Expanding Methodology Plan by Using Data Fusion Method

National Institute for Environmental Study, Japan Seiya Maki

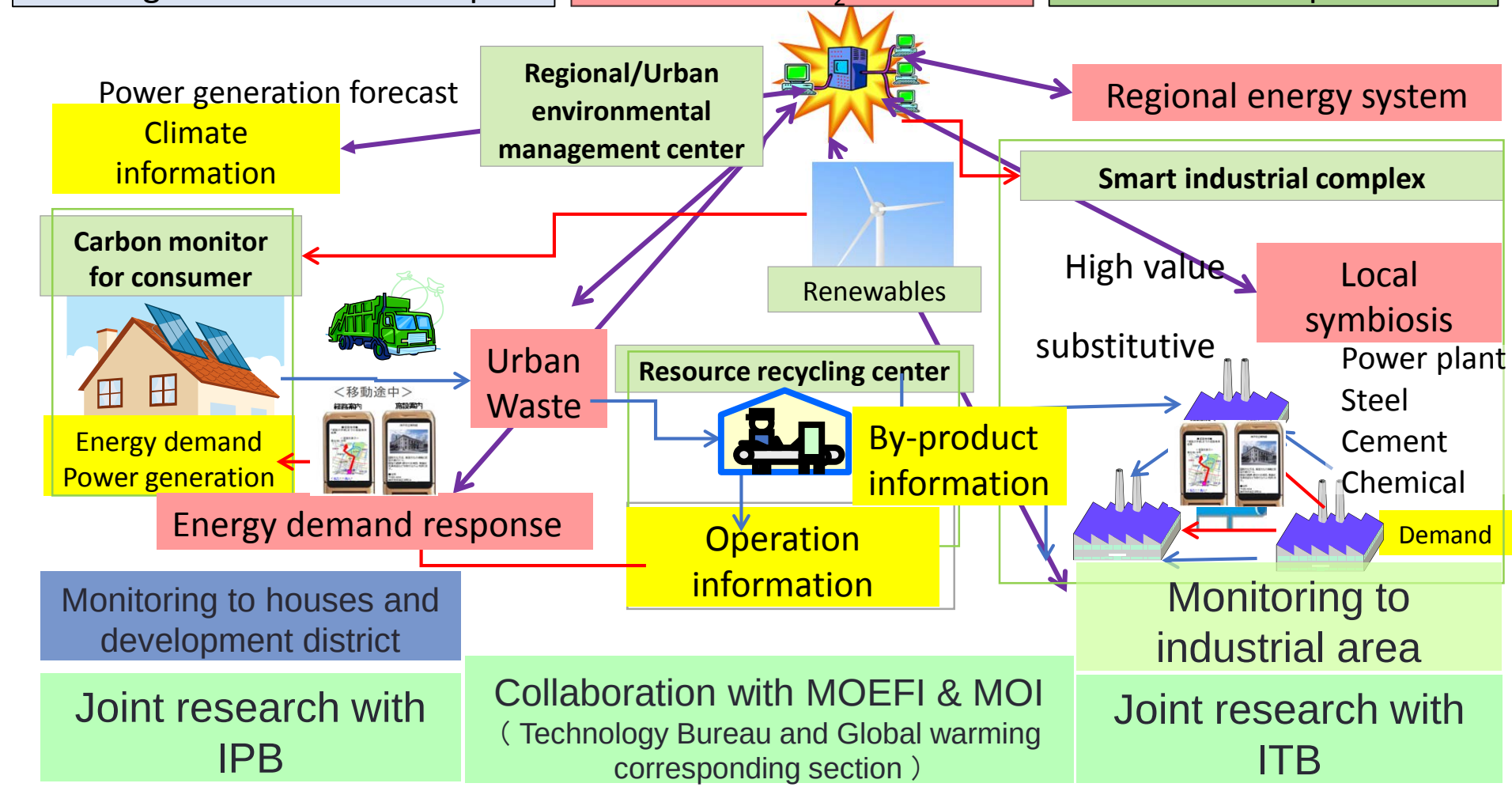
maki.seiya@nies.go.jp

Overview of Our Project: Framework of Indonesia

Low carbon monitoring system for neighborhoods and campus

Integrated information network for CO_2 emissions

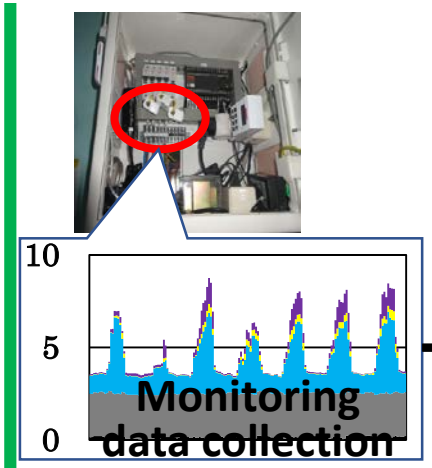
CO_2 monitoring system for industrial parks



- Innovation City or Regional scale Social system for Low-carbon society
- Evaluation CO_2 Reduction policy & Verification of potential by Monitoring Systems

