ICUE2014, Pattaya, Thailand

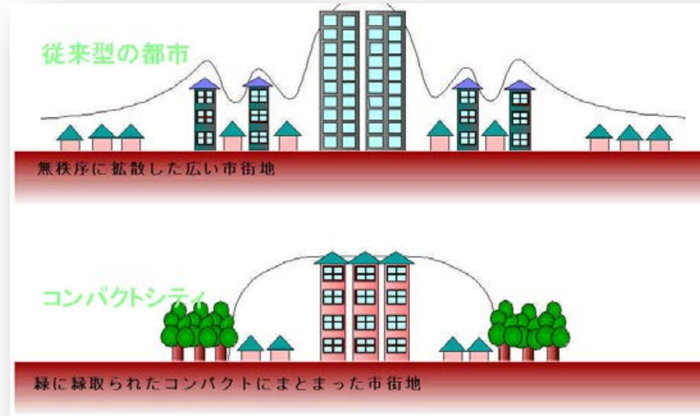
Creating municipality level floor space stock data in Japan by spatial statistics based areal interpolation method

Daisuke Murakami, Hajime Seya, Yoshiki Yamagata

Center for Global Environment Research,
National institute of Environmental Studies, Japan

Future smart city and energy systems

- Designing a future smart city (FSC) is one of the most urgent task.
  - Land use (Compact city)
  - Transport (Electric vehicle)



- FSC must be discussed considering not only land use/transport but also energy systems (Yamagata and Seya, 2013).

Electricity demand estimation

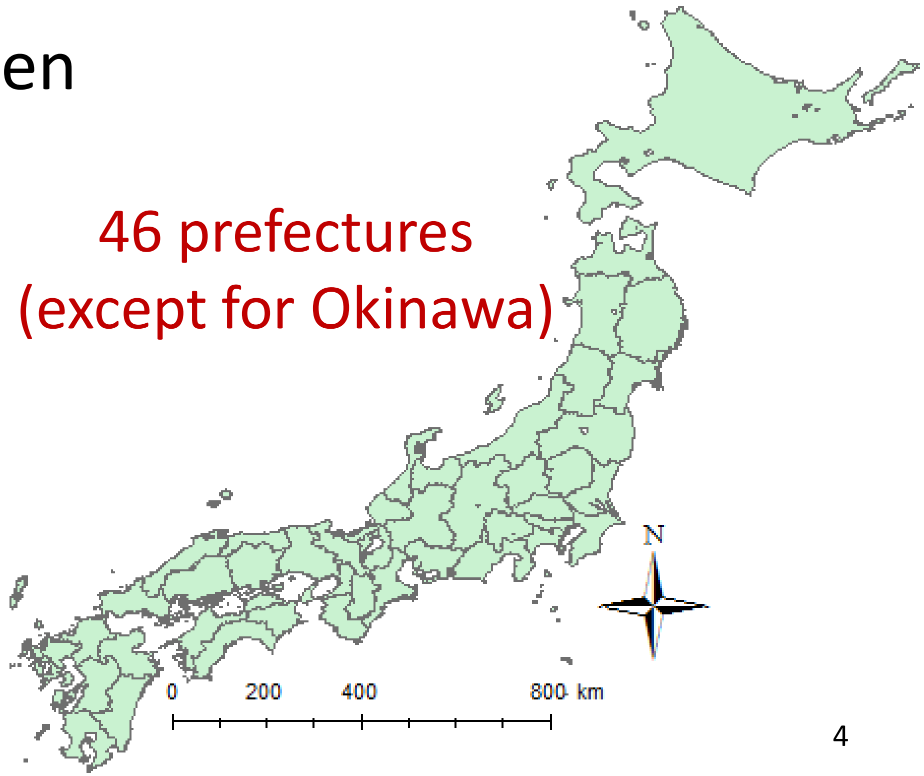
- Regional hourly electricity demand data by sector (e.g., residential) are not available in Japan.
  - To discuss the FSC from the viewpoint of energy systems, electricity demand must be estimated.

- The Japan Institute of Energy estimates the demand using the following equation:
  - *Electricity demand* = *f*(Total floor area in each sector)

Total floor area data in Japan

- Ministry of Land Infrastructure, Transport and Tourism (MLIT) summarizes the total floor area with the following categories:
  - ✓ Residence or non-residence
  - ✓ Wooden or non-wooden
  - ✓ Each completion year
  - ✓ Each prefecture

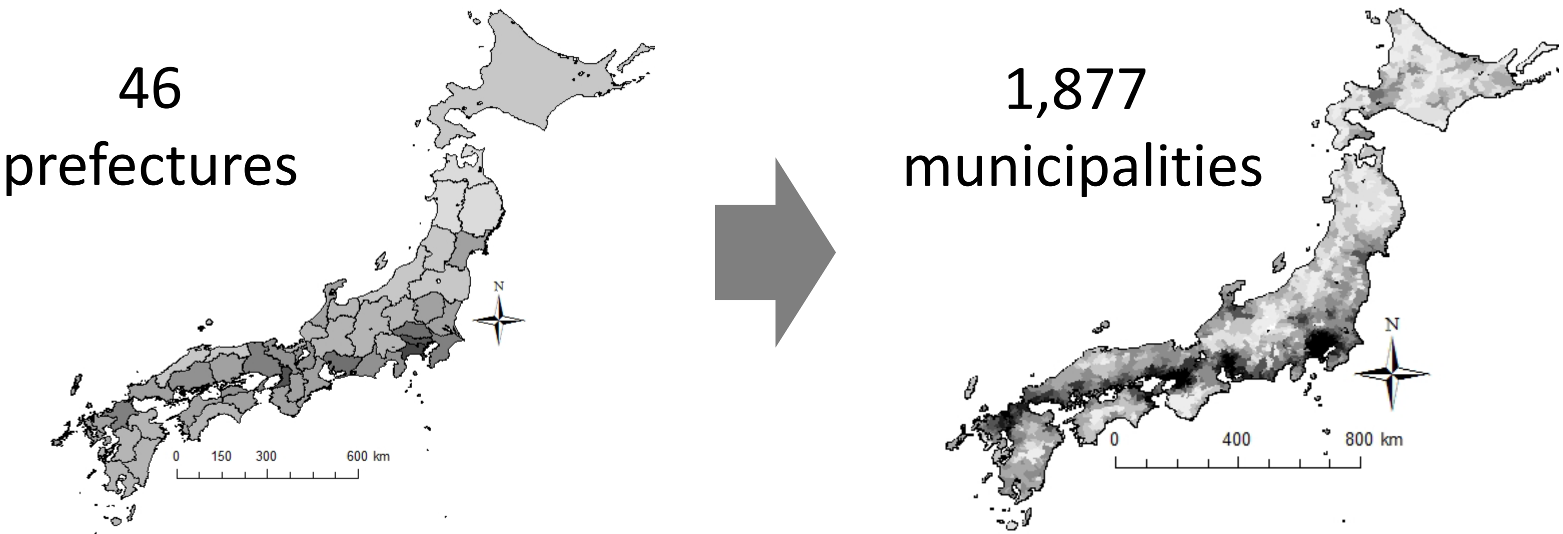
46 prefectures
(except for Okinawa)



