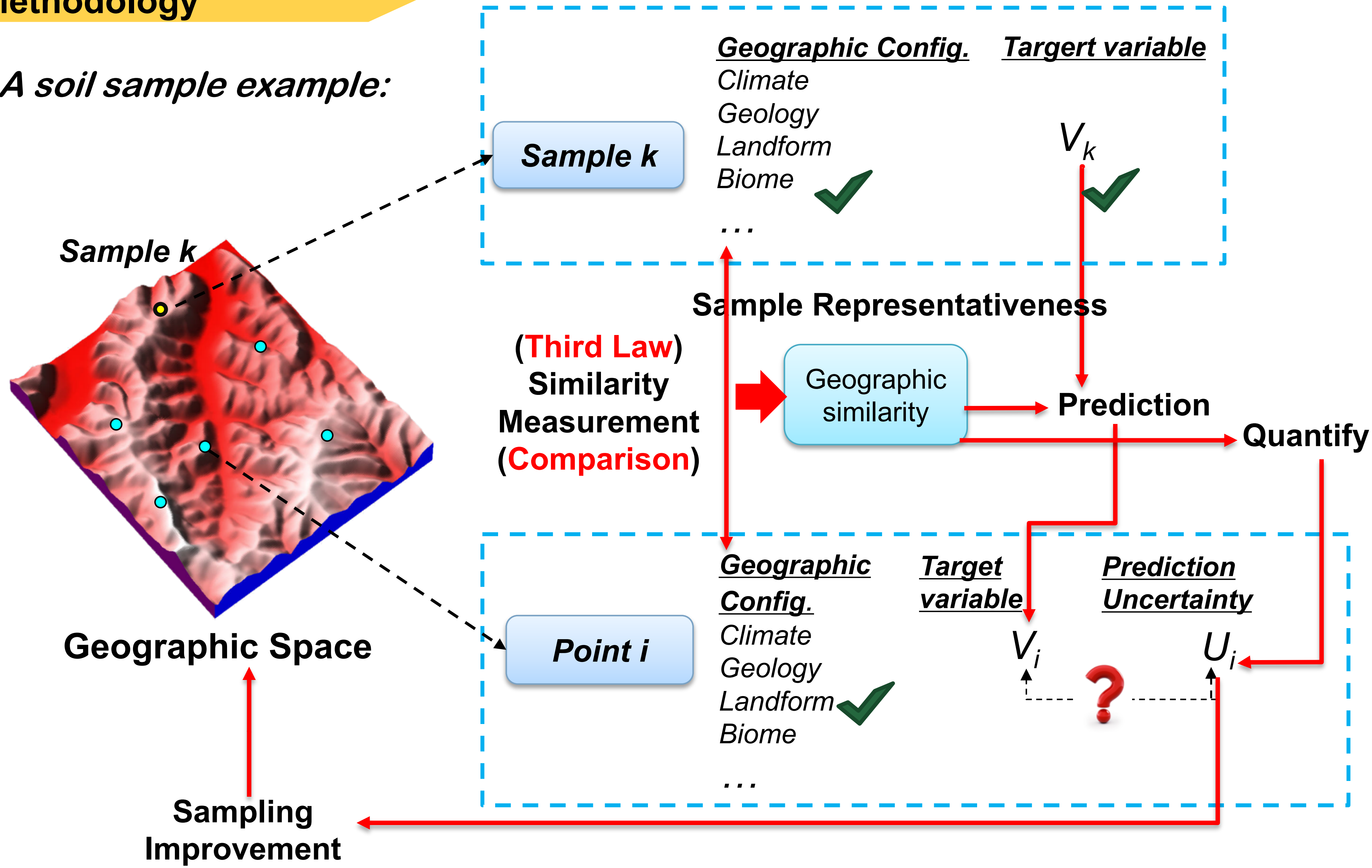
Geographic space

Covariate (feature) space

Third Law and Spatial Prediction

Methodology

A soil sample example:



Outline

1. SP and Laws of Geography

2. New Demands for SP

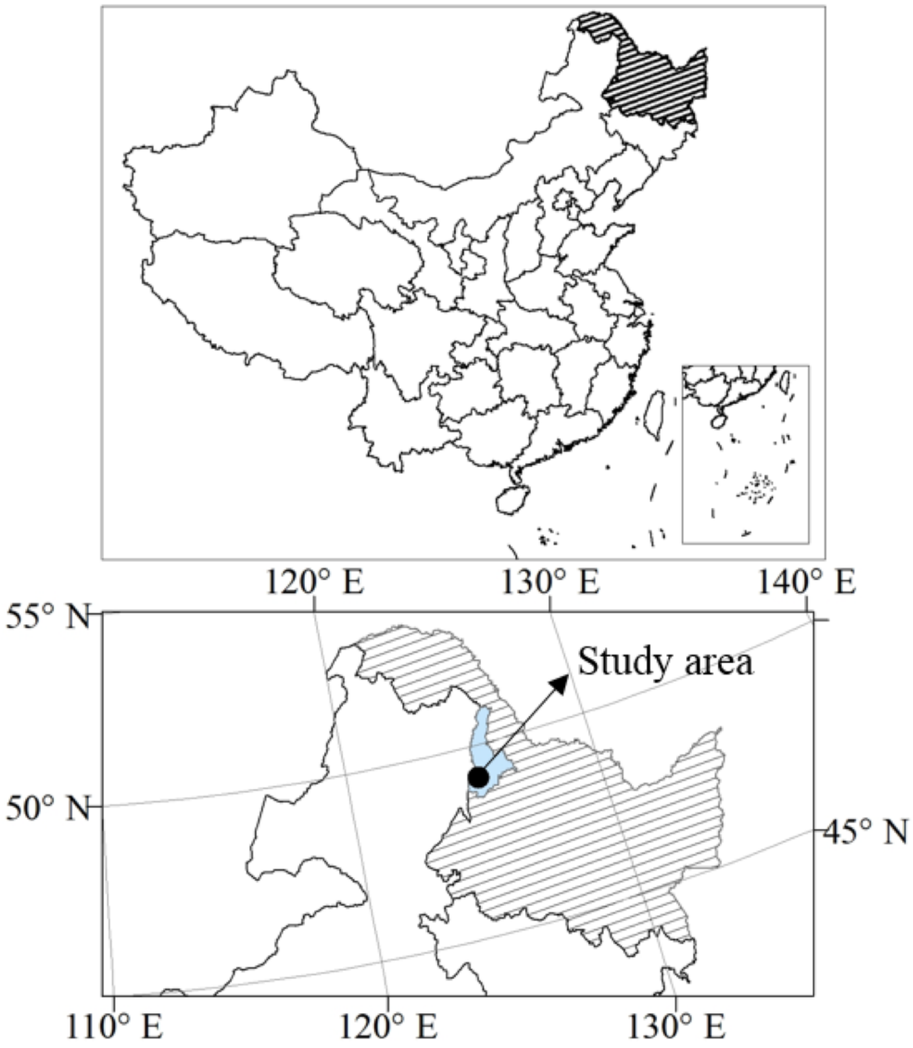
3. Third Law of Geography and SP

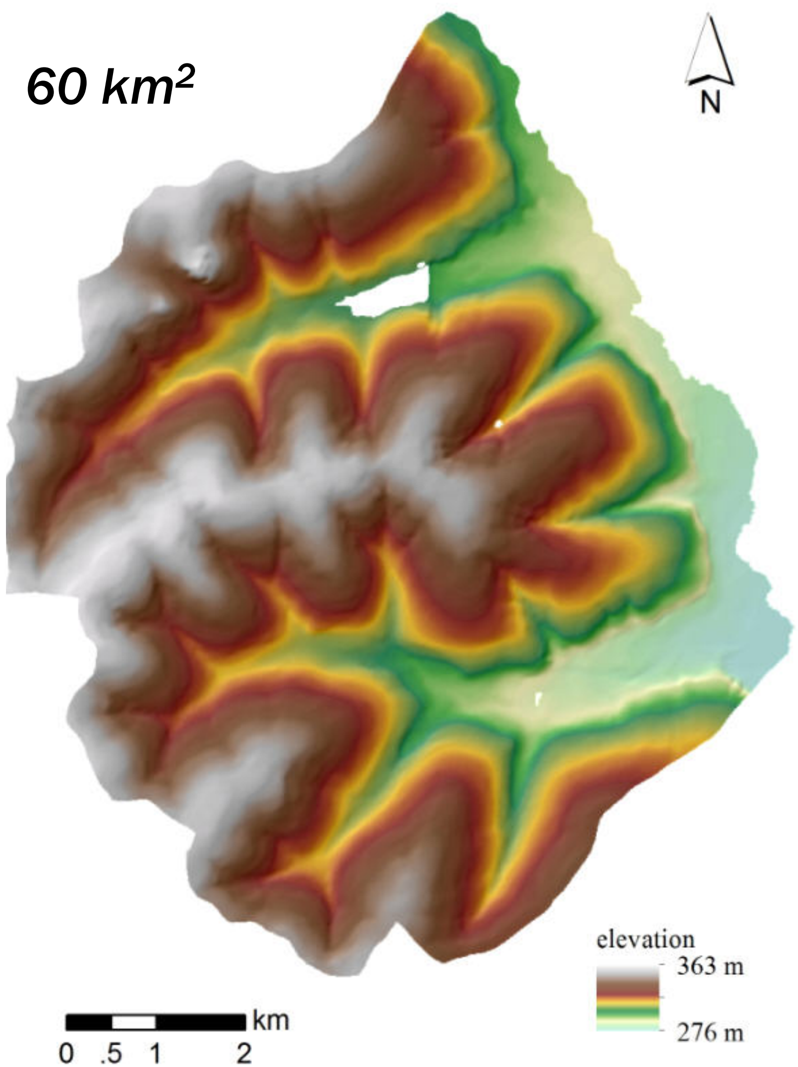
4. Case Studies

5. Implications

Case Study 1: Prediction – soil mapping

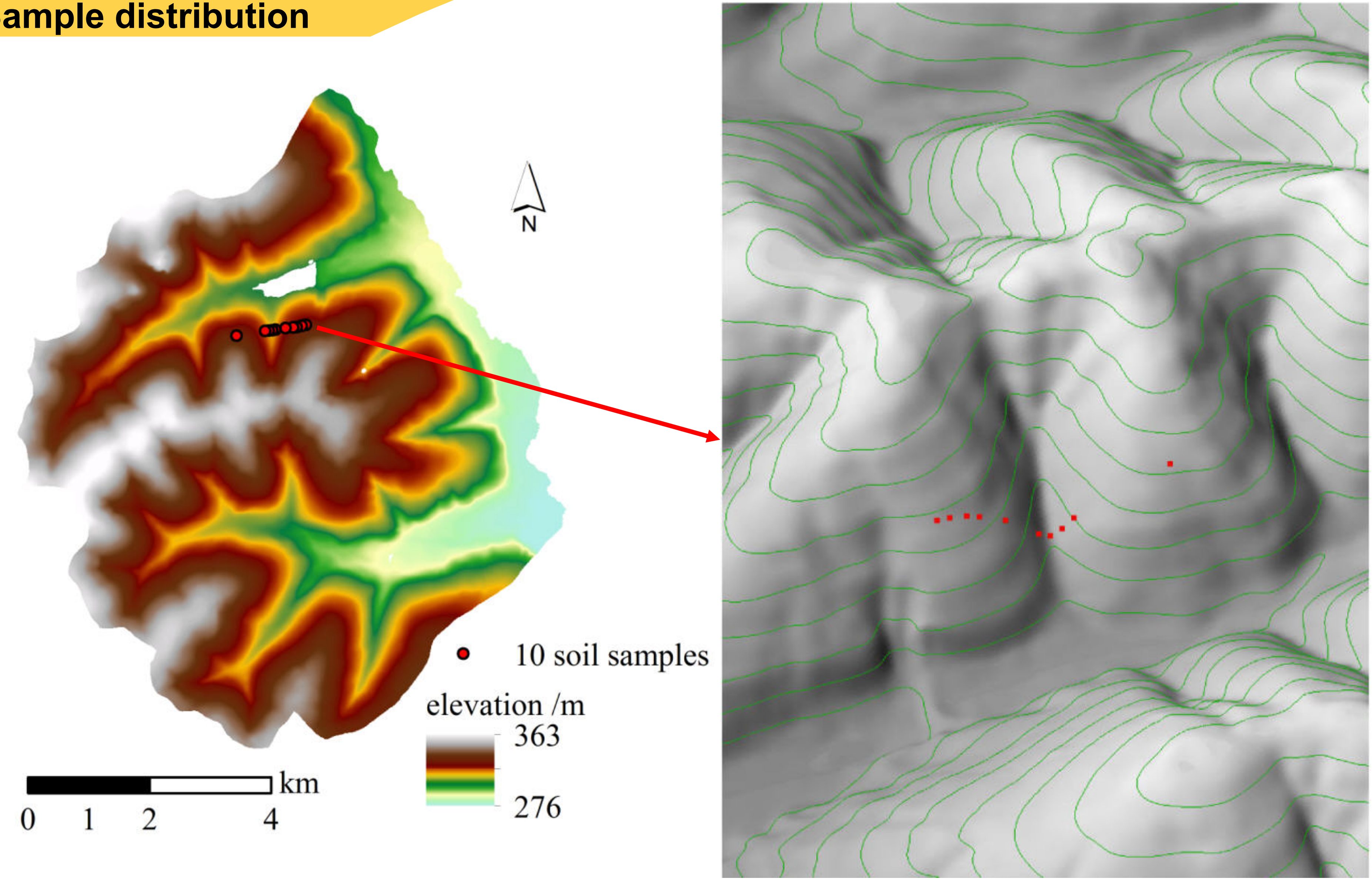
Study area





Case Study 1: Prediction – soil mapping

Sample distribution



Case Study 1: Prediction – soil mapping

Experiment design

Target variable: top soil organic matter content (SOM,%)

Covariates: elevation(e), planform curvature(pl), profile curvature(pr), slope gradient(g), relative position(rp), topographic wetness index(wet)