Analysis flow of this study

Monitoring Data collection & analysis



Statistical analysis

Data fusion

Engineering analysis

Unit data & micro simulation

Unit data development

Constraint collection

DR sim

Spatio - Temporal extended

Questionnaire data

Geo information data

Data fusion

Monitoring data fitting for spatial

Data fusion

Macro Simulation

Spatio-Temporal model

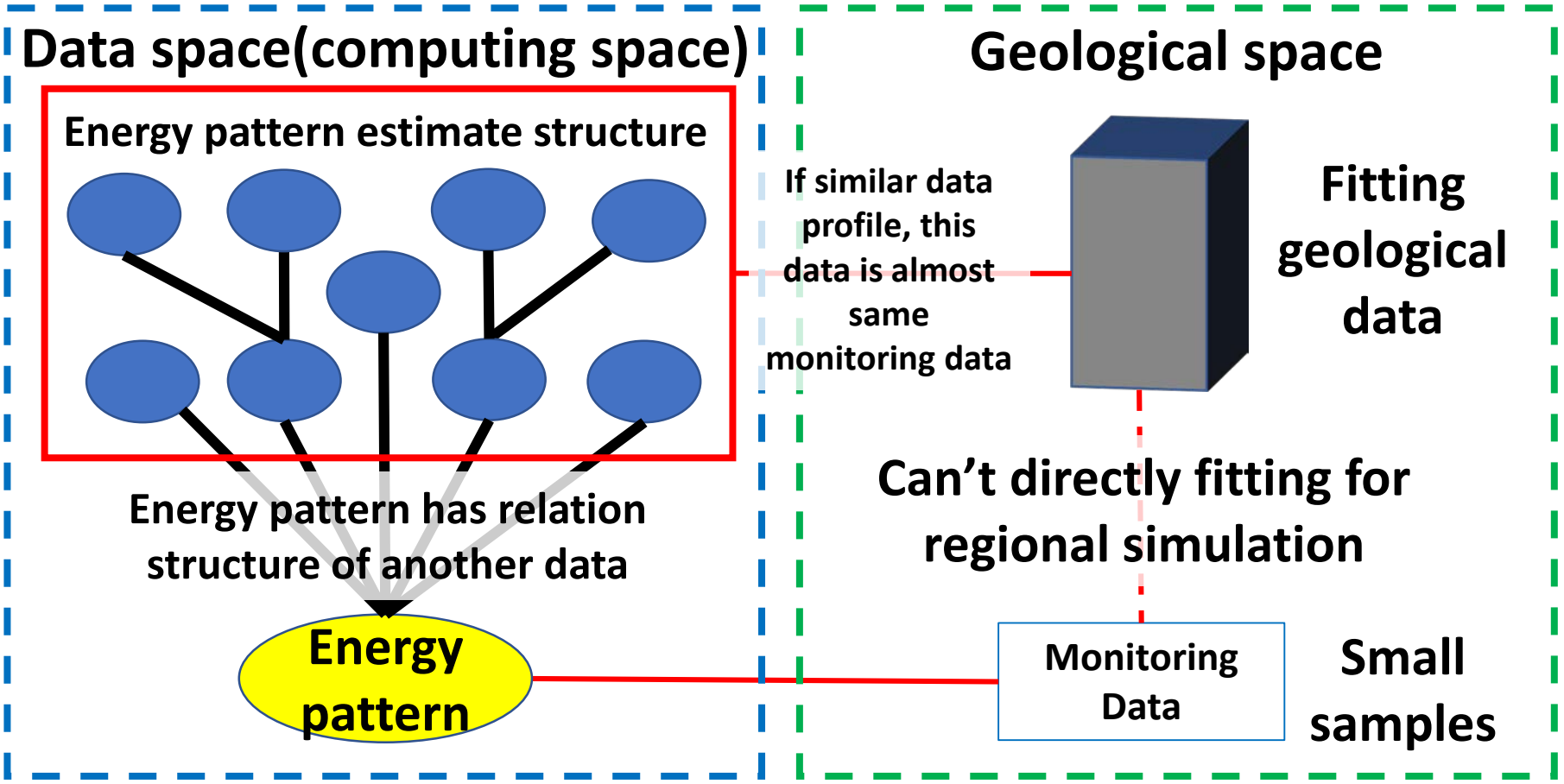
ADR sim

Scenario simulation

What is Data Fusion?

Data fusion is the process of integrating multiple data sources to produce more consistent, accurate, and useful

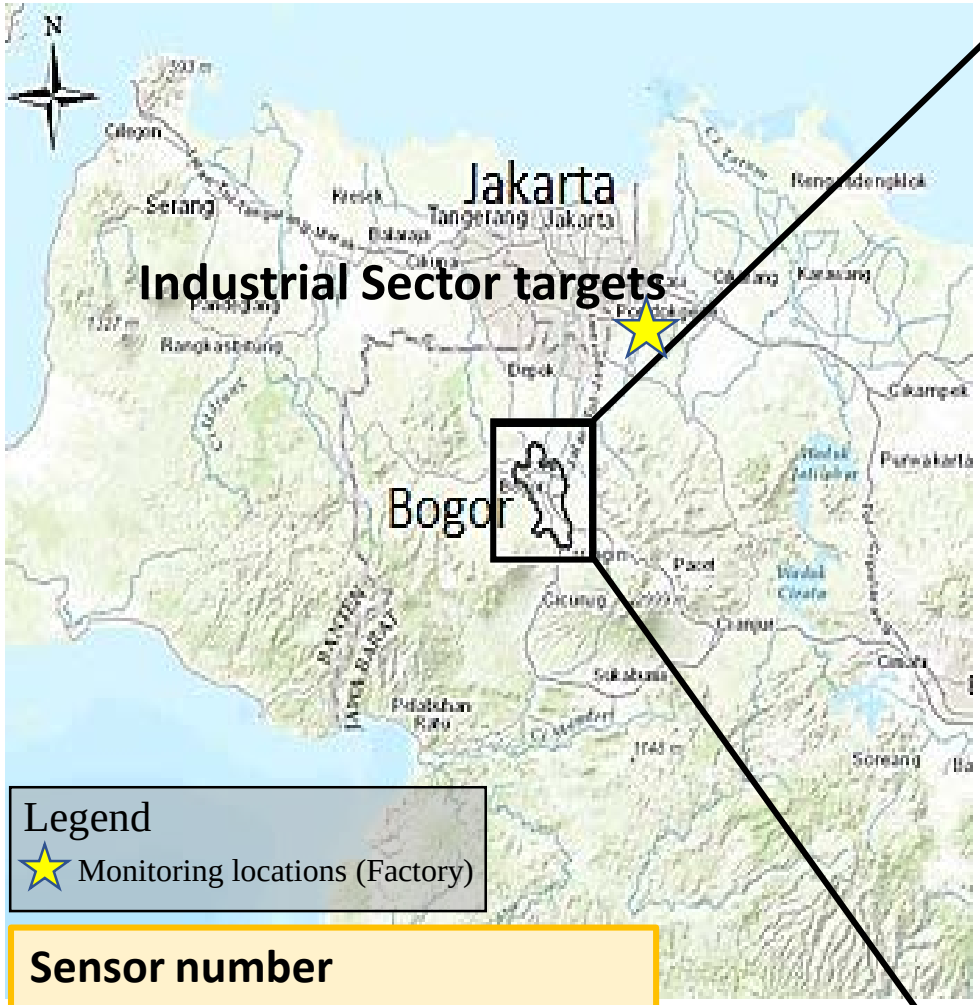
Ex. Extended monitoring data for regional scale



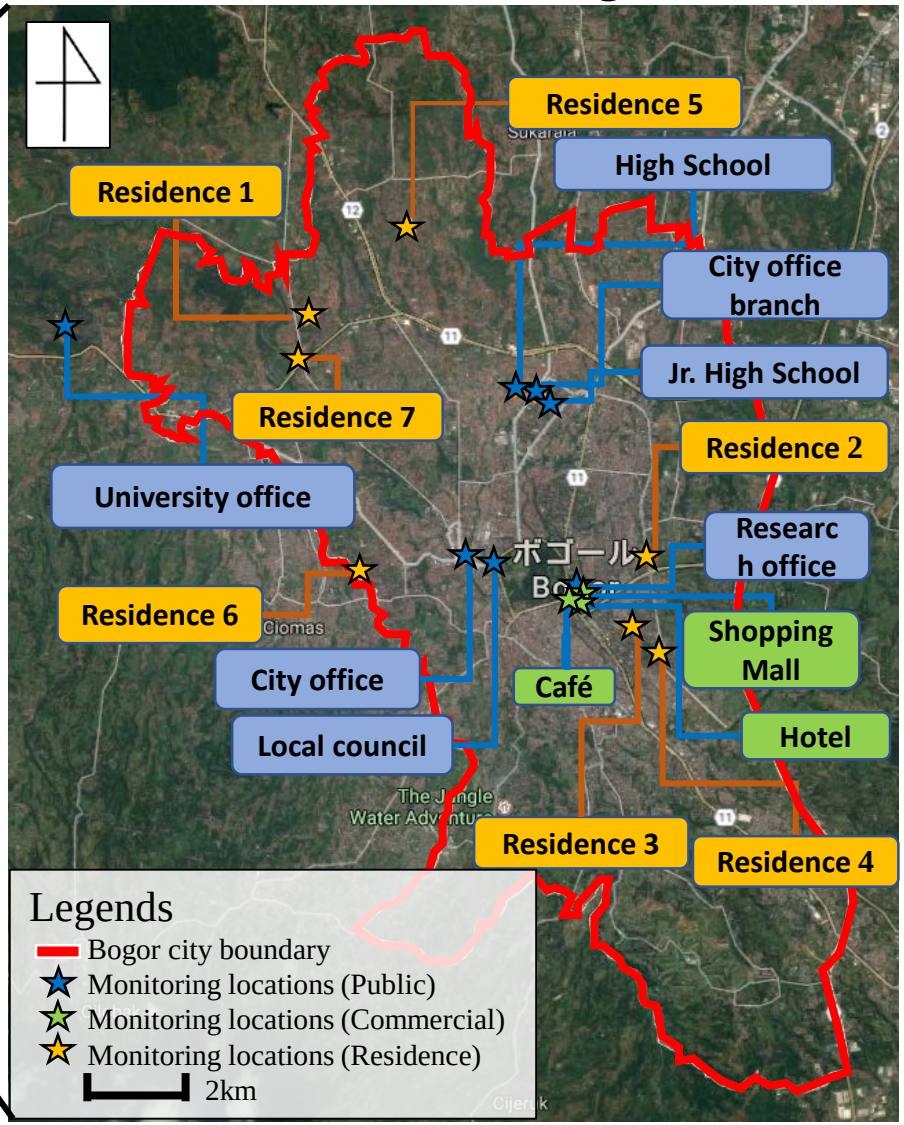
- Similar to Distributed Ledger Technology (DLC, ex. Block chain technology)
- This method could be linked inductive method and deductive method