The spatial resolution is not enough to discuss detailed electricity demand estimation

Objective of this study

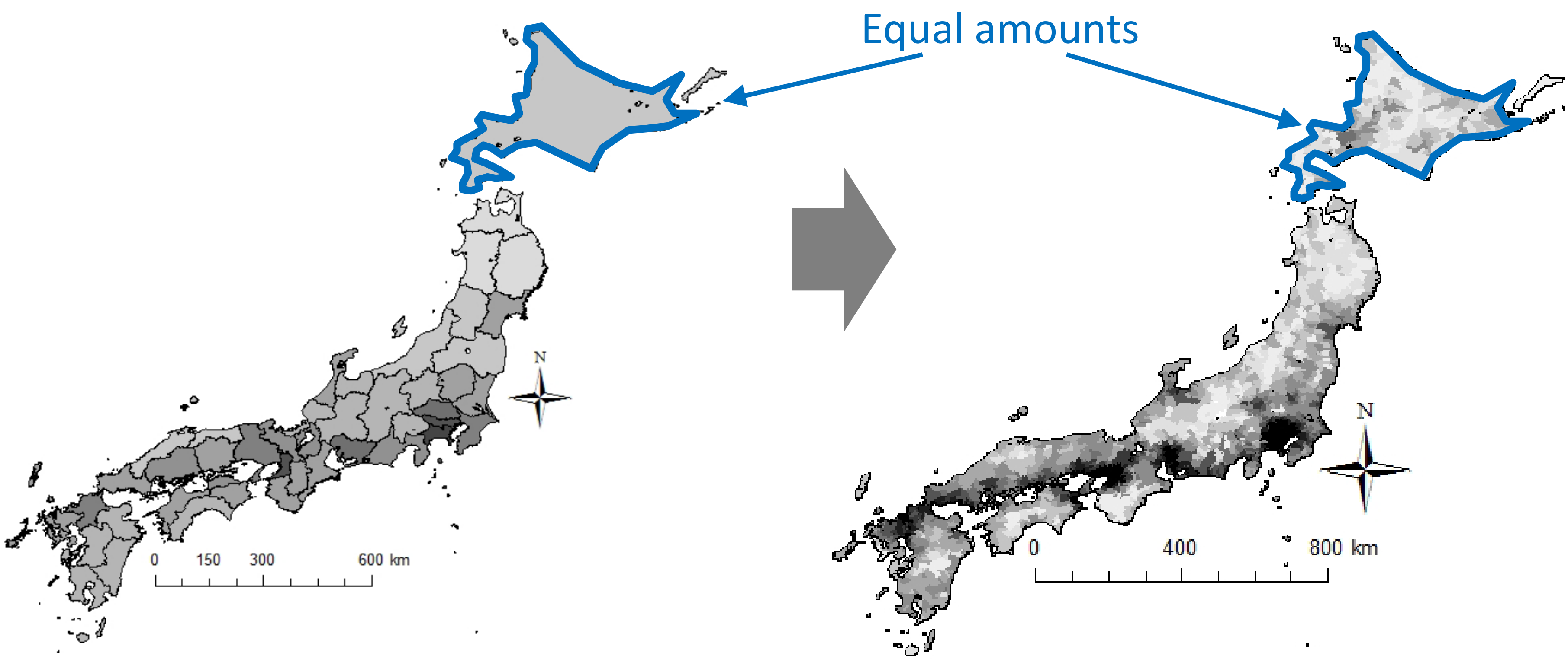
1. Municipal building stocks are estimated by downscaling the prefectural building stock data.
    - Spatial statistical methods are applied to the DS.



2. Municipal electricity demands are estimated utilizing the estimated municipal building stocks.

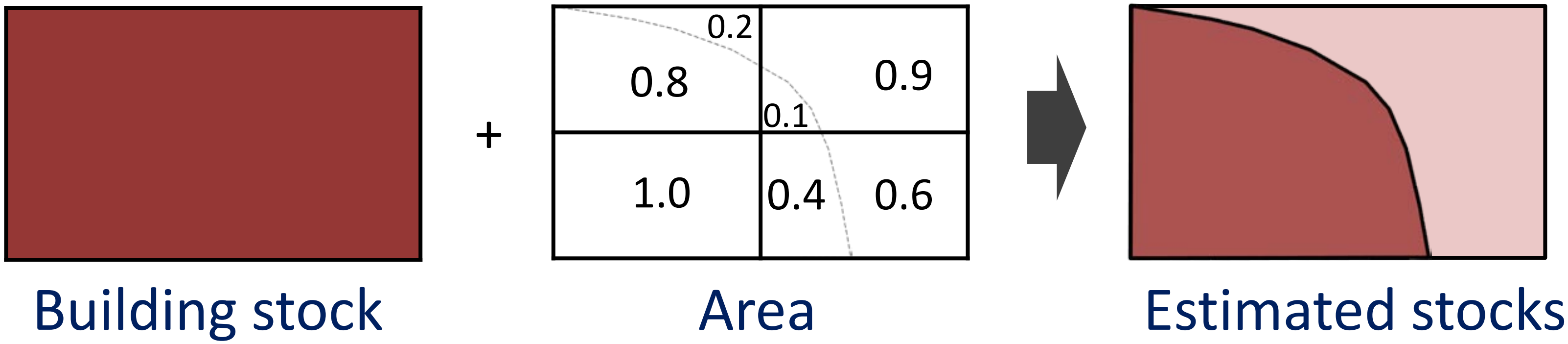
A principle in Downscaling

- Volume preserving property
  - Aggregations of the municipal building stocks estimates must equal to the prefectural actual stock amounts.



(Non- statistical) standard DS methods

- Areal weighting interpolation method
  - Proportional distribution according to areas.



- Dasymetric mapping method
  - Proportional distribution according to an ancillary data.