Models for comparison:

Regression: (statistical, relationship from the entire sample set)

$$SOM_{ij}^{R} = f(\mathrm{e}_{\mathrm{ij}}, \mathrm{pl}_{\mathrm{ij}}, \mathrm{pr}_{\mathrm{ij}}, \mathrm{g}_{\mathrm{ij}}, \mathrm{rp}_{\mathrm{ij}}, \mathrm{wet}_{\mathrm{ij}})$$

Regression-Kriging: (stat and First Law, relationship from entire sample set)

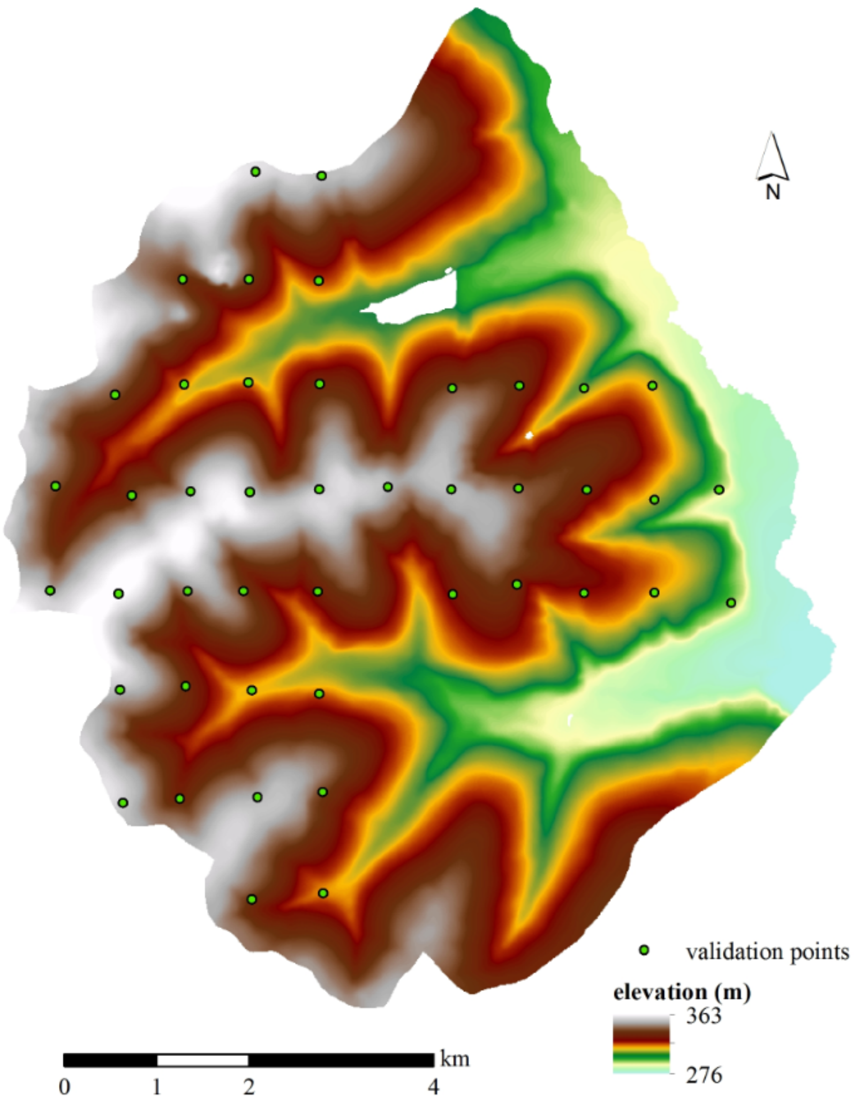
$$SOM_{ij}^{RK} = SOM_{ij}^{R} + SOM_{ij}^{K}$$

iPSM: (Third Law, individual sample comparison)