Monitoring data analysis

Location of Indonesia Research Target



Urban Sector targets

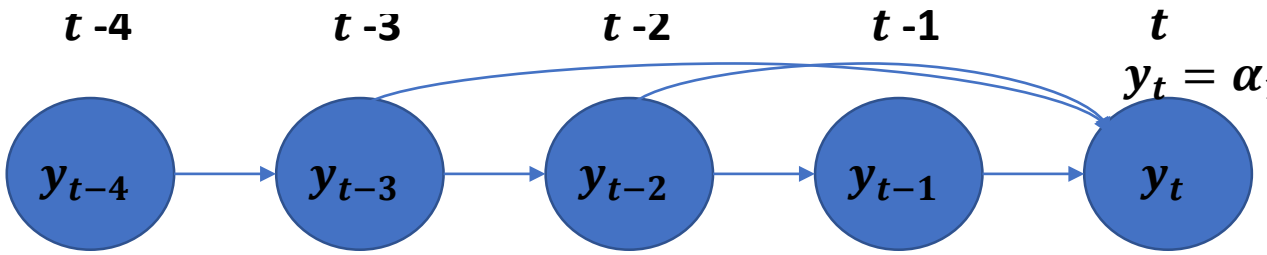


Sensor number

Factory	:	16
Government building	:	106
Residence	:	40
Commercial facilities	:	32

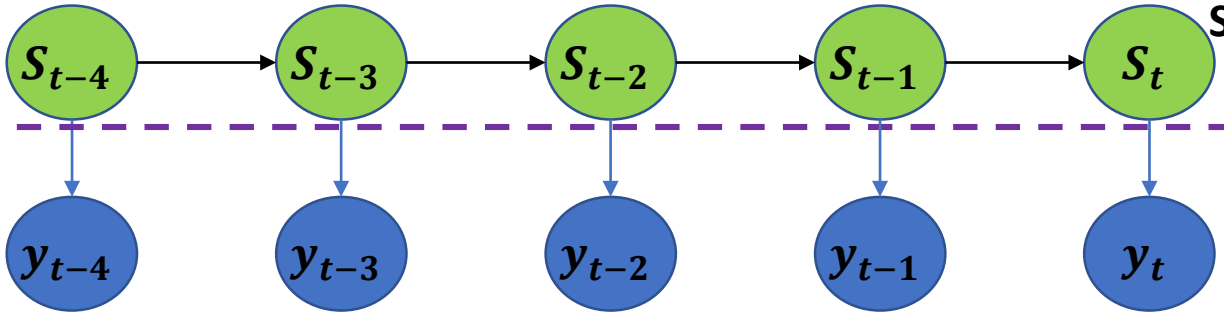
Time-series analysis methods

Auto-Regression Model



- In Auto-regression model, regression equation is created by using past data
- In ARX (Auto-Regression eXgeneous model) is considered to past exogenous data
- In this type model, all of time is predicted by only 1 model

