

Small Area Estimation Modelling for SDGs Indicators

Background



BPS Statistics Indonesia as a national statistics office, support the successful implementation of the sustainable development agenda by providing data that can be used to support the SDGs in Indonesia. The process of producing the data by carried out through regular surveys. However, the implementation of the survey is not able to directly produce the required SDGs indicators



Problem

Lack of samples for indicators estimation with the population. It cause the estimation results will have low precision



Solution

- ❖ Increase the samples: timely, higher cost, more human resources
- ❖ Modeling (*Small Area Estimation*): solution

Small Area Estimation

Definition

Small Area Estimation (SAE) is a statistical technique to estimate the parameters of a subpopulation with a small sample size.

Small Area Estimation itself 'borrows' the power of the accompanying variables which are population in nature to increase the precision of the indicator estimation results in small areas/domains.

Approach

EBLUP

Hierarchical
Bayes

Measurement
Error



Target Indicators

- 1 Percentage of Children < 18 Years Old Living Below the Poverty Line
- 2 Percentage of Infants Age Less than 6 Months Who Get Exclusive Breastfeeding
- 3 Proportion of ever-married women aged 15-49 years whose last delivery was assisted by trained health personnel
- 4 Proportion of ever-married women aged 15-49 years whose last delivery was in a health facility
- 5 Percentage of children aged 12-23 months who received complete basic immunization
- 6 Percentage of Mothers Who Delivered Live Born Children (ALH) in the Last Two Years and Last ALH Born with Low Birth Weight
- 7 Participation rate in organized learning (one year before primary school age)
- 8 Gross Enrollment Rate (GER) for Early Childhood Education (PAUD)
- 9 Percentage of children aged 24-59 months who have participated in the Early Childhood Education program
- 10 Education Completion Level (TPP) SD, SMP, SMA
- 11 Proportion of children under 5 years of age whose births are registered by civil registration institutions

Observation Unit

	Observation Unit
Indicator 1	Children < 18 Years Old
Indicator 2	Infants Age Less than 6 Months
Indicator 3	ever-married women aged 15-49 years
Indicator 4	ever-married women aged 15-49 years
Indicator 5	children aged 12-23 months
Indicator 6	Live Born Children (ALH) in the Last Two Years
Indicator 7	children aged 6 years old
Indicator 8	children aged 3-6 years old
Indicator 9	children aged 24-59 months
Indicator 10	School-age population for every level education
Indicator 11	children under 5 years

Level Estimation: Regency/City in Indonesia
Number of Regencies/Cities: 514

Stages

Preparation

- Identification of Data Availability
- Literature Study

Data Setup

- Setup Live Estimation Data
- Setup data Auxiliary Variables

SAE Pre-Modelling

- Data Exploration
- Variable Selection
- Identify the Appropriate SAE Model

SAE Modeling

- SAE Modeling
- RSE Check
- Result Analysis

Compilation Report

- Preparation of reports

SAE Model Plan which will be used

- SAE EBLUP
- SAE Measurement Error
- SAE HB Beta
- SAE with transformation
- SAE Clustering