Case Study 1: Prediction – soil mapping

Validation samples

44 regularly distributed samples



Case Study 1: Prediction – soil mapping

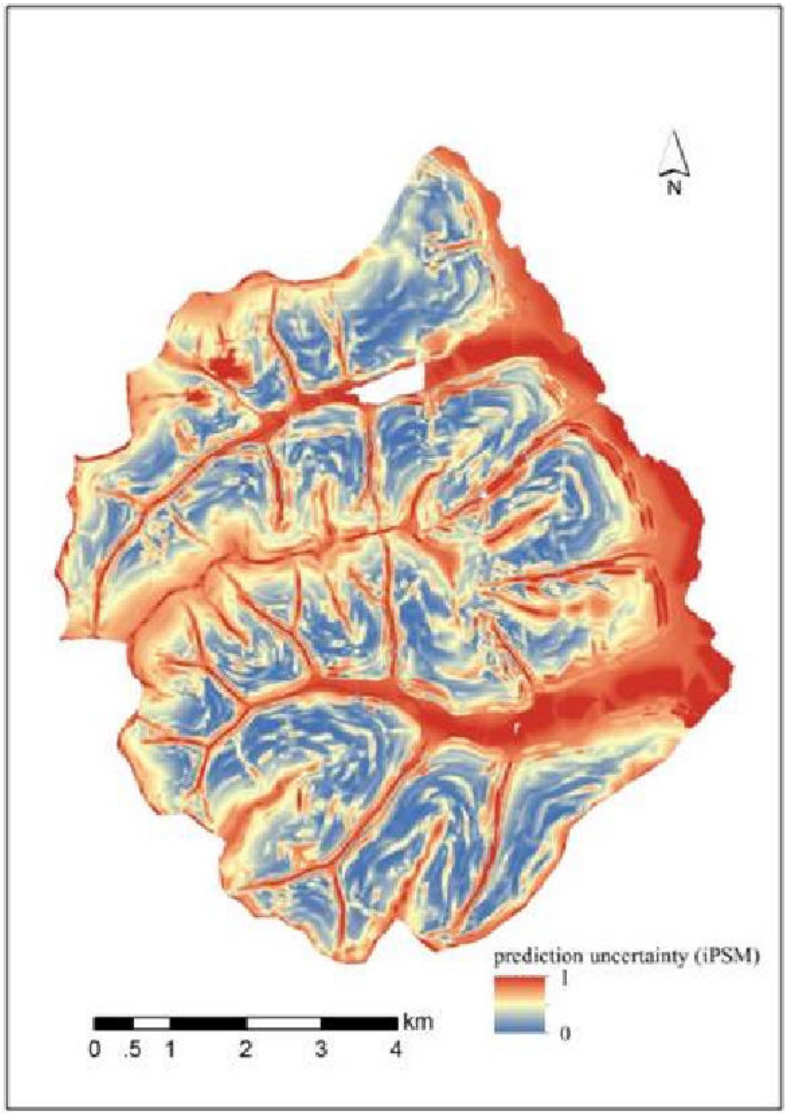
Result comparison

Method	RMSE	MAE
iPSM (Third Law)	1.195	1.009
RK (First Law)	6.095	3.012
LR (Statistical)	6.096	3.012

iPSM does not require samples to be over certain size, nor of specific spatial distribution.