Markov Switching Model

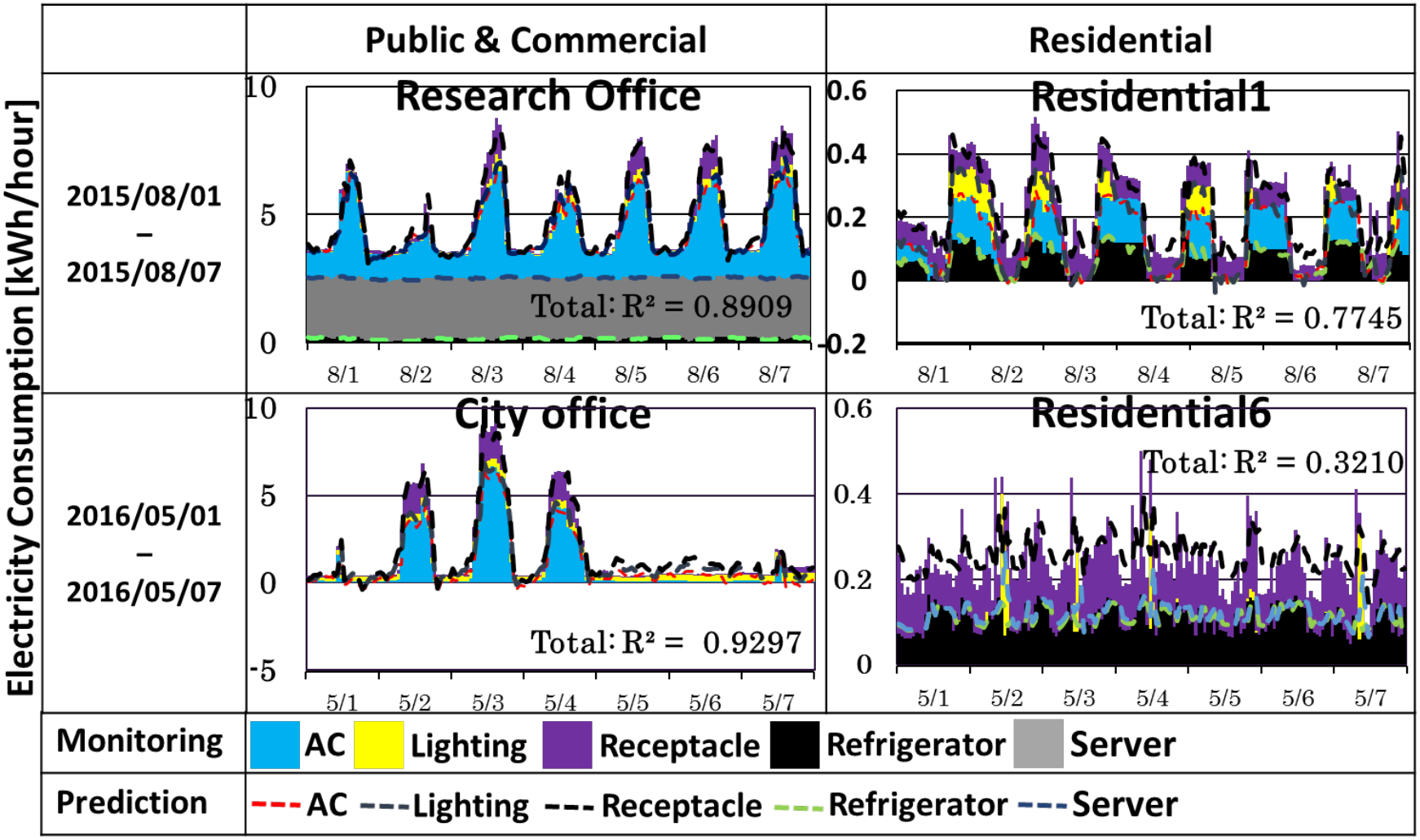


State Variable
Can't observation
Most estimate by data

Regression equation changes for each state variable
The result changes with the probability of the state variable

- Estimate state of electricity consumption and developed equation for each state
- S is State variables that shows electricity consumption pattern
- Future electricity consumption is estimated by using transition probability of S
- Prediciton is used to Conditional Auto Regression model by State variable

Prediction Result by using ARX model



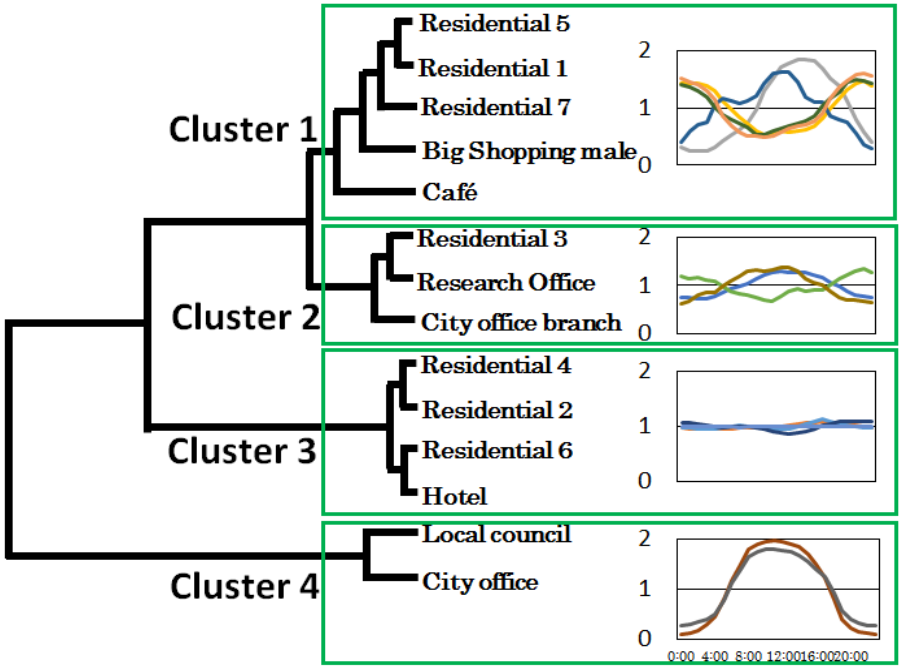
- Development included high R^2 value (over 0.7) and enable Peak prediction model
- Residences have differences bigger than others by small electricity consumption
- Weak periodicity Buildings model have weak performance of peak prediction

Time-series Clustering of Electricity Demand Pattern

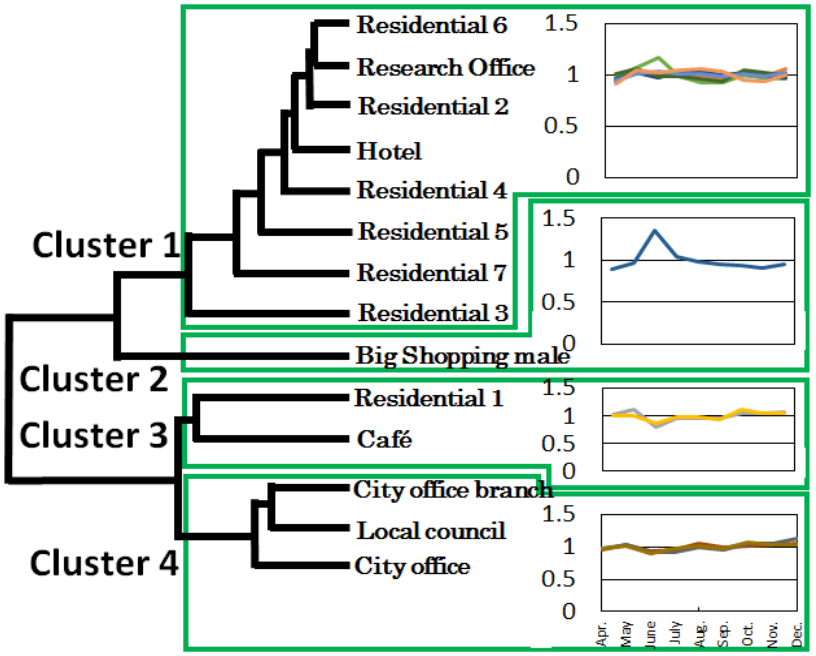
Time-series Clustering

- Time-series clustering by short & long term pattern (baseline: Average)
- Using Dynamic Time Warping (DTW) and Cluster Analysis (Furthest neighbor method)

Short term (Hourly)

