International Journal of GEOMATE, June, 2025 Vol.28, Issue 130, pp.25-34
ISSN: 2186-2982 (P), 2186-2990 (O), Japan, DOI: https://doi.org/10.21660/2025.130.4780
Geotechnique, Construction Materials and Environment

ENHANCED STRATIFIED SAMPLING WITH REMOTE SENSING DATA FOR SOLID WASTE PROJECTS AND RESEARCH SURVEYS

Elprida Agustina[1]*, Emenda Sembiring[1], Anjar Dimara Sakti[2,3], Monica Iyanuar[4]

[1]Air and Waste Management Research Group, Faculty of Civil and Environmental Engineering, Institut Teknologi Bandung, Indonesia;
[2]Geographic Information Sciences and Technology Research Group, Faculty of Earth Sciences and Technology, Institut Teknologi Bandung, Indonesia;
[3]Center for Remote Sensing, Institut Teknologi Bandung, Bandung, Indonesia;
[4]Regional Technical Implementation Unit for Waste Management of Bandung City, Bandung, Indonesia.

*Corresponding Author, Received: 28 Oct. 2024, Revised: 13 Feb. 2025, Accepted: 20 March 2025

ABSTRACT: Surveys are commonly conducted for projects or research related to solid waste management. Stratified sampling allows researchers to obtain a more representative sample from a diverse population, thereby enhancing accuracy and precision, increasing efficiency, and facilitating comparisons between groups. Remote sensing product data refers to information acquired through remote sensing technologies, including socio-economic conditions. In this study, remote sensing product data were utilized to develop an index or classification of areas potentially generating waste and plastic waste. A sample of 400 households was collected, divided into two sampling approaches: 200 households were selected through stratified sampling based on the remote sensing data indexing, while 200 households were chosen randomly. Data distribution analysis was conducted using descriptive statistics, box-plot analysis, and histograms. The data distribution from stratified sampling tends to be more normal, as indicated by a histogram that is symmetrical and bell-shaped (normal distribution) and by a lower number of outliers compared to random sampling. Additionally, hypothesis testing was conducted using t-tests to examine significant differences between the two sampling methods, revealing a significant difference in waste and plastic waste generation data. Modeling using Linear Regression, Random Forest, and XGBoost was conducted, yielding MSE, RMSE, and MAE results. The stratified sampling was better (MSE: 0.13, 0.01, and 0.00, respectively) compared to random sampling (MSE: 0.23, 0.04, and 0.00, respectively). The results show that stratified random sampling, with class divisions based on remote sensing product data, yields more normalized data, less error, and greater stability for modeling analysis.

Keywords: Remote sensing data product, Solid waste surveys, Stratified sampling, Waste sampling

1. INTRODUCTION

Waste is a natural product of human existence, just like water, food, and air, we cannot live without generating any waste [1]. Waste is one of the problems that has been highlighted worldwide from year to year [2]. The amount of waste increases along with the increase in population. The use of products in daily life produces waste that is a source of problems for the environment [3]. Waste becomes a source of problems if not managed properly. To solve the waste problem, many researchers conduct research related to waste technology and waste management.

In conducting waste research, researchers often need valid data, especially primary data. Primary data is usually obtained through observation or surveys. Surveys in waste management projects and research are very necessary to obtain data. There are two ways to conduct a survey to obtain waste management data, the first is through a questionnaire [4, 5] and the second is through interviews [4, 6]. Surveys are conducted to obtain data such as:

- Demographic data (gender, age, etc.) [5];
- Socio-economic data (household size, income, education, occupation, building structure, etc.) [7-9];
- Waste generation data (amount, volume, type, density) [8, 9];
- Waste management data (reduction, collection, separation, treatment, disposal) [11-15];
- Respondent's opinions regarding waste problems [16];
- Respondent's intentions in managing waste [17, 18].

Vassanadumrongdee, S., et al., 2018, conducted a questionnaire survey to investigate the status of separation intention and willingness to pay of Bangkok residents, and face-to-face interviews were used in some cases where the respondents reported difficulty in reading the questionnaire. Respondents are very much needed in conducting a survey. There are two categories of sampling techniques in determining respondents, namely probability sampling techniques and non-probability sampling techniques. Probability sampling techniques include simple random sampling, systematic random sampling, stratified random sampling, and cluster sampling. Non-probability sampling techniques include purposive sampling, quota sampling, snowball sampling, convenience sampling, and self-selection sampling [19]. Probability sampling technique is most commonly found in waste research.

Grazhdani, D., 2016 and Leknoi, U., et al., 2024 were using simple random sampling to collect the data. The simple random sampling technique means that every member of the population has the same probability of being selected for the survey [17, 20-21]. Hu, J., et al., 2024; Cheng, X., et al., 2024; Zhang, S., et al., 2024, were using stratified random sampling [22-24]. Stratified random sampling is usually used to obtain a good representative sample by dividing the population into several subgroups that are individually more homogeneous than the whole population and selecting certain items from each stratum to generate a sample [25]. Lazo, D. P. L., et al., 2022, and Xu, L., et al., 2017 were using systematic random sampling. It has an improvement over the simple random sampling by selecting the sample from the population through a systematic rule [26-27].