Properties of spatial statistics (in general)

- Advantages
  - Properties of spatial data, including spatial dependence and spatial heterogeneity, are easily introduced.
  - Mean squared error (MSE) is minimized in the interpolation.
- Disadvantages
  - They do not consider the volume preserving property.

- Problem
  - How the volume preserving property is satisfied while holding the advantages.

Consideration of spatial properties

- Geographically weighted regression(GWR) is used to capture spatial heterogeneity.

$$\mathbf{y} = \boldsymbol{\mu} + \boldsymbol{\varepsilon} \qquad \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$$

$$\boldsymbol{\mu} = \left\langle \sum_p x_{i,p} \beta_{i,p} \right\rangle_i$$

Similarity among $\beta_{i,p}$s in each municipality are modeled using a kernel function

$\mathbf{y}$: Building stocks	$x_{i,p}$: p-th covariates in unit i
$\langle \bullet \rangle_i$: A vector whose i-th element is $\bullet$	$\beta_{i,p}$: p-th parameter in unit i

  – Spatial heterogeneity is modeled <u>by allowing $\beta_{i,p}$ to vary across geographical space.</u>

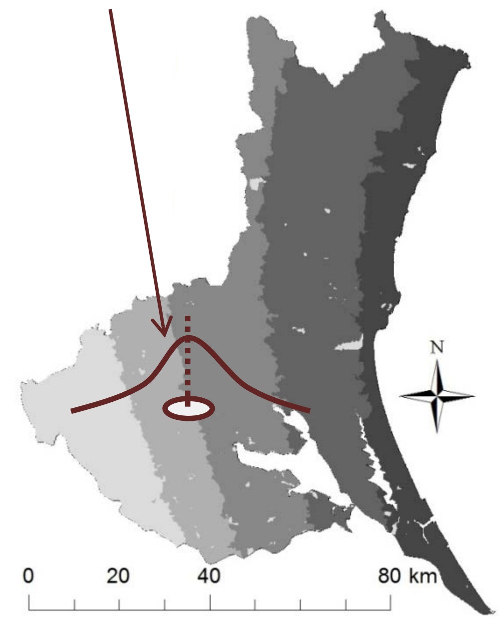
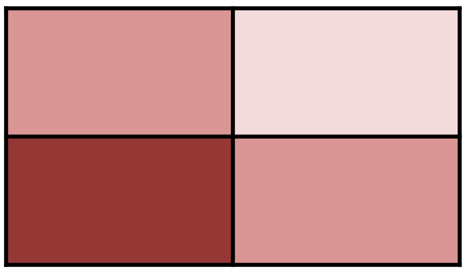


Image of $\beta_{i,p}$s

GWR-based DS model

- GWR is extended as with a model construction in geostatistics.

Prefectural stocks : $\bar{\mathbf{y}}$



Prefectural level GWR

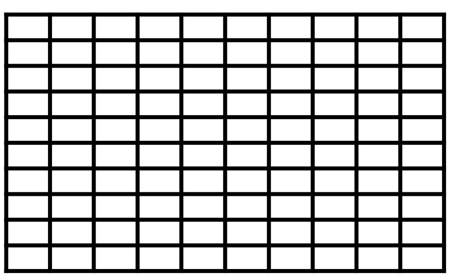
$$\bar{\mathbf{y}} = \mathbf{N}\boldsymbol{\mu} + \mathbf{N}\boldsymbol{\varepsilon} \qquad \mathbf{N}\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{N}\mathbf{N}')$$

N:A matrix for aggregation

DS

Aggregation

Municipal stocks: $\mathbf{y}$



Municipal level GWR

$$\mathbf{y} = \boldsymbol{\mu} + \boldsymbol{\varepsilon} \qquad \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{I})$$

Unknown

- The predictor of $\mathbf{y}$ with minimum MSE.

$$\hat{\mathbf{y}} = \boldsymbol{\mu} + \mathbf{N}'(\mathbf{N}\mathbf{N}')^{-1}(\bar{\mathbf{y}} - \mathbf{N}\boldsymbol{\mu})$$

Properties of the GWR-based DS method

- Consideration of a spatial data property: ✓
- Prediction with minimum MSE: ✓
- Volume preservation: ?
  - It is satisfied if aggregations of municipal building stock estimates ($\hat{\mathbf{y}}$) equal to the prefectural actual building stocks ($\overline{\mathbf{y}}$).
  - Let aggregate $\hat{\mathbf{y}}$ using the aggregation matrix **N** as

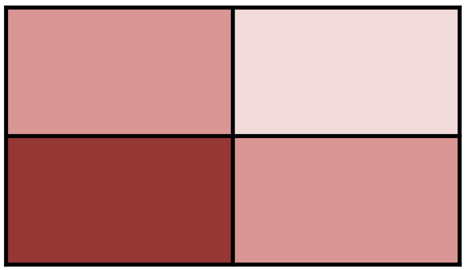
$$\begin{aligned}\mathbf{N}\hat{\mathbf{y}} &= \mathbf{N}[\boldsymbol{\mu} + \mathbf{N}'(\mathbf{N}\mathbf{N}')^{-1}(\overline{\mathbf{y}} - \mathbf{N}\boldsymbol{\mu})] \\ &= \mathbf{N}\boldsymbol{\mu} + \overline{\mathbf{y}} - \mathbf{N}\boldsymbol{\mu} \\ &= \overline{\mathbf{y}}\end{aligned}$$

⟶ Thus, the extended GWR preserves volumes.

Geostatistics (GS)-based DS model

- The GWR-based method is an extension of the GS-based method

Prefectural stocks : $\bar{\mathbf{y}}$



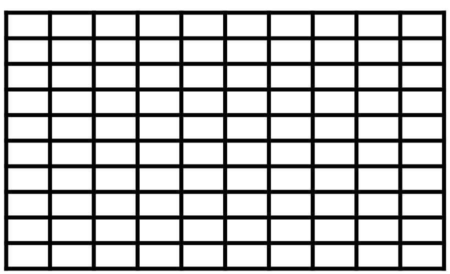
Prefectural level GWR

$$\bar{\mathbf{y}} = \mathbf{NX\beta} + \mathbf{N\varepsilon} \qquad \mathbf{N\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{NCN}')$$

DS

Aggregation

Municipal stocks: $\mathbf{y}$



Municipal level GWR

$$\mathbf{y} = \mathbf{X\beta} + \boldsymbol{\varepsilon} \qquad \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{C})$$