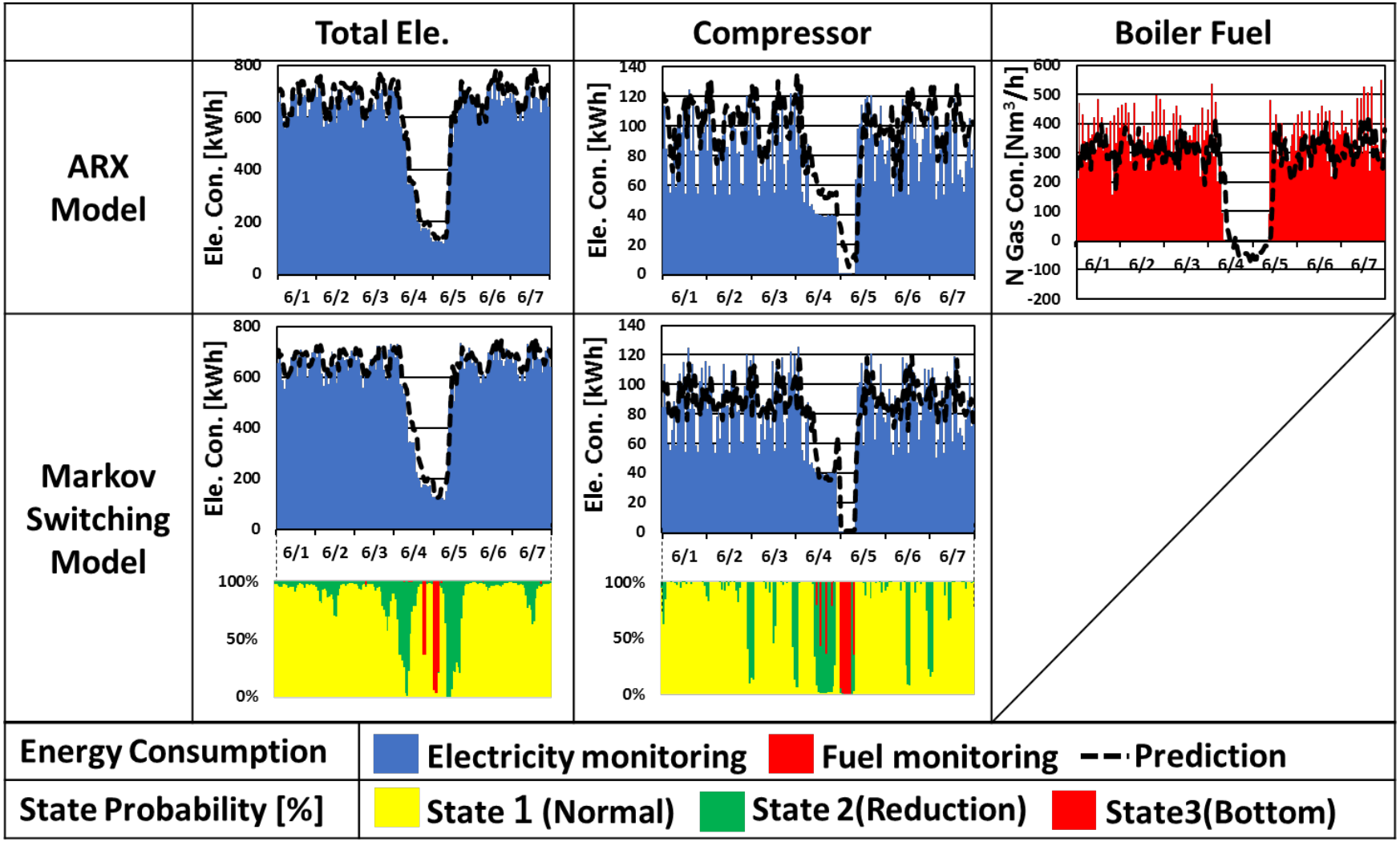


Medium-long term (daily)



- Hourly &Long term variability is divided into 4 clusters respectively by peak and non-peak electricity consumption patterns
- Electricity demand patterns were divided by Peak/Average value

Analysis Result by ARX model and Markov Switching Model (MSwM)



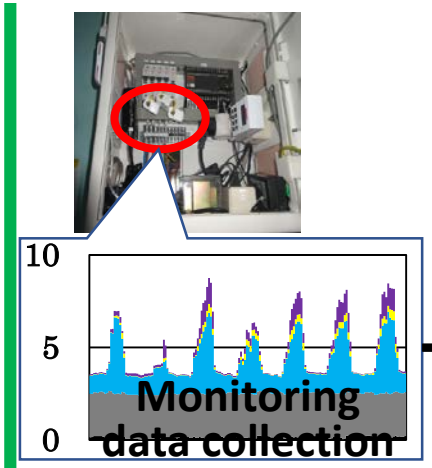
- ARX and MSwM are developed enable Peak prediction model
- MSwM were better fitting prediction than ARX model on these monitoring data
- MSwM were divided 3 another Electricity Consumption States for each hour

Data Fusion method

1. Linkage statistical and engineering method
2. Extended monitoring data for spatial data
3. Energy saving scenario for regional scale

Analysis flow of this study

Monitoring Data collection & analysis



Statistical analysis

Data fusion

Engineering analysis

Unit data & micro simulation

Unit data development

Constraint collection

DR sim

Spatio - Temporal extended

Questionnaire data

Geo information data

Data fusion

Monitoring data fitting for spatial

Data fusion

Macro Simulation

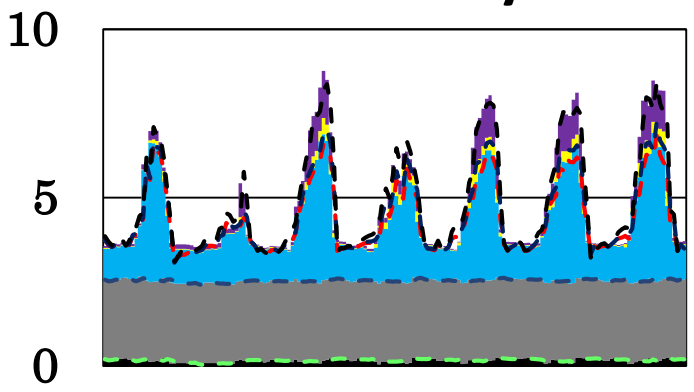
Spatio-Temporal model

ADR sim

Scenario simulation

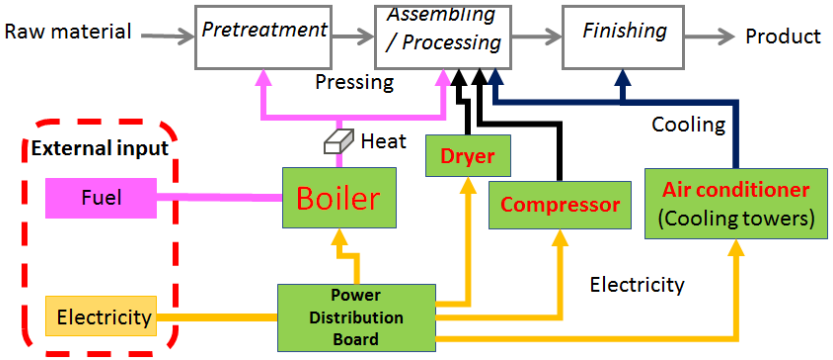
Statistical & Engineering method Data Fusion Image

Statistical analysis



- Analysis by monitoring result
- **Inductive method** analysis

Engineering method



- Analysis by operational process
- **Deductive method** analysis

Data fusion

Operation estimation by monitoring energy data

Estimation

Energy saving potential & Reduction Constraint

Linkage between statistical and engineering method ex. WWTP

Input

- Report data
- Statistics data
- Monitoring data

Model estimation

Activate Sludge model (ASM)

- Water quality modeling
- Basic proses is defined by Statistics
- Diction process profile for fitting water quality report

Minimization Electricity consumption Differential

- Estimate electricity consumption by use each process unit data
- Fitting not only water quality but also each time electricity demand
- Arrangement correlation process & State variables

Markov Switching model

- Electricity prediction modeling
- Monitoring data (Real data)
- Estimate each time Latent state profile → Maybe relation process