Case Study 2: Uncertainty – soil mapping

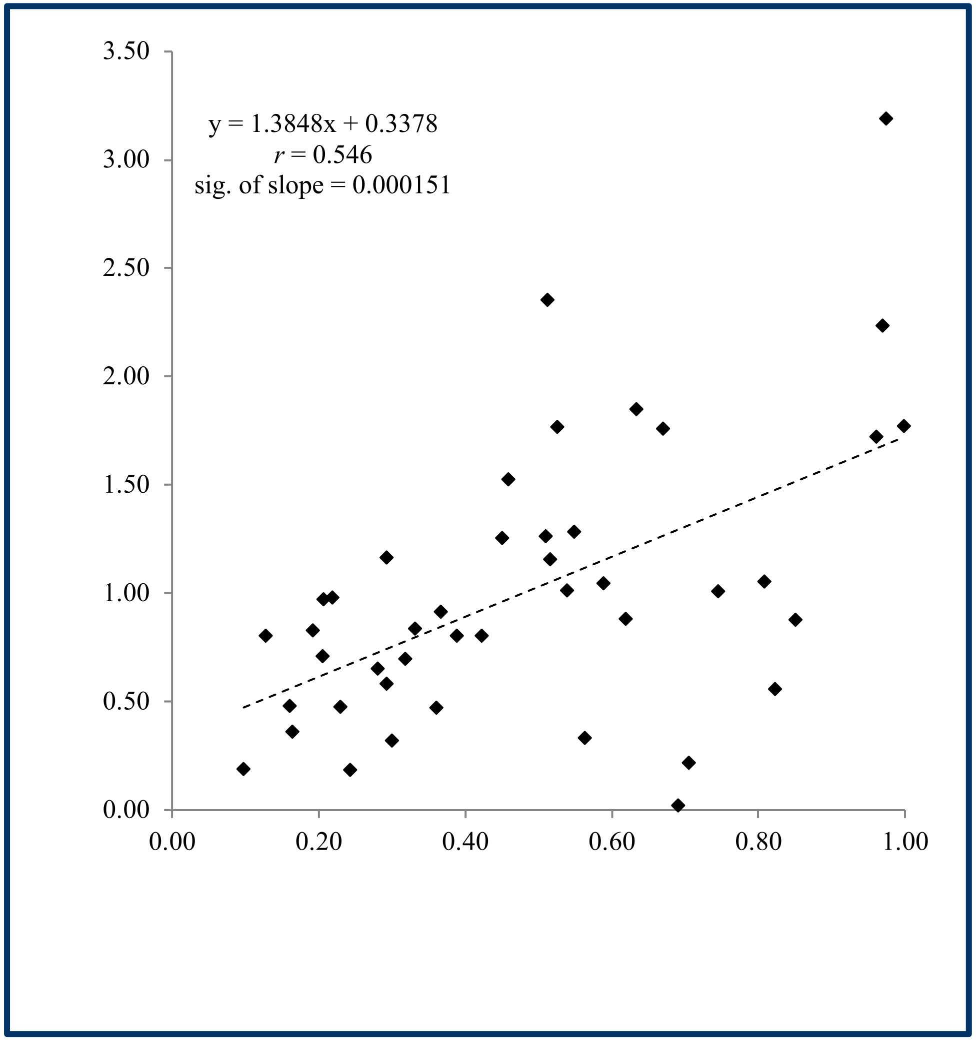
Map of uncertainty



Prediction uncertainty

Case Study 2: Uncertainty – soil mapping

Uncertainty and accuracy



Relationship between residual and uncertainty

Prediction uncertainty is indicative of prediction accuracy.

Outline

1. SP and Laws of Geography

2. New Demands for SP

3. Third Law of Geography and SP

4. Case Studies

5. Implications

Summary and Implications

Under the Third Law, spatial prediction:

1) Does not require samples to be over particular size;
2) Does not require samples to be of specific distribution;
3) Does not require relationships to be stationary;
4) Does provide uncertainty which is effective in
   a) indicating spatial variation of accuracy;
   b) allocating effective sampling efforts;
   c) mitigating spatial bias in citizen data

Implication:

Potentially, a complete new way for spatial prediction?
a transformative way?

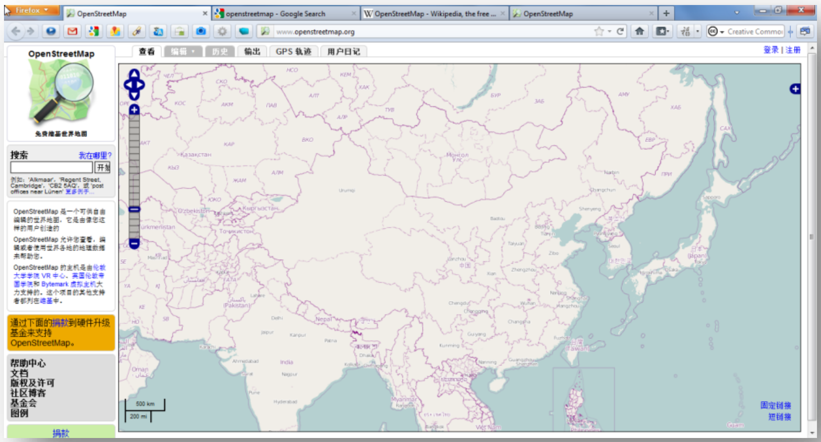
Thank you for your attention!

谢谢！

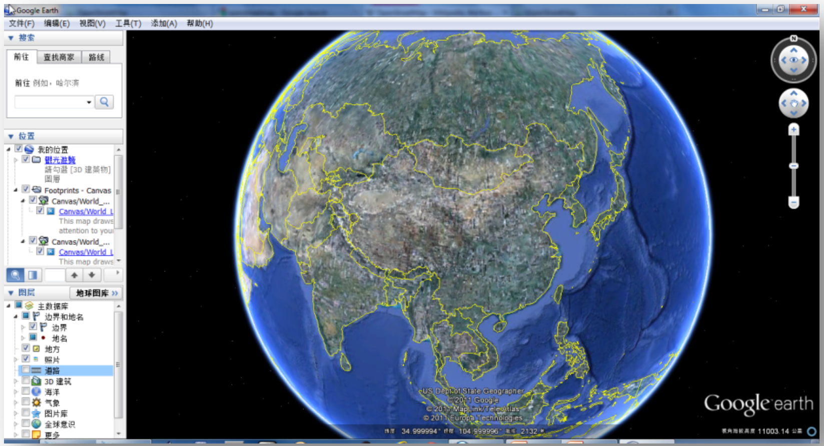
WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

Case Study 3: New data – habitat mapping

Sampling bias (citizen data)