X: Explanatory variables **β**: Parameters

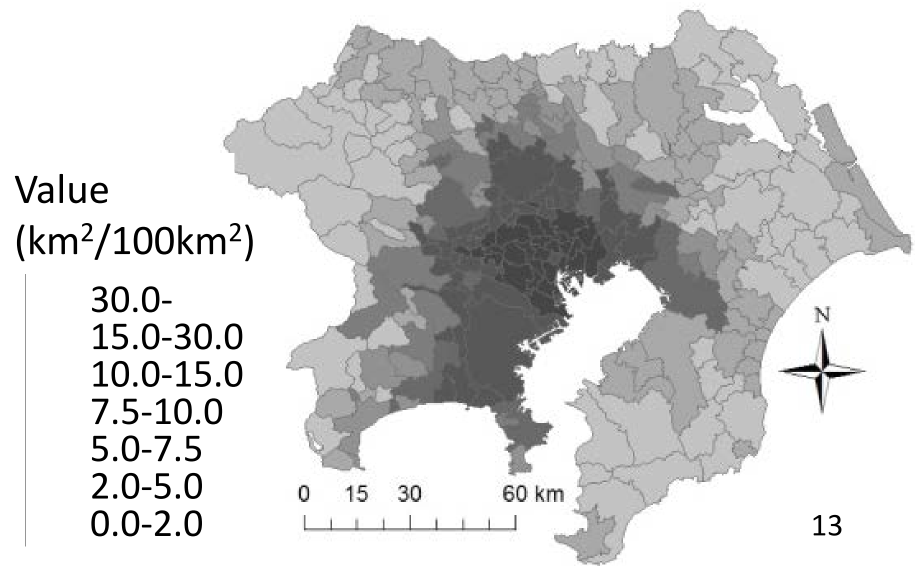
C: A distance matrix describing spatial dependence (i.e., nearby observations are similar)

- The predictor of $\mathbf{y}$ with minimum MSE.

$$\hat{\mathbf{y}} = \mathbf{X\beta} + \mathbf{CN}'(\mathbf{NCN}')^{-1}(\bar{\mathbf{y}} - \mathbf{NX\beta})$$

Municipal building stock estimation

- Municipal building stocks are estimated by downscaling the prefectural stocks data provided by MLIT.
  - To examine accuracy of our downscaling, residential stock data in Tokyo metropolitan area was constructed based on the Fixed asset tax rolls (below figure).



Municipal residential building stock data in Tokyo

Accuracy comparison: measure

$$\text{RMSE} = \sqrt{\frac{1}{1{,}803}\sum_i (\hat{y}_i^{volume} - y_i^{volume})^2}$$

$$\text{RMSE(dens.)} = \sqrt{\frac{1}{1{,}803}\sum_i (\hat{y}_i - y_i)^2}$$

$$\text{RMSPE} = \sqrt{\frac{1}{1{,}803}\sum_i \left(\frac{\hat{y}_i^{volume} - y_i^{volume}}{y_i^{volume}}\right)^2}$$

$$\text{MAE} = \frac{1}{1{,}803}\sum_i |\hat{y}_i^{volume} - y_i^{volume}|$$

$$\text{MAE(dens.)} = \frac{1}{1{,}803}\sum_i |\hat{y}_i - y_i|$$

$$\text{MAPE} = \frac{1}{1{,}803}\sum_i \left|\frac{\hat{y}_i^{volume} - y_i^{volume}}{y_i^{volume}}\right|$$

RMSE=Root mean square error

RMSPE=Root mean square percentage error

MAE = Mean absolute error

MAPE = Mean absolute percentage error

y_i^{volume} : Stock amounts

y_i : Stock density

Accuracy comparison result

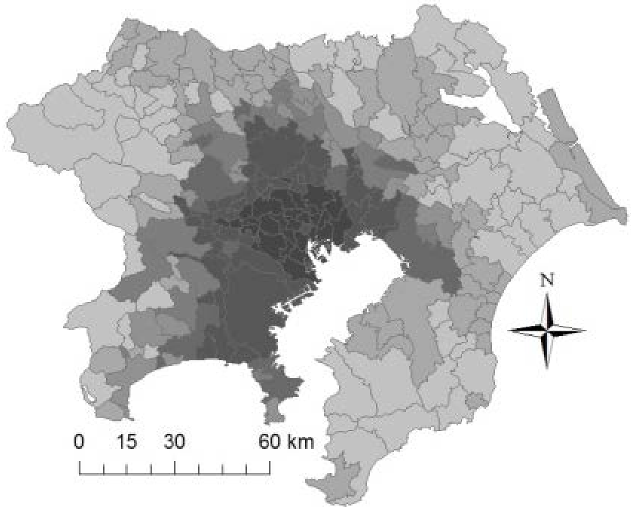
Red :Best   (shaded) :Better than the dasymetric method

	Areal weight	Dasy metric	Without covariates		With covariates	
			GS	GWR	GS	GWR
Covariates	N.A.		None		No. of railway stations Road densities	
Weights	Area	Building land area				
RMSE	7.97×10^{6}	2.62×10^{6}	**1.93×10^{6}**	4.52×10^{6}	2.19×10^{6}	2.76×10^{6}
RMSE(dens.)	1.44×10^{5}	8.50×10^{4}	**7.58×10^{4}**	1.08×10^{5}	1.24×10^{5}	1.22×10^{5}
RMSPE	1.47	8.35×10^{-1}	**5.39×10^{-1}**	6.57×10^{-1}	6.27×10^{-1}	6.22×10^{-1}
MAE	4.26×10^{6}	1.70×10^{6}	1.26×10^{6}	1.93×10^{6}	**1.18×10^{6}**	1.27×10^{6}
MAE(dens.)	9.07×10^{4}	5.12×10^{4}	**4.34×10^{4}**	5.47×10^{4}	4.49×10^{4}	4.51×10^{4}
MAPE	3.29	5.74×10^{-1}	3.75×10^{-1}	4.59×10^{-1}	**3.56×10^{-1}**	3.60×10^{-1}

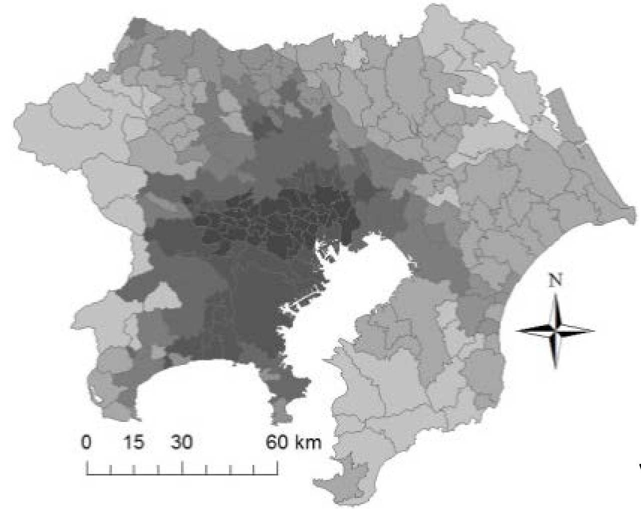
– GS and GWR outperforms the dasymetric method whose accuracy has been demonstrated.