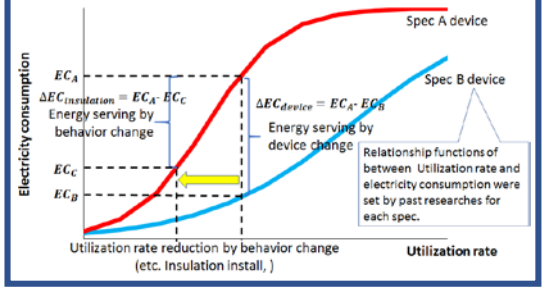
Output

Database of Process prediction by monitoring data

We could estimate future electricity demand, water quality & Process condition

Process Optimization

Differential between optimized state and real one is Energy saving Potential

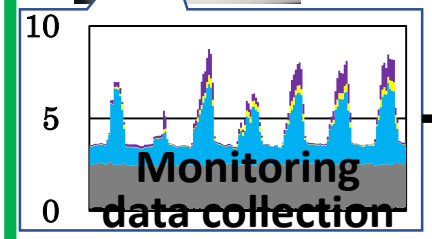


Data Fusion method

1. Linkage statistical and engineering method
2. Extended monitoring data for spatial data
3. Energy saving scenario for regional scale

Analysis flow of this study

Monitoring Data collection & analysis



Statistical analysis

Data fusion

Engineering analysis

Unit data & micro simulation

Unit data development

Constraint collection

DR sim

Spatio - Temporal extended

Questionnaire data

Geo information data

Data fusion

Monitoring data fitting for spatial

Data fusion

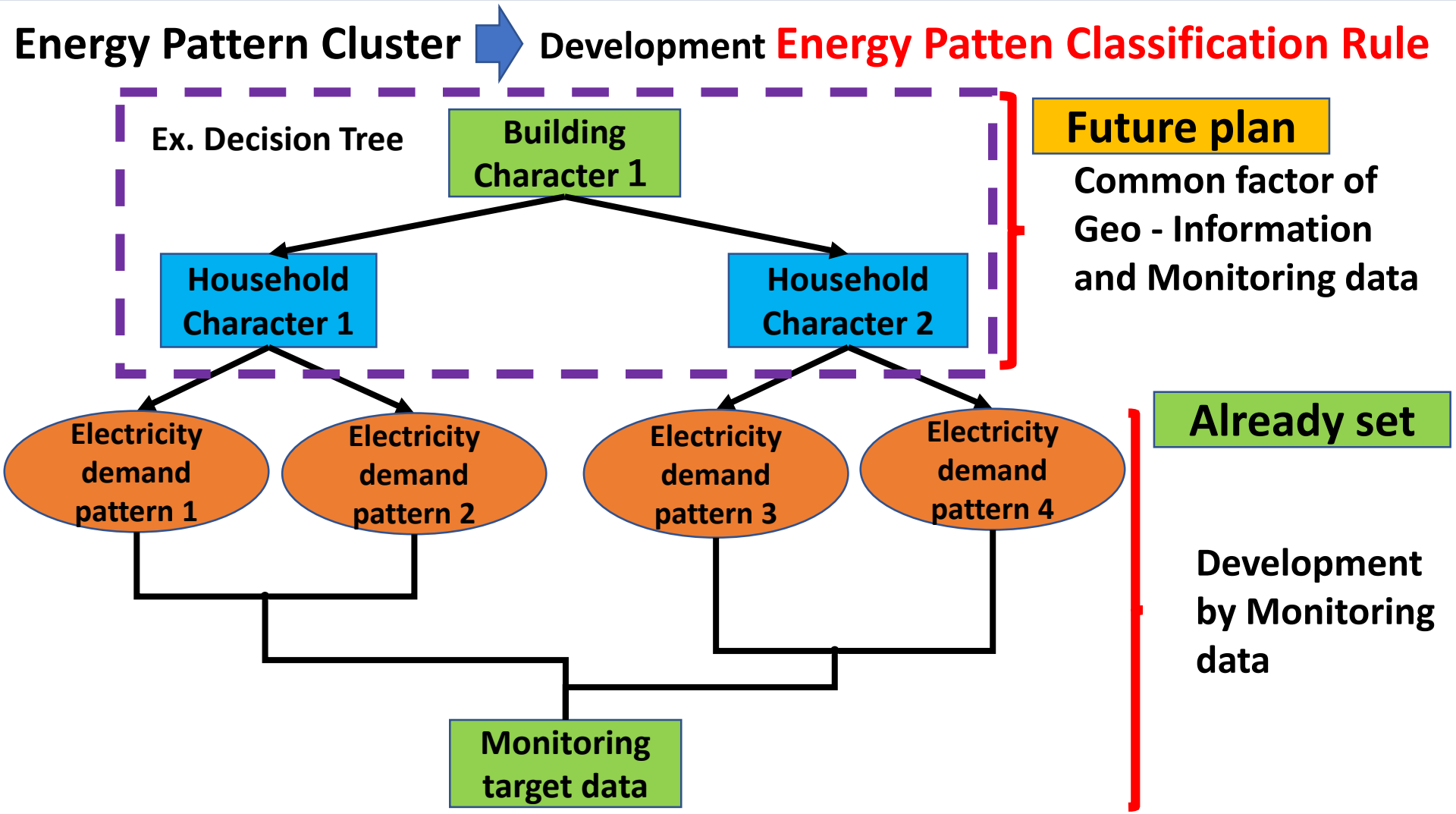
Macro Simulation

Spatio-Temporal model

ADR sim

Scenario simulation

Data fusion image between monitoring & another type data



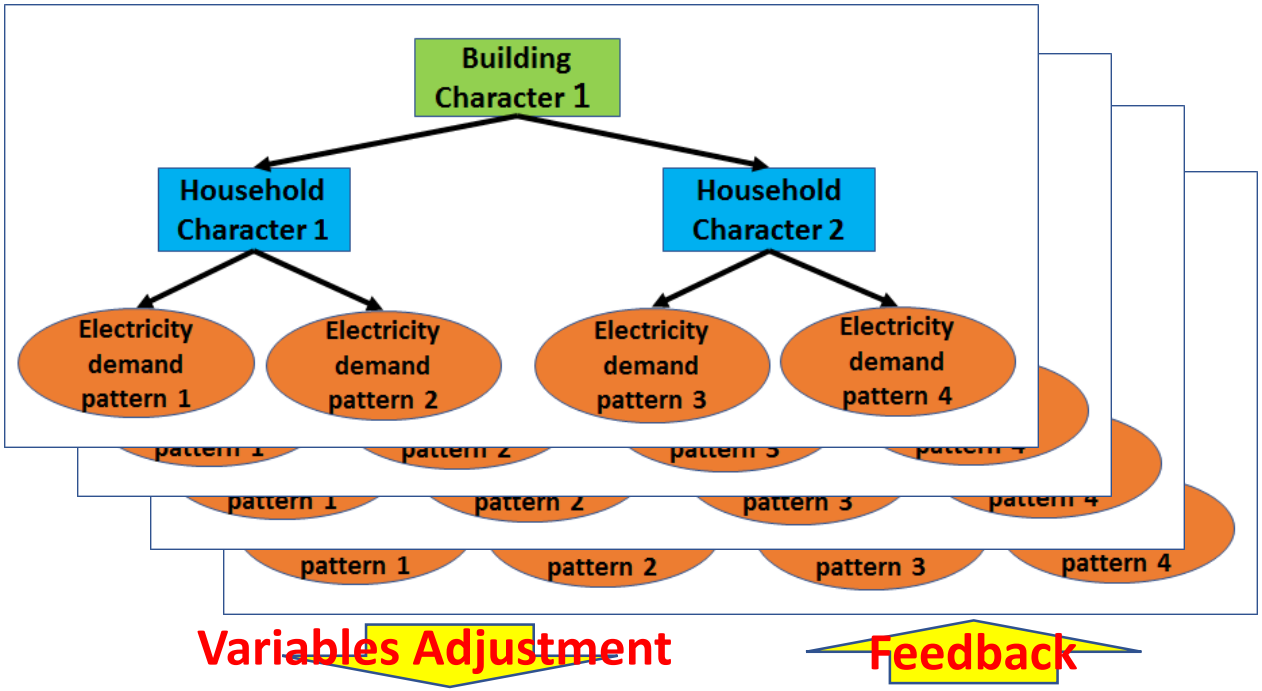
- Development Energy Patten Classification Rule by Monitoring Targets
  → Assuming that this rule could be applied More Extensively Field
- This Rule is essential for Monitoring Result Expanding