OpenStreetMap



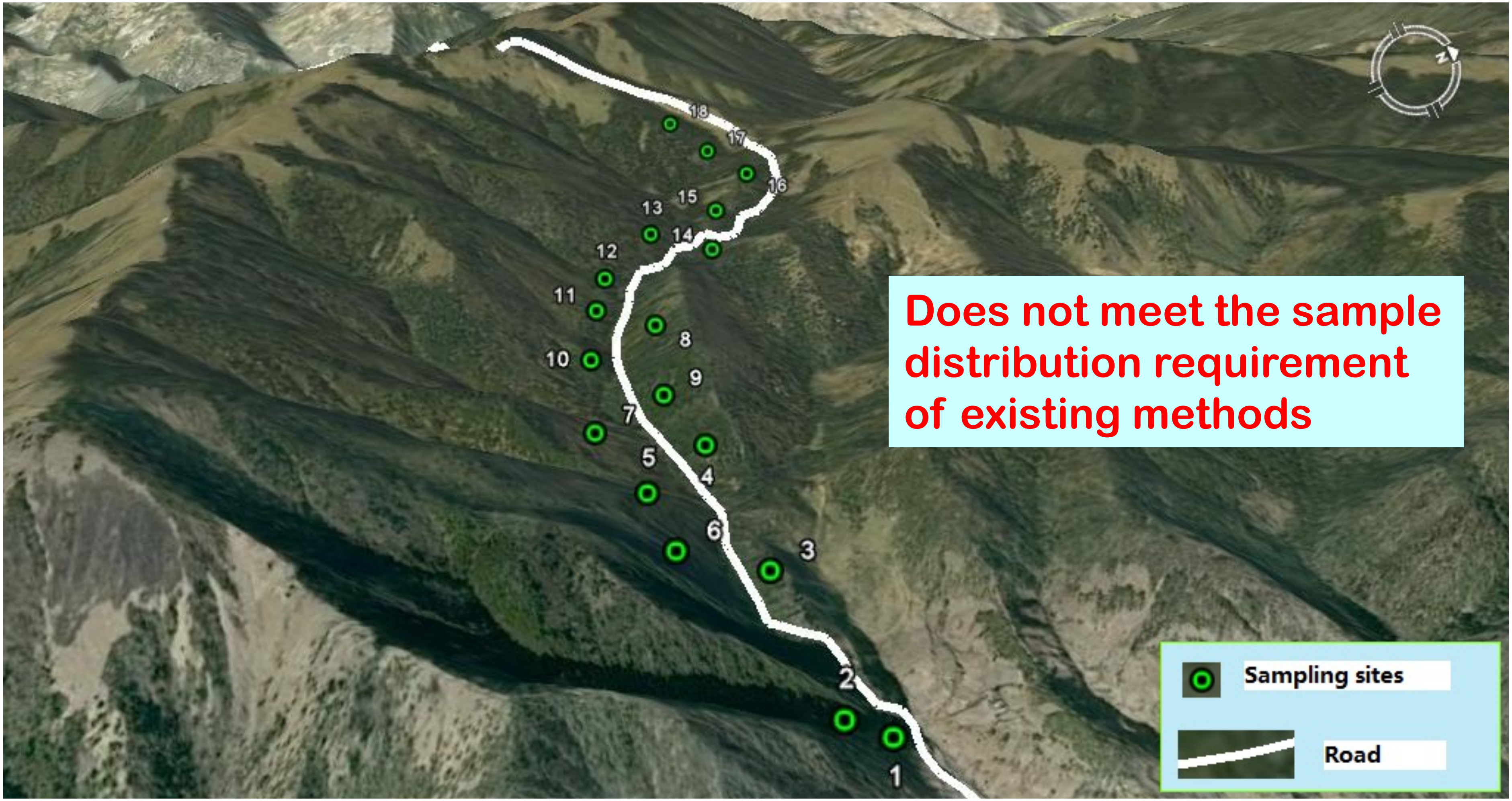
Google Earth

Data collection using location enabled PDAs

Case Study 3: New data – habitat mapping

Sampling bias (citizen data)

By nature volunteered geographic information from citizens are ad-hoc and opportunistic (spatially biased). They often do not meet the requirements of the traditional mapping techniques.



Case Study 3: New data – habitat mapping

Sampling bias (citizen data)

Habitat suitability of the Red-tailed Hawk in Wisconsin based on data from eBird.

Guiming Zhang
(U of Denver)