– Consideration of covariates does not necessarily improves prediction accuracy.

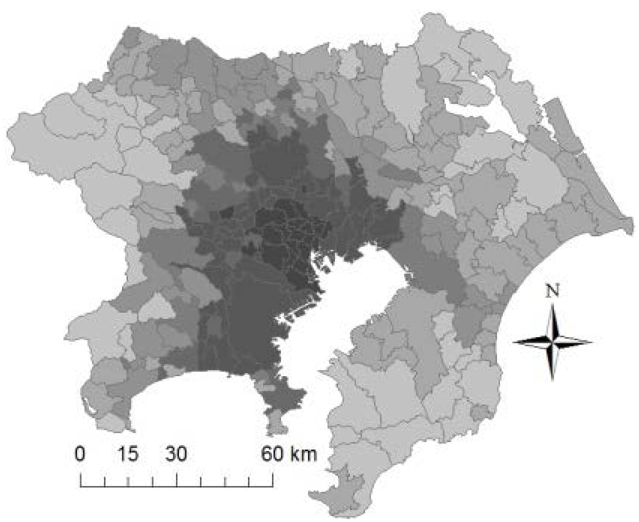
Downscaling results



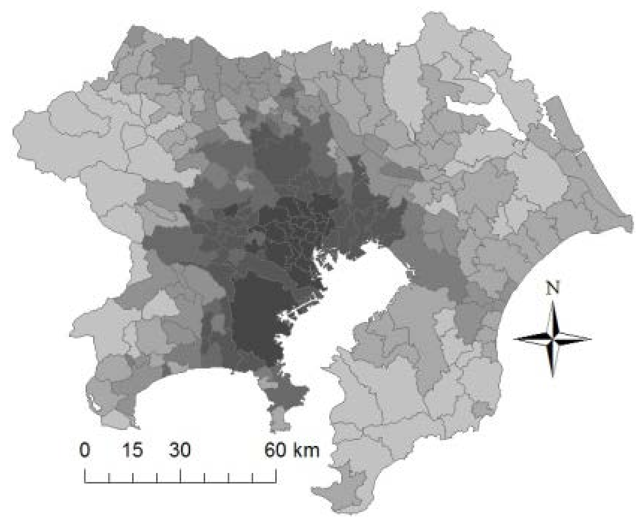
True



Dasymetric method



GS without covariates



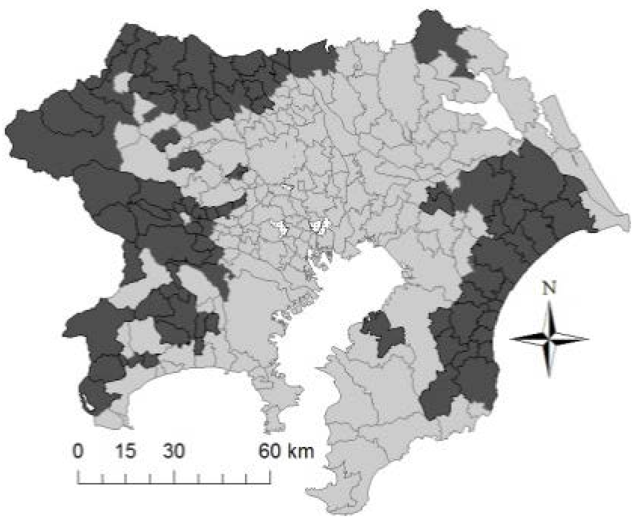
GWR with covariates

Value
($km^2/100km^2$)

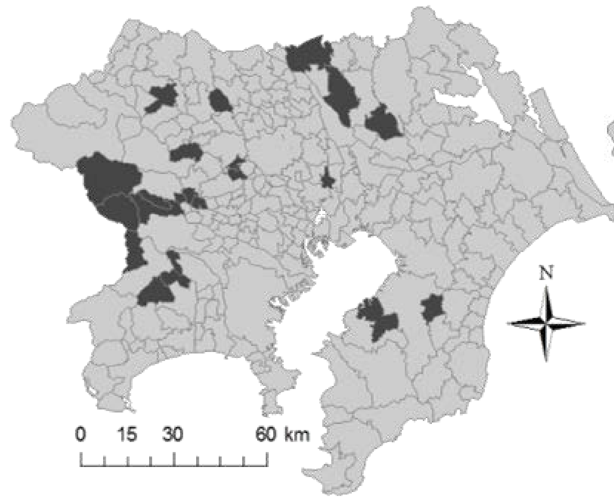
30.0-
15.0-30.0
10.0-15.0
7.5-10.0
5.0-7.5
2.0-5.0
0.0-2.0
0.0