Inage of Data fitting by Ensemble learning method

Accuracy improvement of classification by Ensemble Learning Method

- ✓ Our Monitoring Samples number are **Limited**
- ✓ We could many of Classification rule by **non-Robust Sample number**

- ➢ **Setting Ensemble** of many of Classification Rule
- ➢ Improvement each Classification rule by **fitting exogenous variables**
- ➢ Improvement **Ensemble fitting value** likely to Genetic algorithm method



Ex. Random Forest

Ensemble of Energy Pattern Classification Rule by Geo-information

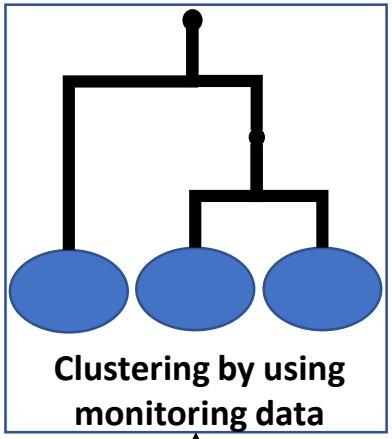
Minimize: Deferential of Statistics data or Macro analysis Result

Expanded model by using monitoring and another type data

Expanded energy demand model by using Machine Learning

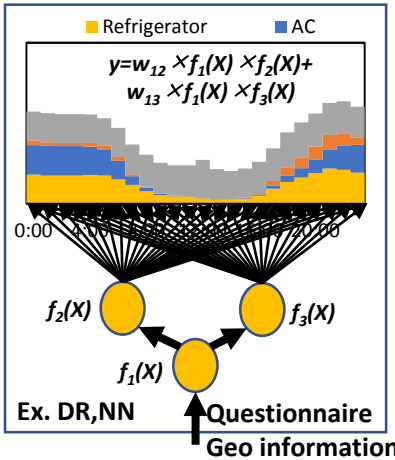
Monitoring analysis phase

①Cluster analysis



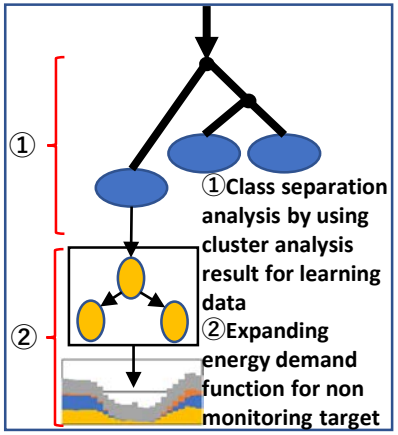
Clustering by using monitoring data

②Energy demand function

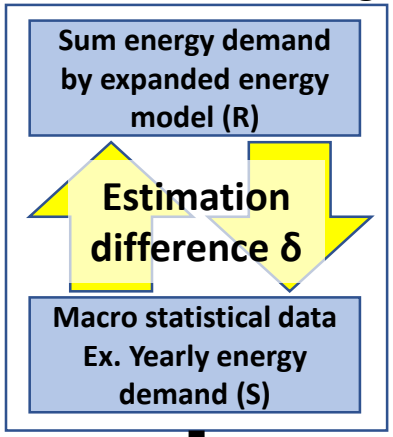


Macro data fitting phase

③Class separation



④Macro data fitting



$\delta = |R-S|$

$\delta \leq \varepsilon$

No

Reset variables of Cluster analysis & Class separation

Ex. Random forest

Yes

Finished

※ε：Acceptable error (set by user)

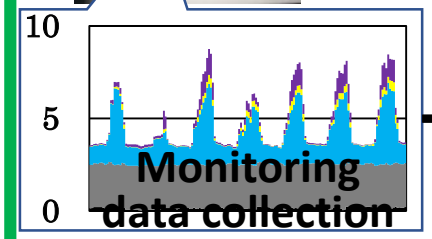
This method is one of a Semi-Supervised Learning

Data Fusion method

1. Linkage statistical and engineering method
2. Extended monitoring data for spatial data
3. Energy saving scenario for regional scale

Analysis flow of this study

Monitoring Data collection & analysis



Statistical analysis

Data fusion

Engineering analysis

Unit data & micro simulation

Unit data development

Constraint collection

DR sim

Spatio - Temporal extended

Questionnaire data

Geo information data

Data fusion

Monitoring data fitting for spatial

Data fusion

Macro Simulation

Spatio-Temporal model

ADR sim

Scenario simulation

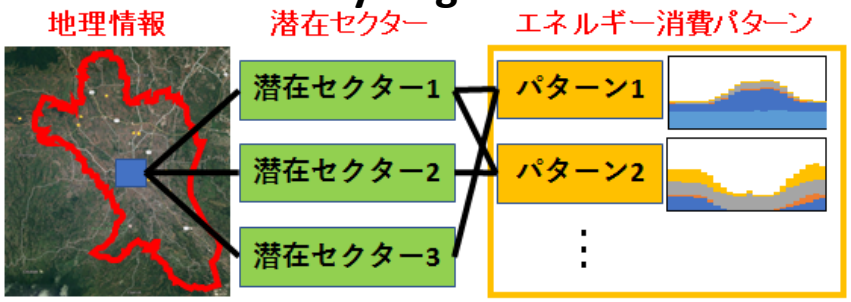
Micro - Macro Linkage model Future Image

Dvelopment Time-series inventory

- Modelling “Average” “time-series variation”
- Modelling Energy serving by DR, change device etc

Spatial extend

Spatial extend model by Bogor data

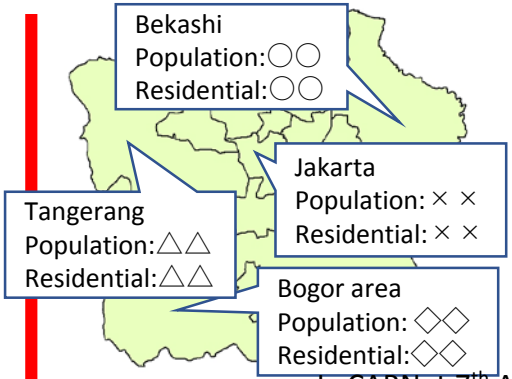


ベイジアンネットによる地理情報からのパターン推定を確率モデル化

- Estimate Latent sector by using geo information and questionnaire.
- Link monitoring results to spatial
- Development probability model by Bayesian net
- Estimate Spatio-Temporal energy consumption