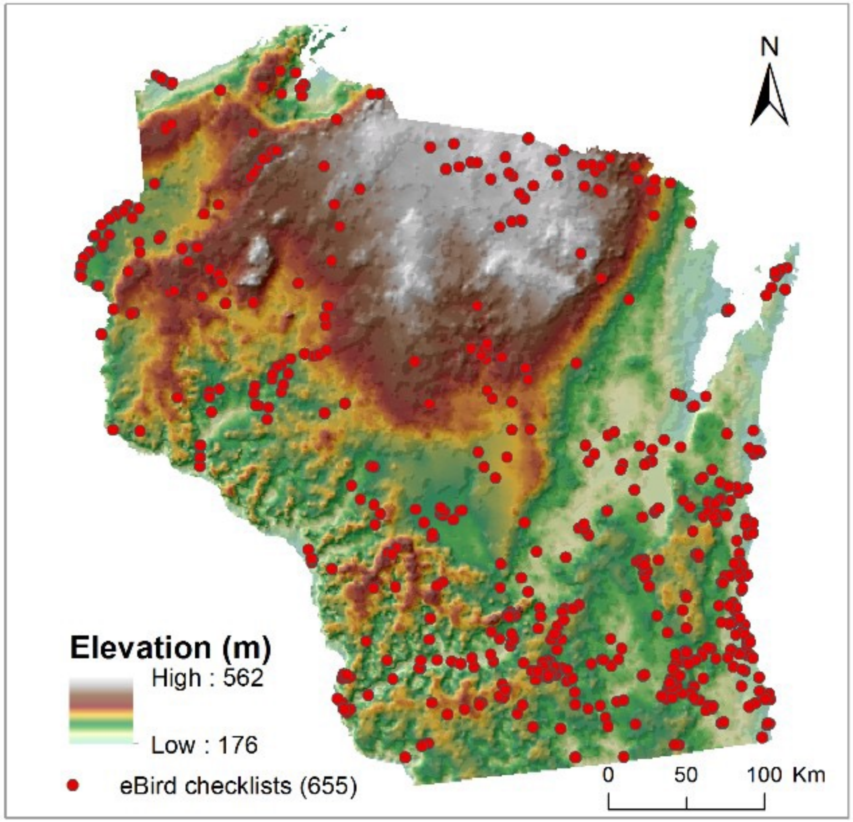
© Jonathan Eckerson | Macaulay Library

© Brian Sullivan | Macaulay Library

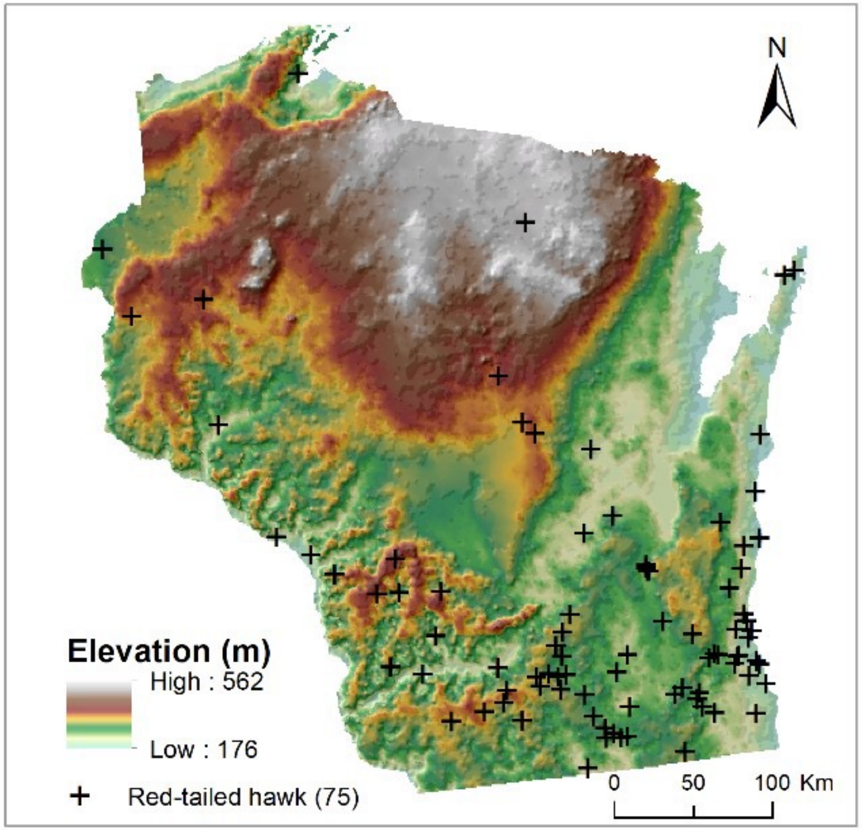
Case Study 3: New data – habitat mapping

Sampling bias (citizen data)