In this study, the researcher used both stratified and random sampling to obtain respondents. Stratified sampling was used to ensure that each stratum of the population was considered with a certain proportion of people of different characteristics. It gives better accuracy from each of the strata, and also gives more detailed and reliable information about the sample [25]. The random sampling was also conducted to make a comparable output and decide the best performance sampling techniques.

Several researchers have previously used remote sensing techniques to solve waste management problems globally [28-30]. The researcher also used the remote sensing technique in this study. Remote sensing system provides information about the Earth's surface using satellite images that can detect and monitor land cover and land use [31]. Remote sensing uses aerial or satellite imagery to obtain data [32]. Remote sensing product data depicting physical environmental conditions such as vegetation, major roads, spatial topology, soil, etc. Haripavan, N., et al., 2023, were using remote sensing techniques to identify the suitable locations for solid waste disposal in urban areas, specifically focusing on the municipality of Gudivada. Fraternali, P., et al., 2024, were detecting illegal solid waste disposal sites by remote sensing [28, 30, 33-35].

Remote sensing data products have been utilized in recent studies to integrate socio-economic variables for modeling plastic waste flows into the ocean and projecting plastic waste generation at the source [36-37]. In this study, remote sensing data products are used to classify the study area, facilitating a more nuanced understanding of spatial variability. The objective of this research is to evaluate the application of stratified sampling in combination with remote sensing data to enhance the representativeness and accuracy of waste generation data, thereby contributing to improved methodologies in solid waste management research and projects.

2. RESEARCH SIGNIFICANCE

Integrating stratified sampling with remote sensing data improves waste generation analysis by enhancing accuracy, reducing errors, and producing normalized data, which is essential for stable machine learning models. Leveraging spatial and socio-economic insights, this approach enables a deeper understanding of waste patterns, supporting targeted interventions and effective waste management policies.

3. METHODOLOGY

3.1. Variable Selection

To maximize the mobilization and distribution of sample collection, zones are defined depending on categories. A variety of geographic characteristics are used to structure the area definition. Table 1 presents the study of parameters and explanations for the classification of sampling areas.

Table 1. Variable identified from recent studies

Variable	Argument
Gross Domestic Product	The higher the GRDP/GDP in an area, the greater the waste generated [38-47]
Population density	The denser the population in an area, the greater the amount of waste generated [48-54]
Populated area/land use	The more activity there is from the community in an area, the greater the amount of waste generated [55-56]

3.2 Data Used in This Study

Each of these attributes is represented using 300 x 300-meter resolution raster maps created using remote sensing data products, which increases data accuracy. Remote sensing product data were accessed via Google Earth Engine by first defining the spatial and temporal resolution. The VIIRS Nighttime Day/Night Band Composites Version 1 data represents GDP/GRP (Gross Regional Product); the Copernicus Land Use dataset's LULC (Land Use Land Cover) data represents populated areas; and the WorldPop Global Project Population data represents population density. Figure shows how these remote sensing data products as can be seen in Table 2 and Figure 1. All remote sensing product data used in this study are reclassified to obtain distinct classes, allowing the indices to be accumulated.

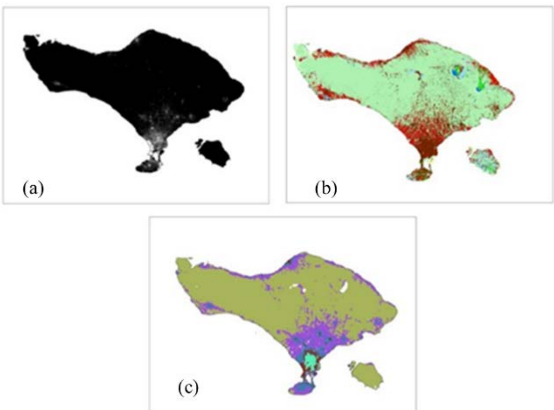
Table 2. Detailed remote sensing product data in this study

Variable	Remote Sensing Product Data	Data Source	Available Spatial Resolution	Spatial Resolution Used	Temporal Resolution
Gross Domestic Product	VIIRS Nighttime Day/Night Band Composites Version 1	[57]	463.83 m^2	30 m^2	1 January - 6 June 2023
Population density	WorldPop Global Project Population Data: Estimated Residential Population	[58]	92.77 m^2	30 m^2	1 January 2020 – 31 December 2020
Urban and Rural Area	Tsinghua FROM-GLC Year of Change to Impervious Surface	[59]	30 m^2	30 m^2	1 January 1985 - 31 December 2018

The night-time light index is divided into four classes (index 1-4) using the Natural Breaks (Jenks) reclassification method. The population density index is also divided into four classes (index 1-4), following the SNI 03-1733-2004 standard for urban residential planning. Meanwhile, the area classification is divided into two classes (index 0 and 1) to distinguish between rural and urban areas. After reclassification, the indices are summed using a raster calculator. The result is in an index range of minimum 2 and maximum 10.

This produces a map of potential plastic waste generation areas, which is then further categorized into three classes:

- Index 2–4: Low potential for plastic waste generation (Low Index Class)
- Index 5–7: Moderate potential for plastic waste generation (Middle Index Class)
- Index 8–9: High potential for plastic waste generation (High Index Class)



(a) VIIRS Nighttime Day/Night Band Composites Version 1
(b) WorldPop Global Project Population Data: Estimated Residential Population
(c) Tsinghua FROM-GLC Year of Change to Impervious Surface

Fig. 1 Remote sensing data product used in this study

Sampling was conducted on 400 households/respondents, divided into two groups: the random sampling group and the stratified sampling group. A total of 200 samples were selected randomly, while another 200 samples were taken using stratified sampling. The stratified sampling was carried out proportionally across three classes: high index, middle index, and low index according to the area categories shown in Figure 2.