Spatial plots of the error ratios (ER)s

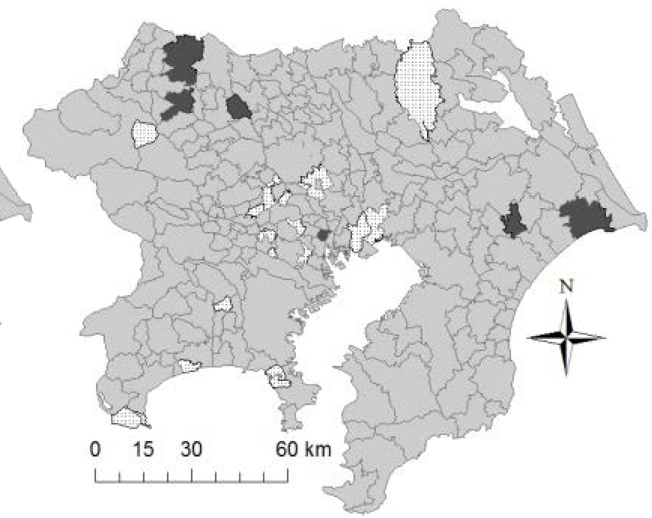
■ : ER＞0.5 　　□ : ER＜－0.5



Dasymetric method



GS without covariates



GWR with covariates

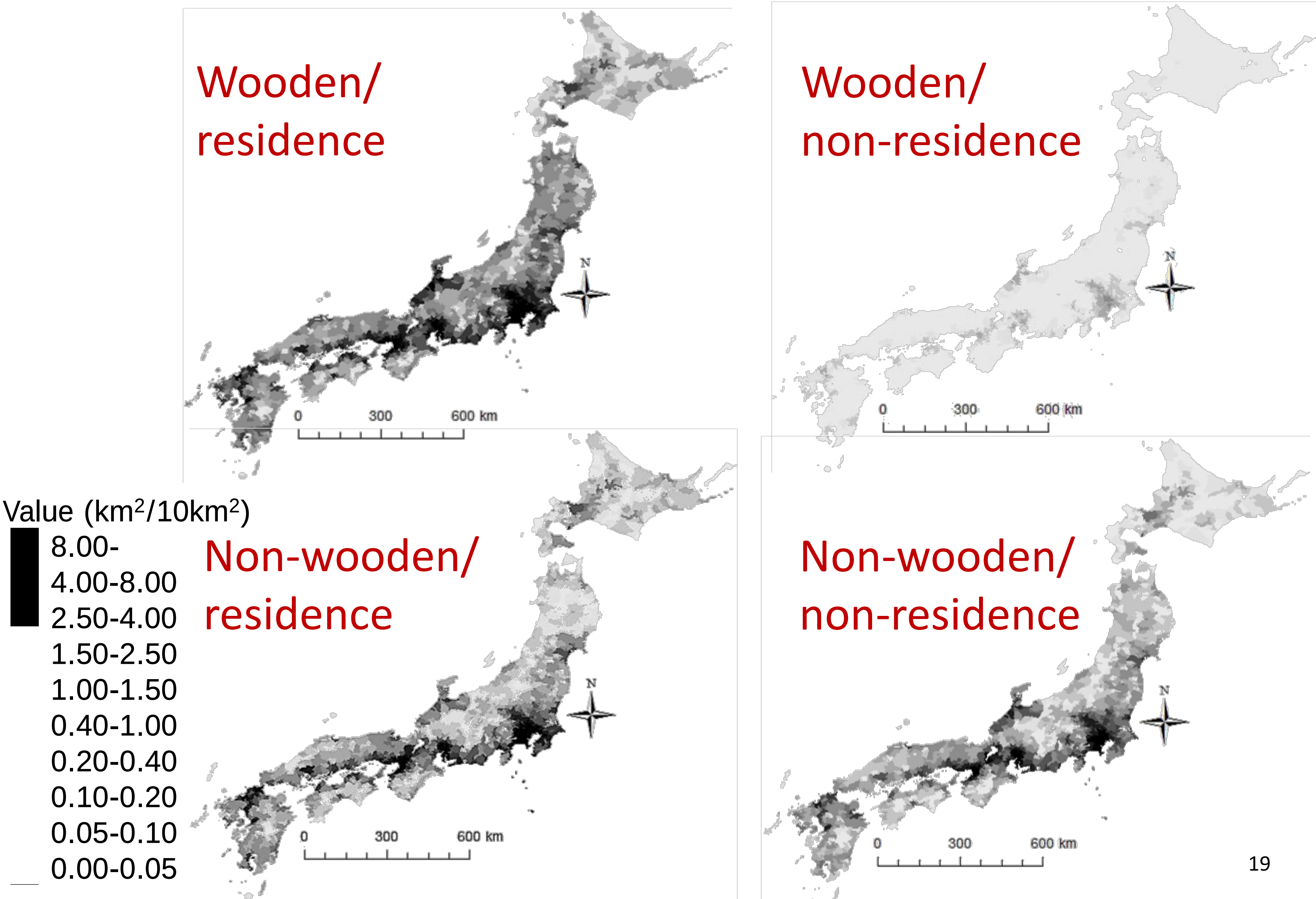
- Dasymetric : Over-estimations are found in non-urban areas.
- GS and GWR: Such over-estimations are not found and accurate.

Discussion

- Effectiveness of the GWR-based DS method was conformed.
  - Accuracy
  - Efficiency in capturing spatial pattern of data

- GS without covariates, which was the most accurate, is adopted to the municipal building stock estimation all over Japan.
- Then, the electricity demands are estimated using the estimated stock data.