Citizen observation efforts

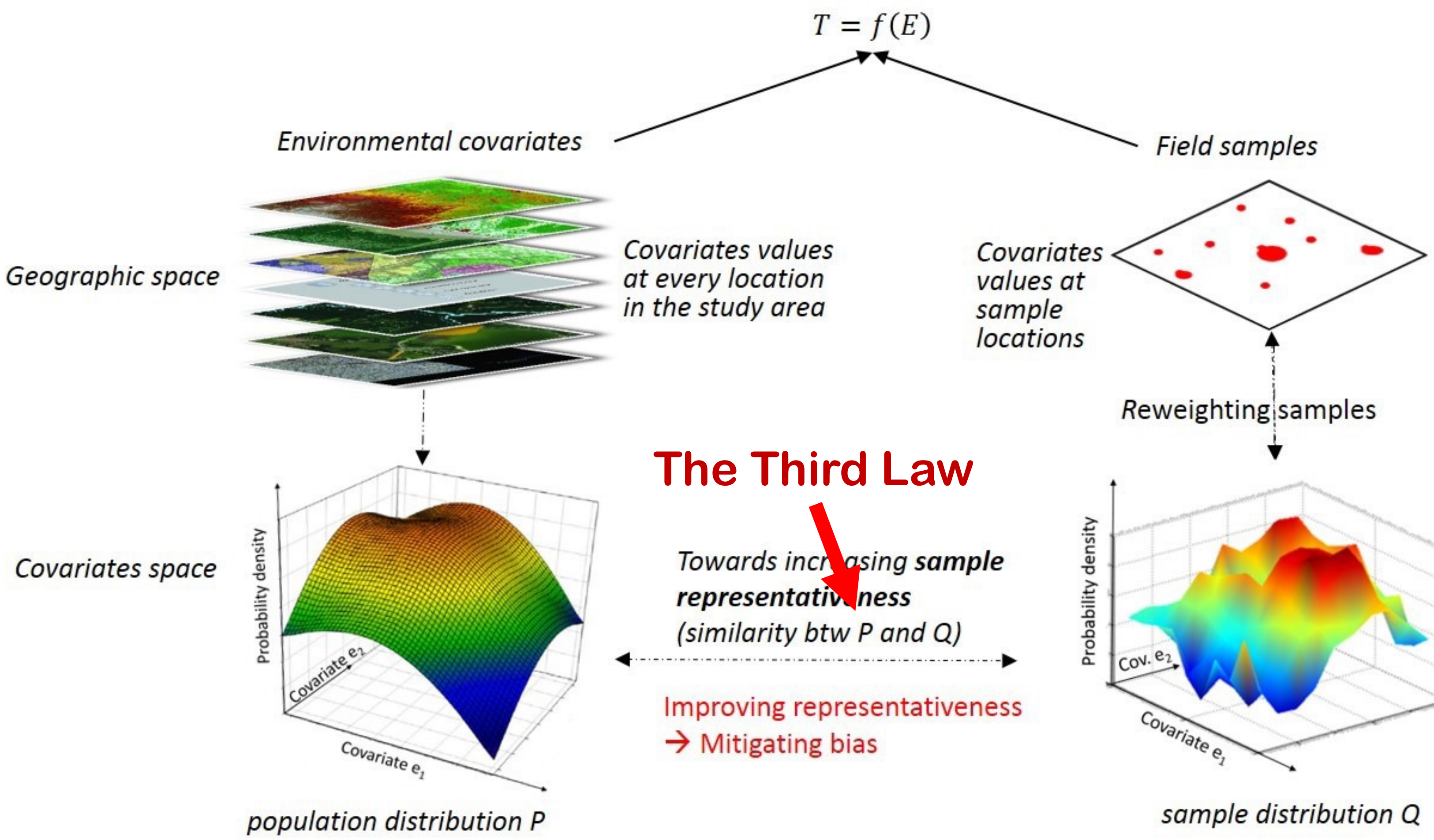


Red-tailed hawk observed



Case Study 3: New data – habitat mapping

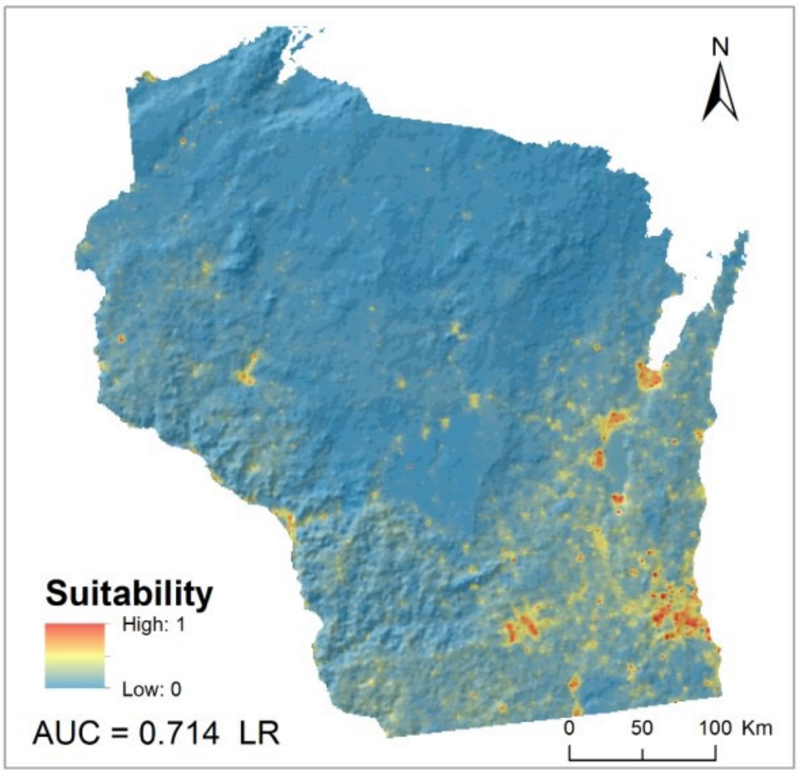
Sampling bias (citizen data)



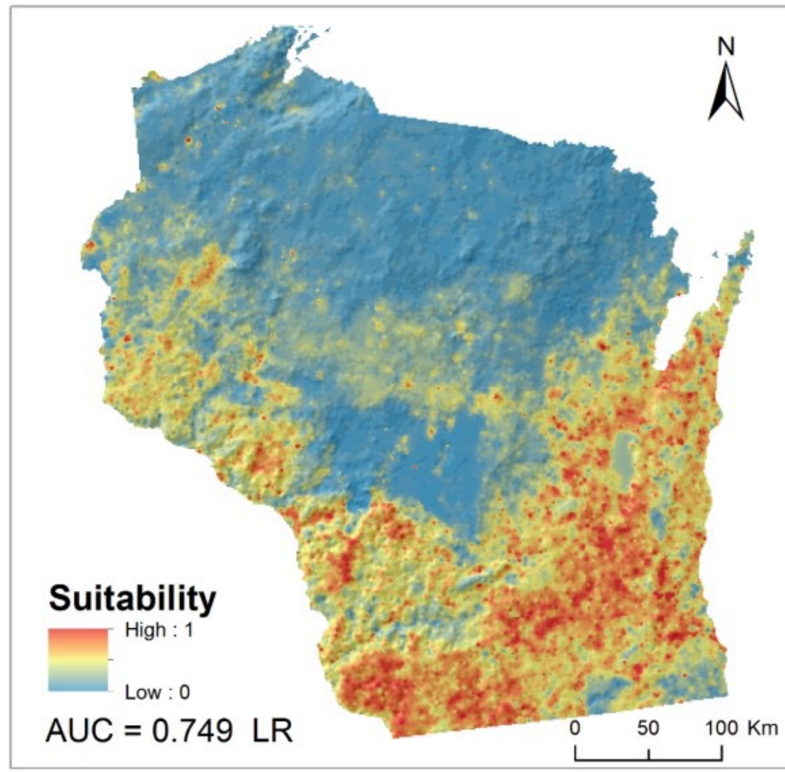
Case Study 3: New data – habitat mapping

Sampling bias (citizen data)

Without bias mitigation



With bias mitigation



Effective in mitigating spatial bias

Acknowledgements

Supports listed below are greatly appreciated:

1) Vilas Associate Award, Hamel Faculty Fellow Award, Manasse Endowed Chair Professorship from University of Wisconsin-Madison

2) Outstanding Oversea Scholarship Award from the Chinese Academy of Sciences

3) Outstanding Professorship Award from Nanjing Normal University

4) Natural Science of China: 40971236, 41023010, 41431177

5) Ministry of Science and Technology of China: 2007CB407207, 2007BAC15B01, 2011AA120305, 2015CB954102, 2010DFB24140

Laws of Geography

First Law of Geography:

"Everything is related to everything else, but near things are more related to each other" (Tobler, 1970)

Tobler W R. 1970. "A Computer Movie Simulating Urban Growth in the Detroit Region", *Economic Geography*, 46(Supp 1):234-240.

Second Law of Geography:

"Spatial heterogeneity could be called as the second law of geography" (Goodchild, 2004)

Goodchild, M.F., 2004. "The validity and usefulness of laws in geographic information science and geography." *Annals of the Association of American Geographers*, 94: 300-303

Third Law of Geography:

'The more similar geographic configurations of two points (areas), the more similar the values (processes) of the target variable at these two points (areas)' (Zhu et al., 2018)

Zhu, A-Xing, Guonian Lu, Jing Liu, Cheng-Zhi Qin, Chenghu Zhou, 2018. "Spatial prediction based on Third Law of Geography." Annals of GIS, 2018, VOL. 24, NO. 4, 225–240, DOI: 10.1080/19475683.2018.1534890

Other Papers

On spatial prediction based on the Third Law

Zhu, A.X., J. Liu, F. Du, S. J. Zhang, C.Z. Qin, J. Burt, T. Behrens, T. Scholten, 2015. "Predictive soil mapping with limited sample data", *European Journal of Soil Sciences*, Vol. 66, 535–547.

Zhang, Shu-Jie, A-Xing Zhu*, Jing Liu, Cheng-Zhi Qin, Yi-Ming An, 2016. "An heuristic uncertainty directed field sampling design for digital soil mapping", *Geoderma*, Vol. 267, pp. 123-136.

Li, Yan, A-Xing Zhu, Zhou Shi, Jing Liu, Fei Du, 2016. "Supplemental sampling for digital soil mapping based on prediction uncertainty from both the feature domain and the spatial domain", *Geoderma*, 284, 73-84.