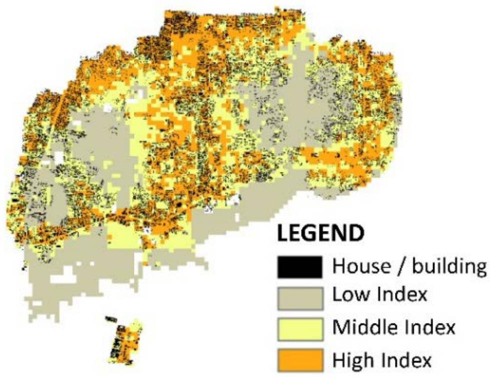


Fig. 2 The classification of sampling area

The details of variable obtained in this study as can be seen in Table 3.

Table 3. Details of variable obtained

Variable	Data Type	Unit
House area	Discrete	m^2
Education level	Continue	No unit
Income	Discrete	IDR/month
Family member	Discrete	Person
GDP	Continue	Index (No unit)
Area Characterization	Ordinal	Index (No unit)
Elevation	Discrete	Meter
Population Density	Discrete	Person/m^2

3.3 Primary Waste Generation and Composition Sampling

The sampling of plastic waste generation and composition at the source will be conducted using stratified random sampling, based on the area classification as outlined in Figure III.4. The sampling method follows the SNI-19-3964-1994 standard, which outlines methods for collecting and measuring urban waste generation and composition.

Plastic waste generation will be measured over a period of eight consecutive days. The waste composition, by weight, will also adhere to the SNI-19-3964-1994 standard. Sampling will take place in eight villages within Denpasar City, namely Pemogan Village, Pedungan Village, Sanur Kauh Village, Sidakarya Village, Sanur Village, Sesetan Village, Renon Village, and Panjer Village.

3.4 Statistical Analysis

The analysis conducted included descriptive statistics: box-plot analysis and histograms. Hypothesis testing was performed using t-tests to examine significant differences between the two sampling methods, revealing a significant difference in waste generation data. The model was applied to two target variables (Y = dependent variables) for each method, specifically waste generation, and plastic waste generation, while the rest variables listed in Table 3 represent the independent features (X).

4. RESULTS AND DISCUSSION\

The waste sampling and collection process adhered to ethical sampling practices, whereby respondents were first asked about their willingness and commitment to have their waste collected for eight consecutive days. Not all respondents initially selected through the mapping process were willing to participate, so adjustments were made in the field by replacing respondents in accordance with the predefined class or classification.

3.1 Data Distribution of Sampling and Questionnaire Results

Descriptive analysis of the sampling results using random and stratified sampling can be seen in Tables 4 and 5. Both tables indicate that the standard deviation (std) produced by the stratified sampling approach is significantly lower than that of the random sampling. A lower standard deviation implies that the data points are more consistent and predictable, making it easier to understand or make forecasts based on that data. Conversely, a higher standard deviation can indicate greater uncertainty and variability, which may complicate drawing conclusions or making predictions. A histogram is a graph that present frequencies off all variable values with vertical connected bar chart [60]. The histogram used to help visualize the distribution of variable for both the random and stratified sampling datasets.

Table 4. Descriptive analysis for random sampling

	House area	Edu cation level	Income	Family member	GDP	Area Characte rization	Elevation	Populati on Density	Waste Generation (kg)	Plastic Waste Generation (kg)
count	199.00	199.00	199.00	199.00	199.00	199.00	199.00	199.00	199.00	199.0
mean	560.20	3.88	7,011,055	4.20	2.20	0.20	271.39	17.86	1.18	0.04
std	808.34	1.08	5,856,394	1.80	2.93	0.40	330.94	15.24	0.75	0.10
min	70.00	1.00	300,000	2.00	0.01	0.00	5.00	0.00	0.26	0.00
25%	250.00	4.00	3,000,000	3.00	0.63	0.00	23.50	6.74	0.52	0.00
50%	400.00	4.00	5,000,000	4.00	1.16	0.00	92.00	11.31	0.98	0.00
75%	600.00	5.00	8,950,000	5.00	2.26	0.00	438.00	28.42	1.67	0.35
max	10000.0	6.00	30,000,000	10.00	26.22	1.00	1309.00	91.00	3.47	0.96

Table 5. Descriptive analysis for stratified sampling

	House area	Edu cation level	Income	Family member	GDP	Area Charac terizati on	Elevati on	Populatio n Density	Waste Generation (kg)	Plastic Waste Generation (kg)
count	199.00	199.00	199.00	199.00	199.00	199.00	199.00	199.00	199.00	199.0
mean	232.16	4.14	6,391,960	4.07	2.56	0.60	11.09	77.58	1.01	0.15
std	136.49	0.81	4,083,366	1.32	1.95	0.50	63.04	19.81	0.51	0.07
min	60.00	2.00	2,000,000	1.00	0.12	0.00	1.00	11.23	0.30	0.01
25%	150.00	4.00	4,000,000	3.00	1.24	0.00	5.00	63.01	0.67	0.10
50%	200.00	4.00	6,000,000	4.00	1.98	1.00	6.00	77.36	0.90	0.14
75%	300.00	5.00	6,000,000	5.00	3.59	1.00	9.00	93.86	1.24	0.18
max	800.00	6.00	20,000,000	8.00	11.41	1.00	895.00	125.64	3.59	0.52