Data Source

Data **SUSENAS** KOR and KP Module 2020



basic socioeconomic
characteristics of households
and household members

Village Potential Data **PODES** 2020

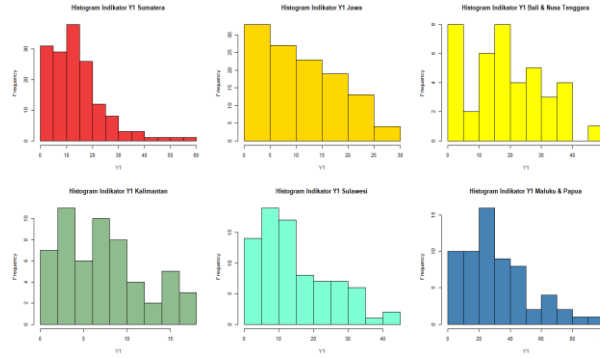


Village Characteristics

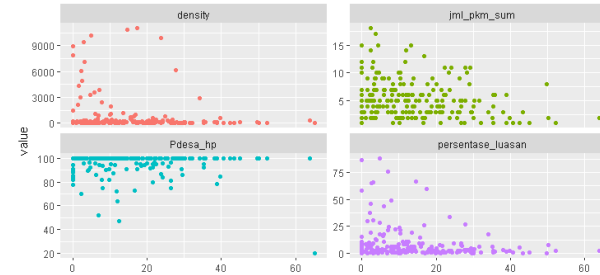
Other Data



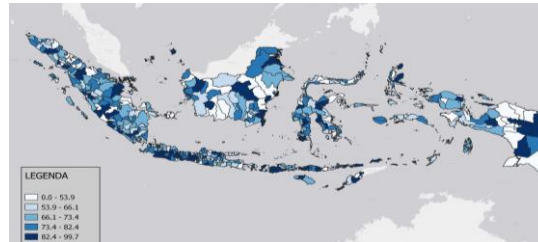
Data Characteristics



Non-Normal
Data Distribution



Non-linearity



Unsampled Observation

Data Condition at a Glance



19% Regions have RSE above 25% overall for all indicators. This RSE problem mostly occurs in Eastern Indonesia.



There are areas that are unsampled (no samples) or no samples with certain categories are found. This makes the RSE of the estimate unable to be generated



Some indicators are not normally distributed at the national level. So that another SAE model is used and regional grouping is carried out.

Analysis Results

Percentage of Children < 18 Years Old Living Below the Poverty Line

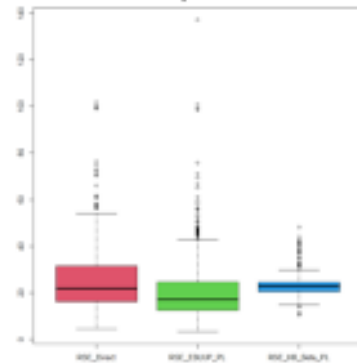
5 number summary Estimasi

Statistik	Direct	EBLUP Pulau	HB Beta Pulau
Min	0	0.060	2.290
1 st Qu.	5.833	5.933	8.219
Median	11.974	11.650	12.307
Mean	13.995	14.579	15.112
3 rd Qu.	20.081	19.642	17.750
Max	91.553	80.519	63.776

5 number summary RSE

Statistik	Direct	EBLUP Pulau	HB Beta Pulau
Min	4.499	3.42	10.58
1 st Qu.	16.247	12.65	20.66
Median	21.884	17.04	22.68
Mean	26.002	21.16	23.34
3 rd Qu.	31.437	24.77	24.57
Max	101.853	137.13	48.16
NA	2		

Boxplot RSE



Line chart perbandingan RSE



Jumlah kabupaten/kota menurut kelompok nilai RSE

	Direct	EBLUP Pulau	HB Beta Pulau
RSE ≤ 25%	305	348	432
25% < RSE ≤ 50%	169	143	82
RSE > 50%	38	23	0
NA	2		

Analysis Results

Proportion of ever-married women aged 15-49 years whose last delivery was assisted by trained health personnel

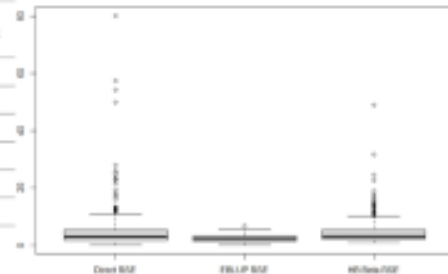
5 number summary Estimasi

Statistik	Direct	EBLUP	HB Beta
Min	0.04613	0.4997	0.05097
1 st Qu.	0.86565	0.9151	0.87811
Median	0.95095	0.9518	0.93326
Mean	0.89856	0.9389	0.89184
3 rd Qu.	0.97623	0.9774	0.95832
Max	0.99909	0.9991	0.98793