Estimation result of the municipal stocks in 2007



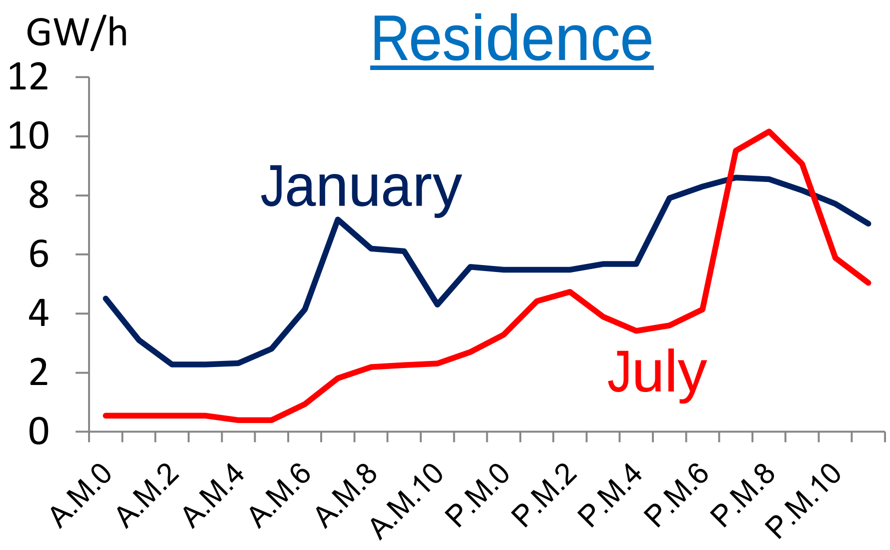
Electricity demand estimation

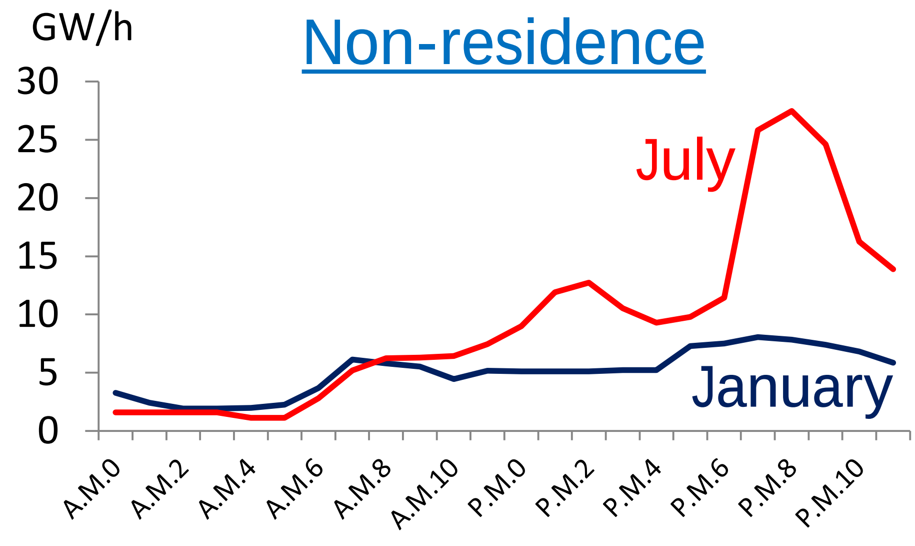
- Estimation equation of Japan Institute of Energy:

Hourly electricity demand =

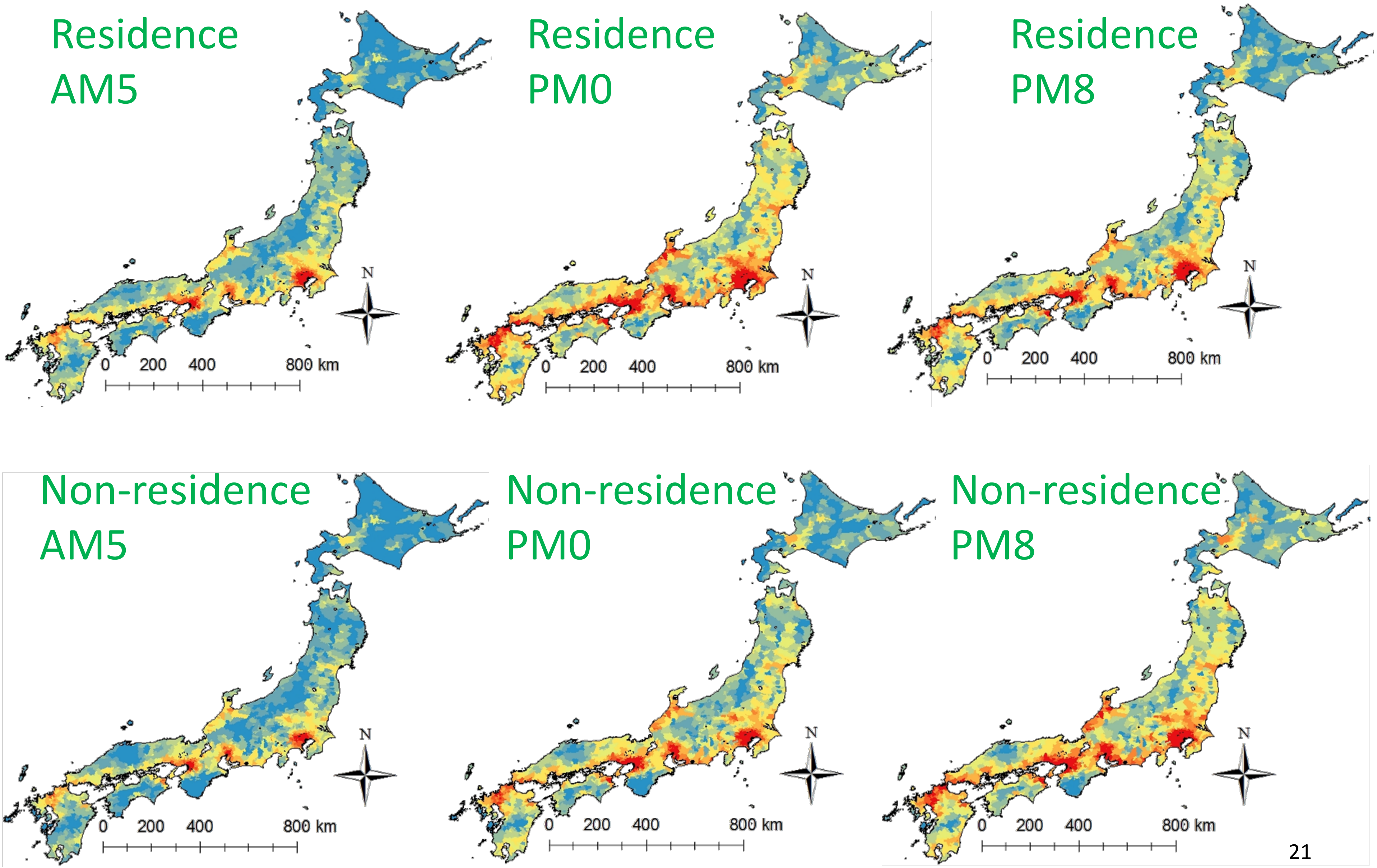
Building stock amount × *unit hourly electricity demand*