A distribution that is symmetrical and bell-shaped (normal distribution) is generally desirable, as it indicates that the model can effectively predict outcomes across the range of values. In Figure 3, it can be observed that the data for income, GDP, and family members show a similar distribution between random and stratified sampling. However, the population density, household waste generation, and plastic waste generation in households exhibit different data distributions. The significant difference in data distribution is suspected to be due to the inadequate division of classes/indices, which does not accurately represent the population classes. The presence of classes with overly large ranges leads to a reduction in the opportunity to capture maximum or minimum data within those classes. Nevertheless, stratified sampling provides a more evenly distributed result.

Boxplots are a graphical representation used to display the distribution of a dataset and identify its key statistical characteristics, including the median, quartiles, and potential outliers. Boxplots are effective for looking at distributions over time by creating a series of boxplots, or for comparing subgroups by generating side-by-side boxkplots [61]. They visually summarize the central tendency, variability, and symmetry of the data by showing the minimum and maximum values, the interquartile range (IQR), and any extreme values that fall outside of the expected range [62]. Box-plot analysis for both sampling approaches was conducted, and as shown in Figure 4, the data distributions obtained from the two methods are quite different. From the results obtained, both approaches produced outliers, indicating that data needs to be removed first to improve the data distribution. The presence of outlier data can complicate conclusions and introduce bias into the analysis. The number of outliers in random sampling is 14 (fourteen) data points, while in stratified sampling, there are 6 (six) data points.

The variance analysis reveals that Stratified Sampling has significantly lower variance for both waste generation (0.26) and plastic generation (0.0056) compared to Random Sampling (variance of 56.68 and 1.07, respectively). This indicates that Stratified Sampling produces more consistent and balanced data, likely preserving proportions from different subgroups, which can improve model stability and reduce the influence of outliers. In contrast, Random Sampling, with its higher variance, reflects a broader spread of data, which may capture a wider diversity of values but can also result in less consistent representation across subgroups. Therefore, Stratified Sampling is preferable if the goal is to achieve more stable and accurate predictions by minimizing data variability, while Random Sampling may still be useful for capturing a more diverse range of information.

The Shapiro-Wilk test evaluates whether a dataset conforms to a normal distribution, with the null hypothesis assuming normality. A p-value below the significance threshold (commonly 0.05) indicates rejection of the null hypothesis, demonstrating that the data is not normally distributed.

The Shapiro-Wilk test result is:

- Random Sampling: Shapiro Result (statistic=0.41, pvalue=5.64 e-25)
- Stratified Sampling: Shapiro Result (statistic=0.91, pvalue=1.66e-09)

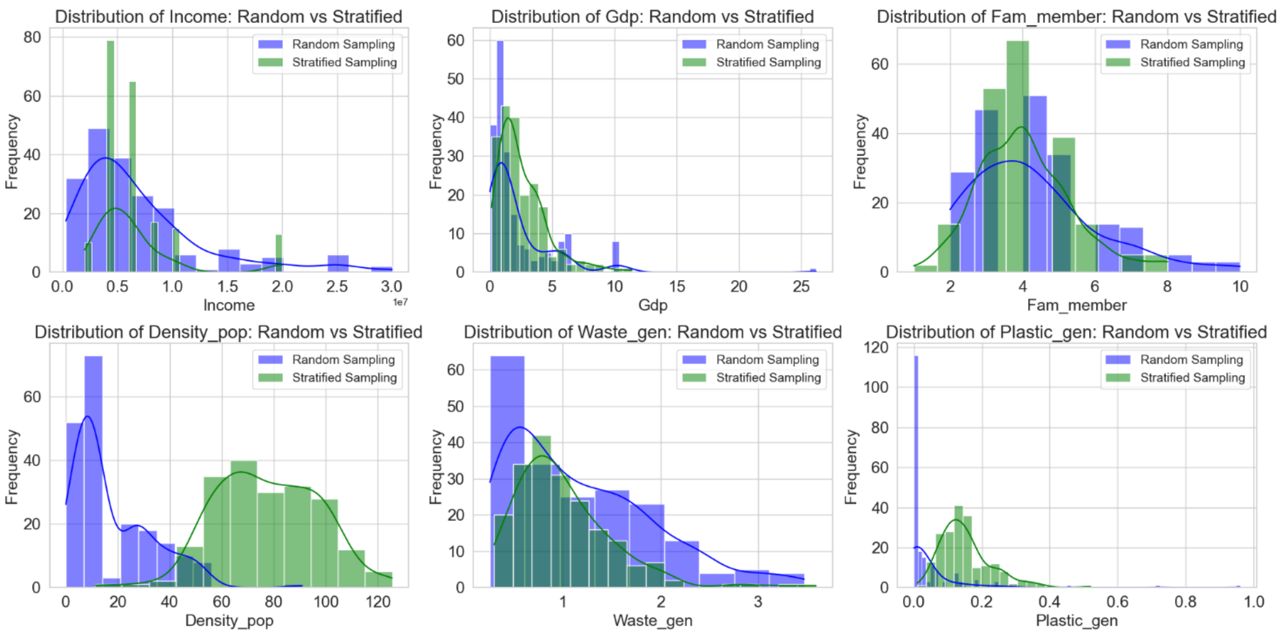


Fig. 3 Data distribution for random and stratified sampling

In this analysis, the Random Sampling dataset exhibits an extremely small p-value (5.64e-25), firmly rejecting normality. Similarly, the Stratified Sampling dataset, with a p-value of 1.67e-09, also rejects the null hypothesis, indicating it is not normally distributed. Nevertheless, the larger p-value for Stratified Sampling suggests its distribution is comparatively closer to normal than that of Random Sampling. These findings highlight that while both datasets deviate significantly from normality, Stratified Sampling may offer a marginally better approximation to a normal distribution.