Spread out

Extend for Jabodetabek

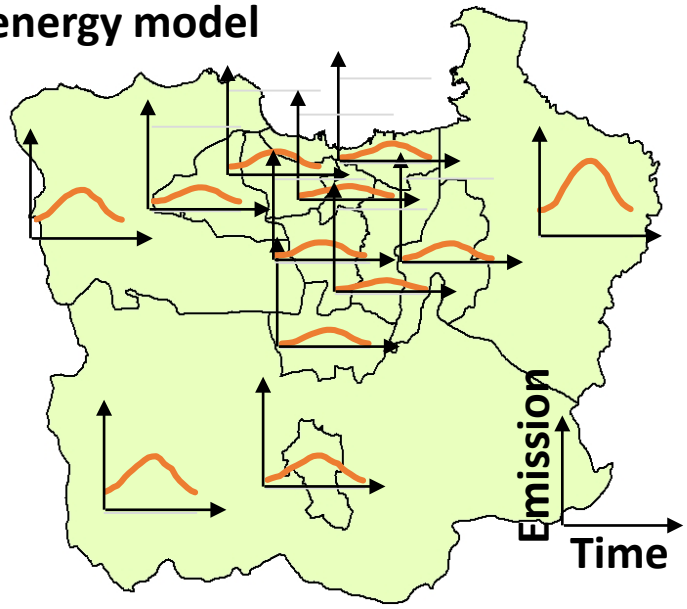


- Collect Jabodetabek region data likely to statics
- Using Probability spatio-temporal model by developed Bogor data
- Estimate Japodebek energy consumption

Serving effect

Regional estimate

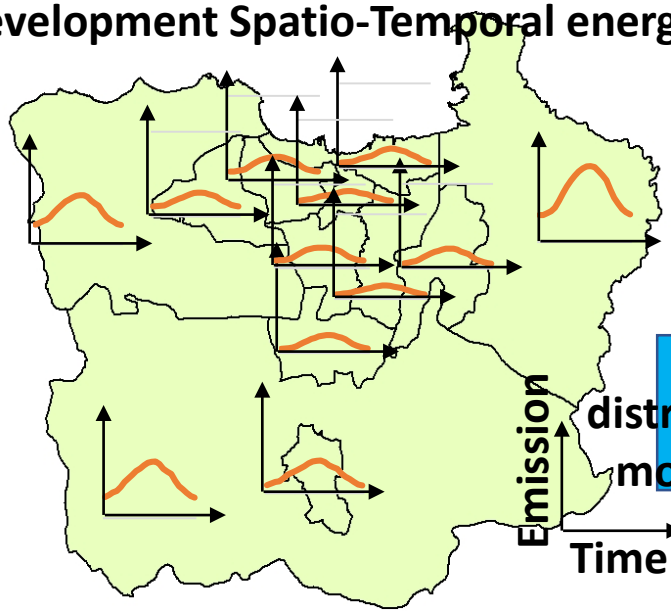
Development Spatio-Temporal energy model



- Linking data are “Average” “Daily variation” “Term variation”.
- Visualize regional spatio-temporal energy consumption.
- Creation CO_2 emission potential map.
- Development temporal serving energy behavior inventor.
- Development probability model of serving energy behavior utilization rate.
- Modelling CO2 emission point, and Linkage GOSAT and Cos Satellite site monitoring result

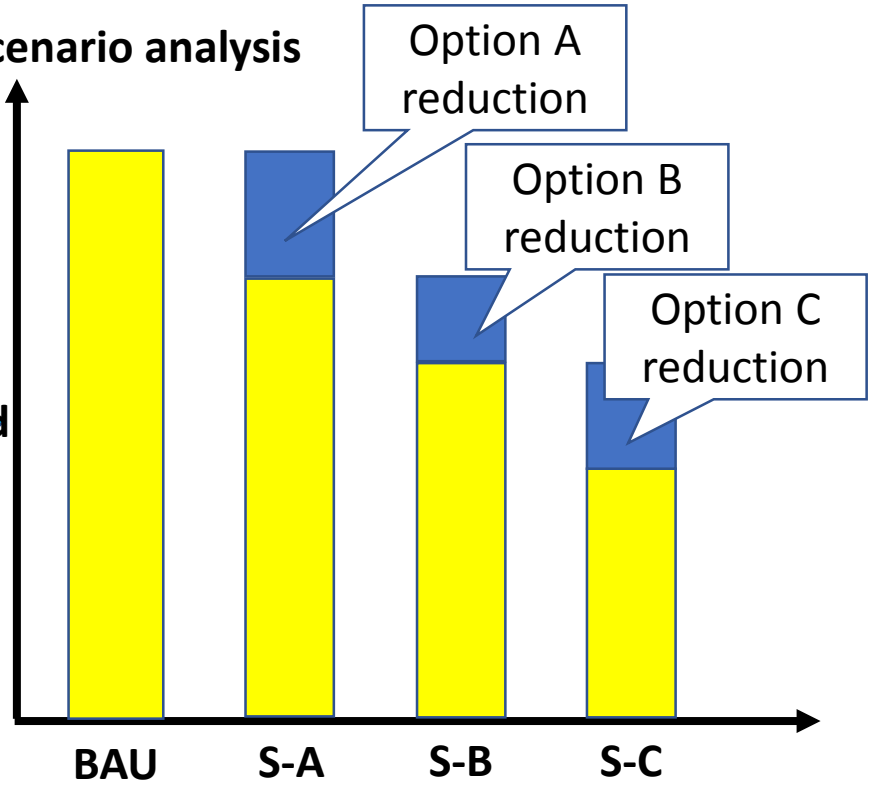
Micro - Macro Linkage model output

Development Spatio-Temporal energy model

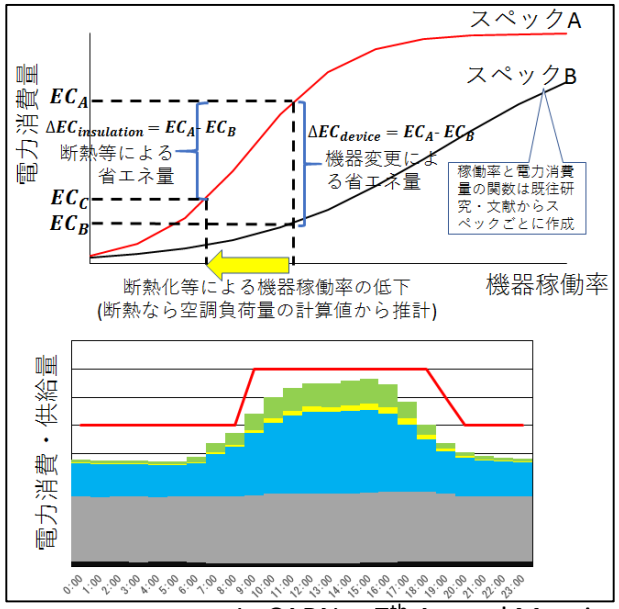


This energy distribution is based monitoring result

Scenario analysis



Energy saving & Demand response function



Including Micro energy saving analysis

- **Micro detail data**
- **Demand response constraints**
- **Time-series data for spatial map**

→ **Could Refer to many of Option**

Estimation many of Detailed energy saving scenario

Thank you for your attention

maki.seiya@nies.go.jp