- Total electricity demands in Japan

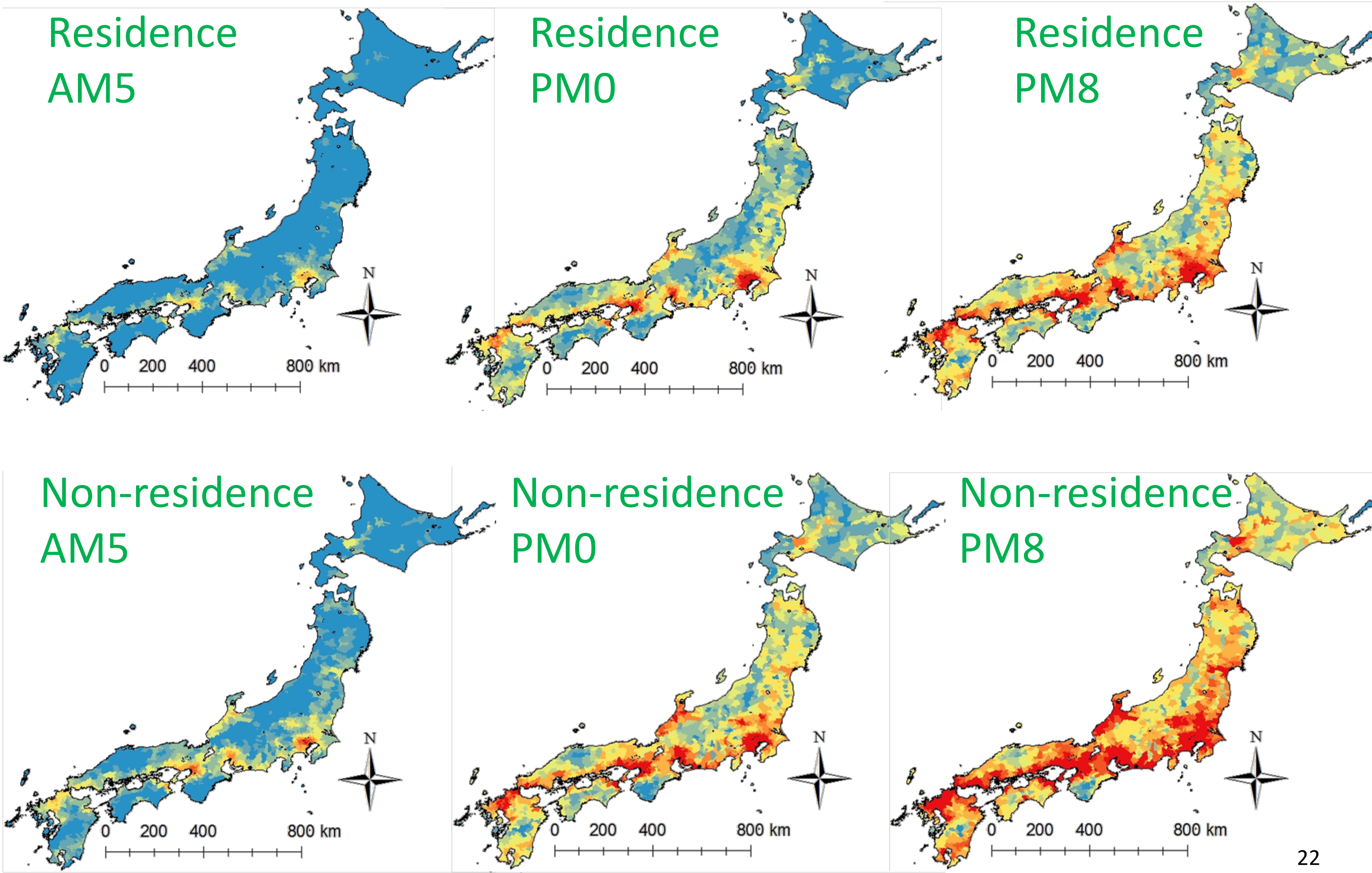




Estimated residential electricity demands (January)



Estimated non-residential electricity demands (July)



Concluding remarks

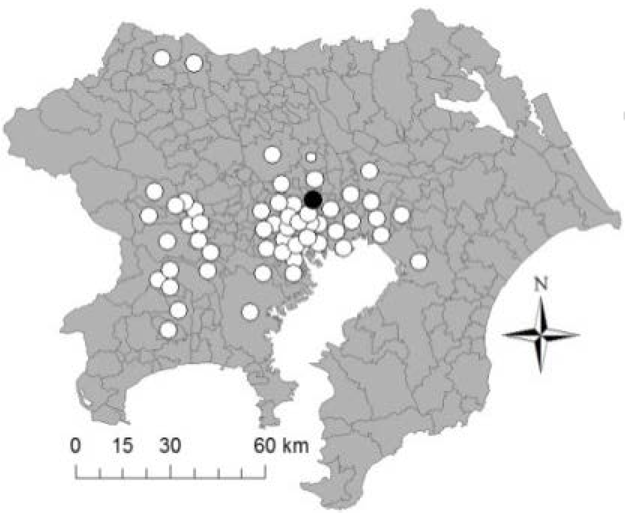
Effectiveness of spatial statistical (i.e., GWR-based and GS-based) methods is demonstrated.

Municipal electricity demands are estimated using a spatial statistical methods.

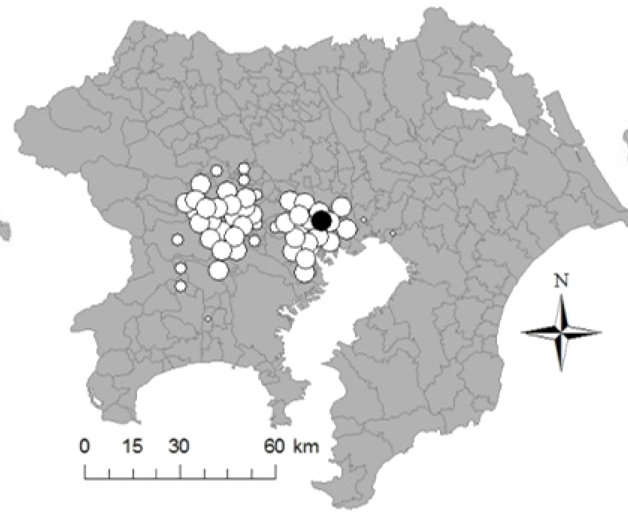
- Task in future studies
  - Employment of the estimated data to discussions of urban planning, including the achievement of FSC.
  - Development of methods for DS that successfully utilize covariates.
  - Computationally efficient electricity demand estimation

Unexplained spatial pattern

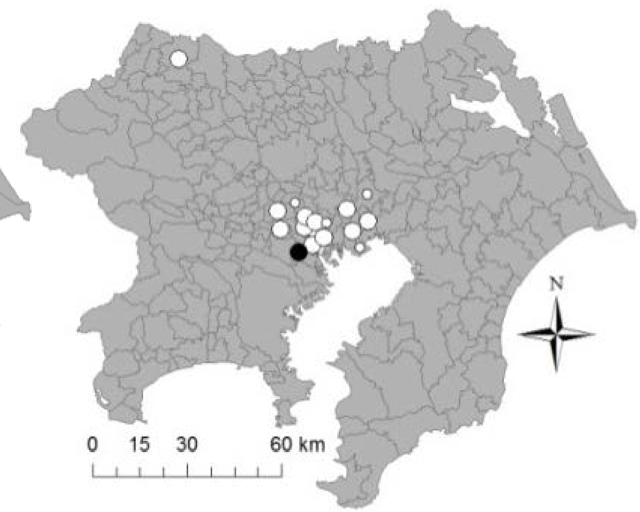
- Residual spatial dependences are tested by the local Moran statistics.



Dasymetric method



GS without covariates



GWR with covariates

- Dasymetric : Residual Spatial dependence is significant in many municipalities.
- GS and GWR : Residual spatial dependence is significant in the central Tokyo area only.