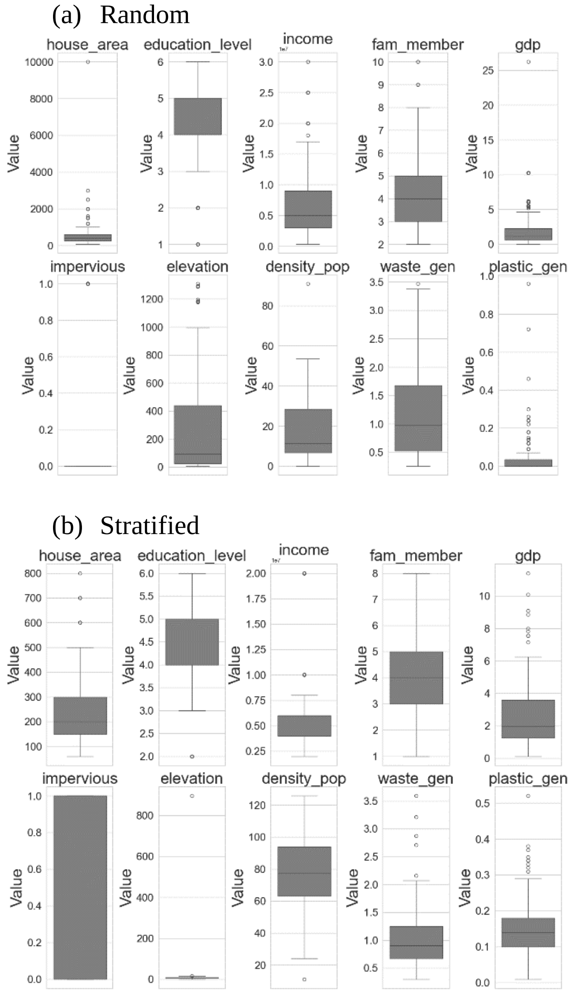


Fig. 4 Box-plot analysis for random and stratified sampling

3.2. Hypothesis Testing

Hypothesis tests are used to make conclusions about all variables based on sample data. The process is start with make the formulation null and alternative hypotheses, select an appropriate statistical test such as t-test, calculate a test statistic, and determine the statistical significance of the results [63]. In this study, the T-test results help compare the means of each feature between the two sampling methods (random sampling and stratified sampling). The t-statistic shows the magnitude and direction of the difference between the means, while the p-value tells us whether that difference is statistically significant. A p-value less than 0.05 usually indicates a significant difference. The results of the hypothesis testing can be seen in Table 6.

Variable such as house area, education level, area characterization, elevation, population density, waste generation, and plastic waste generation show statistically significant differences between the two sampling methods, meaning one method samples a different population characteristic than the other. Features like income, family members, and GDP do not show significant differences, meaning these features are relatively similar across the two methods. Stratified and random sampling methods lead to significant differences in many of the features, which suggests that one method may be better at capturing certain population characteristics.

Table 6. Hypothesis testing output for random and stratified sampling

Feature	T-Statistic	P-Value
House area	5.64	3.15 e-08*
Education level	-2.60	9.44 e-03*
Income	1.22	2.21 e-01
Family member	1.09	2.76 e-01
GDP	-1.43	1.52 e-01
Area Characterization	-9.06	5.69 e-18*
Elevation	10.89	2.27 e-24*
Population Density	-33.69	2.23 e-118*
Waste generation	2.53	1.1 e-02*
Plastic waste generation	12.20	2.71 e-29*

Status: * means significant difference

3.3. Machine Learning Model Performance with Linear Regression

The training and testing process uses a 70:30 split, where 70% of the data is allocated for training the models and the remaining 30% is used for testing. The models employed are Linear Regression and 2 (two) non-linear model that are Random Forest, and XGBoost, each fitted to the training data. Predictions are then made on the test set, and their performance is evaluated using Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE). These metrics provide insights into the models' accuracy, with lower values indicating better predictive performance. The results are subsequently compared to assess the effectiveness of different sampling methodologies in predicting the target variable. The result of the process as can be seen in Table 7.

Stratified sampling consistently yields better results compared to random sampling across all evaluated models. This is evident from the substantially lower error metrics, including Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE), achieved with stratified sampling. For example, Linear Regression under stratified sampling achieves an MSE of 0.13 and for random sampling the MSE is 0.23. Similarly, Random Forest's performance improves from an MSE of 0.04 with random sampling to 0.01 with stratified sampling. Even the highly robust XGBoost model shows enhanced precision with stratified sampling (RMSE=0.00) compared to random sampling (RMSE=0.0015). These results highlight that stratified sampling, by ensuring balanced representation of subgroups within the dataset, provides a more reliable foundation for accurate predictive modeling.