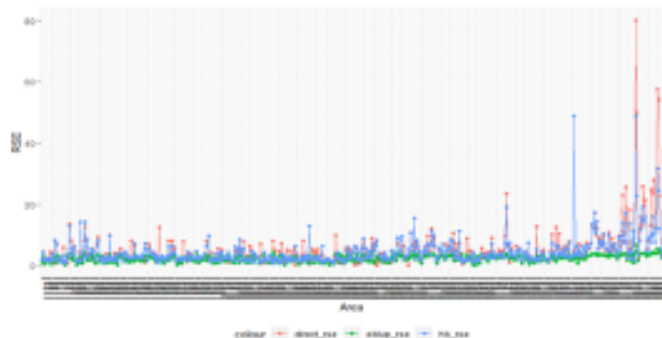
5 number summary RSE

Statistik	Direct	EBLUP	HB Beta
Min	0.09209	0.088	0.7604
1 st Qu.	1.55743	1.509	2.1414
Median	2.89744	2.327	3.0662
Mean	4.56554	2.333	4.4027
3 rd Qu.	5.26793	3.145	5.2565
Max	80.16078	6.719	48.8772

Boxplot RSE



Line chart perbandingan RSE



Jumlah kabupaten/kota menurut kelompok nilai RSE

	Direct	EBLUP	HB Beta
RSE ≤ 25%	505	514	511
25% < RSE ≤ 50%	4	0	3
RSE > 50%	3	0	0
NA	2	0	0

Results Modeling

No	Indicator	Number of Regions with RSE < 25%	
		Direct Estimate	SAE Modeling
1	Percentage of Children < 18 Years Old Living Below the Poverty Line	305	404
2	Percentage of Infants Age Less than 6 Months Who Get Exclusive Breastfeeding	397	514
3	Proportion of ever-married women aged 15-49 years whose last delivery was assisted by trained health personnel	507	514
4	Proportion of ever-married women aged 15-49 years whose last delivery was in a health facility	500	513

Results Modeling

No	Indicator	Number of Regions with RSE < 25%	
		Before Modeling	After Modeling
5	Proportion of children under 5 years of age whose births are registered by civil registration institutions	501	509
6	Percentage of children aged 24-59 months who have participated in the Early Childhood Education program	215	59
7	Level of Completion of Education (TPP)		
	- SD	514	514
	- JUNIOR HIGH SCHOOL	512	514
	- SENIOR HIGH SCHOOL	495	514
8	Percentage of children aged 12-23 months who received complete basic immunization	399	482

Results Modeling

No	Indicator	Number of Regions with RSE < 25%	
		Before Modeling	After Modeling
9	Percentage of Mothers Who Delivered Live Born Children (ALH) in the Last Two Years and Last ALH Born with Low Birth Weight	50	274
10	Participation rate in organized learning (one year before primary school age)	509	512
11	Gross Enrollment Rate (GER) for Early Childhood Education (PAUD)	491	514

Results Modeling



Increased precision

The results from SAE show that the RSE from a high direct estimate can be reduced by SAE, so that more areas have RSE below 25%



Estimate in the non-sample region

Non-sample areas that cannot be estimated using direct estimation can be estimated using *clustering*

Groundcheck Results



Objectives: to check the suitability of variables and comparisons between regions



Methods: Focus Group Discussion with city government and indepth-Interview with local government



- **Result:** the majority of informants think that the auxiliary variables used in SAE modeling are quite appropriate to use. But there are still some differences of comparisons between regions based on SAE modeling results and local government information

Opportunity Next



Additional data

Utilization of explanatory variables from other data to improve estimation precision (sectoral data)



Spatial factor

It is worth trying to incorporate spatial information into SAE modeling



ground check

It is necessary to validate and verify the current SAE modeling results through the method *ground check* appropriate



Data consistency with BPS official release

Further validation is needed regarding the estimation results to be consistent with official BPS releases at the provincial and national levels



SUSTAINABLE
DEVELOPMENT
GOALS

THANKYOU



BADAN PUSAT STATISTIK