Table 7. Performance indicator for each model

Evalua-tion Para-meter	Sampling Method	Linar Regre ssion	Random Forest	XG Boost
MSE	Random	0.23	0.04	0.00
	Stratified	0.13	0.01	0.00
RMSE	Random	0.48	0.20	0.0015
	Stratified	0.36	0.10	0.00
MAE	Random	0.35	0.14	0.001
	Stratified	0.26	0.06	0.00

4. CONCLUSION

Stratified sampling is an effective method for obtaining a comprehensive understanding of a population, despite the relatively complex initial process. This complexity arises from the need to thoroughly understand the study area to establish appropriate strata classifications. By dividing the population into distinct subgroups, stratified sampling ensures that each stratum is adequately represented, thereby enhancing the overall accuracy and reliability of the results. Moreover, the use of remote sensing product data has been shown to significantly improve the accuracy and representativeness of waste generation sampling compared to random sampling method. Remote sensing provides detailed spatial information that allows researchers to identify and delineate different strata based on various environmental and socio-economic factors. This integration of remote sensing data not only facilitates a more precise sampling approach but also enhances the understanding of waste generation patterns within the study area. As a result, combining stratified sampling with remote sensing technology leads to more robust and insightful findings regarding waste generation. The results show that stratified random sampling, with class divisions based on remote sensing product data, yields more normalized data, with less error, and greater stability for machine learning analysis (for projection and classification needs). Despite these advantages, differences in data distribution compared to random sampling indicate that the limited number of classes may influence these results.

5. FURTHER RESEARCH

A difference in the range distribution (max and min value) between stratified and random sampling data may be caused due to the limited number of classes formed. Further research opportunity is to analyze the optimal number of classes needed to obtain data that is representative of the population.

6. ACKNOWLEDGEMENTS

The authors would like to thank Kemendikbudristek (Centre for Education Services - Puslapdik) and LPDP (Indonesia Endowment Funds for Education) for providing facilities during the research activities [grant number: 202101122512]. They would also like to thank Institut Teknologi Bandung for providing support during research activities.

7. LIST OF ABBREVIATIONS

List of abbreviations used in this study:

- MSE – Mean Squared Error
- VIIRS – Visible Infrared Imaging Radiometer Suite
- GDP/GRP – Gross Domestic Product / Gross Regional Product
- LULC – Land Use and Land Cover
- SNI – Standard Nasional Indonesia (National Standard of Indonesia)
- STD – Standard Deviation
- GDP – Gross Domestic Product
- IQR – Interquartile Range

8. REFERENCES

[1] Hasan S. E., Introduction to Waste Management: A Textbook, John Wiley and Sons, 2022, pp. 1-435.

[2] Ferronato N., and Torretta V., Waste Mismanagement in Developing Countries: A Review of Global Issues, International Journal of Environmental Research and Public Health, Vol. 16, Issue 6, 2019, pp. 1-28.

[3] Abdel-Shafy H. I., and Mansour M. S. M., Solid Waste Issue: Sources, Composition, Disposal, Recycling, and Valorization, Egyptian Journal of

Petroleum, Vol. 27, Issue 4, 2018, pp. 1275-1290.

[4] Wegedie K. T., Households Solid Waste Generation and Management Behavior in Case of Bahir Dar City, Amhara National Regional State, Ethiopia, Cogent Environmental Science, Vol. 4, Issue 1, 2018, pp. 1-18.

[5] Cruz M. B., Saikawa E., Hengstermann M., Ramirez A., McCracken J. P., and Thompson L. M., Plastic Waste Generation and Emissions from the Domestic Open Burning of Plastic Waste in Guatemala, Environmental Science: Atmospheres, Vol. 3, Issue 1, 2023, pp. 156-167.

[6] Pathak G., Nichter M., Hardon A., Moyer E., Latkar A., Simbaya J., Pakasi D., Taqueban E., and Love J., Plastic Pollution and the Open Burning of Plastic Wastes, Global Environmental Changes, Vol. 80, 2023, pp. 1-9.

[7] Wang H., Liu X., Wang N., Zhang K., Wang F., Zhang S., Wang R., Zheng P., and Matsushita M., Key Factors Influencing Public Awareness of Household Solid Waste Recycling in Urban Areas of China: A Case Study, Resources, Conservation and Recycling, Vol. 158, 2020, pp. 1-9.

[8] Trang P. T. T., Dong H. Q., Toan D. Q., Hanh N. T. X., and Thu N. T., The Effects of Socio-Economic Factors on Household Solid Waste Generation and Composition: A Case Study in Thu Dau Mot, Vietnam, Energy Procedia, Vol. 107, 2017, pp. 253-258.

[9] Suthar S., and Singh P., Household Solid Waste Generation and Composition in Different Family Size and Socio-Economic Groups: A Case Study, Sustainable Cities and Society., Vol. 14, 2015, pp. 56-63.

[10] Villalba L., Donalisio R. S., Basualdo N. E. C., and Noriega R. B., Household Solid Waste Characterization in Tandil (Argentina): Socioeconomic, Institutional, Temporal and Cultural Aspects Influencing Waste Quantity and Composition, Resources, Conservation and Recycling, Vol. 152, 2020, pp. 1-8.

[11] Gimenez A., Ares G., Chheang S. L., and Jaeger S. R., Reduction in Household Food Waste Through an Information Strategy: Findings from Consumers in Australia, Food Quality and Preferences, Vol. 111, 2023, pp. 1-14.

[12] Reyna-Bensusan N., Wilson D. C., and Smith S. R., Uncontrolled Burning of Solid Waste by Households in Mexico is a Significant Contributor to Climate Change in the Country, Environmental Research, Vol. 163, 2018, pp. 280-288.

[13] Govindan K., Zhuang Y., and Chen G., Analysis of Factors Influencing Residents' Waste Sorting Behavior: A Case Study of Shanghai, Journal of Cleaner Production, Vol. 349, 2022, pp. 1-12.

[14] Li W., Yan D., Li L., Wen Z., Liu M., Lu S., and Huang Q., Review of thermal Treatments for the Degradation of Dioxins in Municipal Solid Waste Incineration Fly Ash: Proposing a Suitable Method for Large-Scale Processing, Science of the Total Environment. Vol. 875, 2023, pp. 1-14.

[15] Taye A., Assefa E., and Simane B., Analysis of Practices and Factors of Solid Waste Management Among Urban Households of Addis Ababa City, Ethiopia, Environmental Challenges, Vol. 14, 2024, pp. 1-11.

[16] Sarker A., Baul T. K., Nath T. K., Karmakar S., and Paul A., Household Solid Waste Management in a Recently Established Municipality of Bangladesh: Prevailing Practices, Residents' Perceptions, Attitude and Awareness, World Development Sustainability, Vol. 4, 2024, pp. 1-12.

[17] Vassanadumrongdee S., and Kittipongvises S., Factors Influencing Source Separation Intention and Willingness to Pay for Improving Waste Management in Bangkok, Thailand, Sustainable Environment Research, Vol. 28, Issue 2, 2018, pp. 90-99.

[18] Liu Z., Yang J. Z., Bloomfield A., Clark S. S., and Shelly M. A., Predicting Recycling Intention in New York State: The Impact of Cognitive and Social Factors, Environmental Development, Vol. 43, 2022, pp. 1-11.

[19] Dhivyadeepa E., Sampling Techniques in Educational Research, Laxmi Book Publication, India, 2015, pp. 1-141.

[20] Grazhdani D., Assessing the Variables Affecting on the Rate of Solid Waste Generation and Recycling: An Empirical Analysis in Prespa Park, Waste Management, Vol. 48, 2016, pp. 3-13.

[21] Leknoi U., Yiengthaisong A., and Likitlersuang S., Social Factors Influencing Waste Separation Behaviour Among the Multi-Class Residents in a Megacity: A Survey Analysis from a Community in Bangkok, Thailand, Sustainable Futures, Vol. 7, 2024, pp. 1-11.

[22] Hu J., Miao L., Han J., Zhou W., and Qian X., Waste Separation Behavior with a New Plastic Category for the Plastic Resource Circulation: Survey in Kansai, Japan, Journal of Environmental Management, Vol. 349, 2024, pp. 1-10.

[23] Cheng X., Zhang L., Li W., and Long R., Intentions and Actions in Household Waste Separation: A Machine Learning Approach on the Gap and Determinants, Environmental Impact Assessment Review, Vol. 107, 2024, pp. 1-13.

[24] Zhang S., Luo Y., and Zhang P., A Comparative Study of Factors Influencing Residents' Waste Sorting Behavior in Urban and Rural Areas of China, Heliyon, Vol. 10, 2024, Issue 9.

[25] Etikan I., and Bala K., Sampling and Sampling Methods, Biometrics and Biostatistics International Journal, Vol. 5, Issue 6, 2017, pp. 215-217.
[26] Lazo D. P. L., and Gasparatos A., Factors Influencing Household-Level Positive and Negative Solid Waste Management Practices in Rapidly Urbanizing Cities: Insights from Santa Cruz De La Sierra, Bolivia, Environmental Research: Infrastructure and Sustainability, Vol. 2, Issue 1, 2022, pp. 1-18.
[27] Xu L., Ling M., Lu Y., and Shen M., External Influences on Forming Residents' Waste Separation Behaviour: Evidence from Households in Hangzhou, China, Habitat International, Vol. 63, 2017, pp. 21-33.
[28] Hannan M. A., Al Mamun M. A., Hussain A., Basri H., and Begum R. A., A Review on Technologies and Their Usage in Solid Waste Monitoring and Management Systems: Issues and Challenges, Waste Management, Vol. 43, 2015, pp. 509-523.
[29] Singh A., Remote Sensing and GIS Applications for Municipal Waste Management, Journal of Environmental Management, Vol. 243, 2019, pp. 22-29.
[30] Gautam S., Brema J., and Dhasarathan R., Spatio-Temporal Estimates of Solid Waste Disposal in an Urban City of India: A Remote Sensing and GIS Approach, Environmental Technology and Innovation, Vol. 18, 2020, pp. 1-12.
[31] Abbas Z., and Jaber H. S., Accuracy Assessment of Supervised Classification Methods for Extraction Land Use Maps Using Remote Sensing and GIS Techniques, in IOP Conference Series: Materials Science and Engineering, Vol. 745, Issue 1, 2020, pp. 1-14.
[32] Suif Z., Jelani J., and Ahmad N., GIS and Remote Sensing Integration for Soil Erosion Assessment Based on a Rusle Model in UPNM Catchment, Geomate Journal, Vol. 26, Issue 114, 2024, pp. 84-91.
[33] Van der Weide R., Blankespoor B., Elbers C., and Lanjouw P., How Accurate is a Poverty Map Based on Remote Sensing Data? An Application to Malawi, Policy Research Working Paper, 2022, pp. 1-61.
[34] Haripavan N., and Dey S., Application of Remote Sensing and Geographic Information System in Solid Waste Management for Gudivada Municipality, Andhra Pradesh, India, Waste Management Bulletin, Vol. 1, Issue 3, 2023, pp. 128-140.
[35] Fraternali P., Morandini L., and González S. L. H., Solid Waste Detection, Monitoring and Mapping in Remote Sensing Images: A Survey, Waste Management, Vol. 189, 2024, pp. 88-102.
[36] Sakti A. D., Rinasti A. N., Agustina E., Diastomo H., Muhammad F., Anna Z., and Wikantika K., Multi-Scenario Model of Plastic Waste Accumulation Potential in Indonesia Using Integrated Remote Sensing, Statistic and Socio-Demographic Data, ISPRS International Journal of Geo-Information, Vol. 10, Issue 7, 2021, pp. 1-23.
[37] Agustina E., Sembiring E., and Sakti A. D., Evaluation of Formal Waste Reduction Facility Location Compared to Recyclable Plastic Waste Generation in Denpasar City, Bali, Indonesia, in E3S Web of Conferences, Vol. 485, 2024, pp.1-14.
[38] Al-Salem S. M., Al-Nasser A., and Al-Dhafeeri A. T., Multi-Variable Regression Analysis for the Solid Waste Generation in the State of Kuwait, Process Safety and Environmental Protection, Vol. 119, 2018, pp. 172-180.
[39] Popli K., Park C., Han S. M., and Kim S., Prediction of Solid Waste Generation Rates in Urban Region of Laos Using Socio-Demographic and Economic Parameters with a Multi-Linear Regression Approach, Sustainability, Vol. 13, Issue 6, 2021, pp. 1-15.
[40] Wang K., Zhao X., Peng B., and Zeng Y., Spatio-Temporal Pattern and Driving Factors of Municipal Solid Waste Generation in China: New Evidence from Exploratory Spatial Data Analysis and Dynamic Spatial Models, Journal of Cleaner Production, Vol. 270, 2020, pp. 1-11.
[41] Fan L., Abbasi M., Salehi K., Band S. S., Chau K. W., and Mosavi A., Introducing an Evolutionary-Decomposition Model for Prediction of Municipal Solid Waste Flow: Application of Intrinsic Time-Scale Decomposition Algorithm, Engineering Applications of Computational Fluid Mechanics, Vol. 15, Issue 1, 2021, pp. 1159-1175.
[42] Hossain K. B., Chen K., Chen P., Wang C., and Cai M., Socioeconomic Relation with Plastic Consumption on 61 Countries Classified by Continent, Income Status and Coastal Regions, Bulletin of Environmental Contamination and Toxicology, Vol. 107, Issue 4, 2021, pp. 786-792.
[43] Xiao S., Dong H., Geng Y., Tian X., Liu C., and Li H., Policy Impacts on Municipal Solid Waste Management in Shanghai: A System Dynamics Model Analysis, Journal of Cleaner Production, Vol. 262, 2020, pp. 1-12.
[44] He R., Sandoval-Reyes M., Scott I., Semeano R., Ferrao P., Matthews S., and Small M. J., Global Knowledge Base for Municipal Solid Waste Management: Framework Development and Application in Waste Generation Prediction, Journal of Cleaner Production, Vol. 377, 2022, pp. 1-11.

[45] Zhang C., Dong H., Geng Y., Liang H., and Liu X., Machine Learning Based Prediction for China's Municipal Solid Waste under the Shared Socioeconomic Pathways, Journal of Environmental Management, Vol. 312, 2022, pp. 1-13.
[46] Younes M. K., Nopiah Z. M., Basri N. A., Basri H., Abushammala M. F., and Younes M. Y., Landfill Area Estimation Based on Integrated Waste Disposal Options and Solid Waste Forecasting Using Modified ANFIS Model, Waste Management, Vol. 55, 2016, pp. 3-11.
[47] Wu F., Niu D., Dai S., and Wu B., New Insights into Regional Differences of the Predictions of Municipal Solid Waste Generation Rates Using Artificial Neural Networks, Waste Management, Vol. 107, 2020, pp. 182-190.
[48] Anilkumar P. P., and Chithra K., Land Use Based Modelling of Solid Waste Generation for Sustainable Residential Development in Small/Medium Scale Urban Areas, Procedia Environmental Sciences, Vol. 35, 2016, pp. 229-237.
[49] Van Emmerik T., Tramoy R., Van Calcar C., Alligant S., Treilles R., Tassin B., and Gasperi J., Seine Plastic Debris Transport Tenfolded During Increased River Discharge, Frontiers in Marine Science, Vol. 642, 2019, pp. 1-7.
[50] Widyarsana I. M. W., Damanhuri E., and Agustina E., Municipal Solid Waste Material Flow in Bali Province, Indonesia, Journal of Material Cycles and Waste Management, Vol. 22, Issue 2, 2020, pp. 405-415.
[51] Ghayebzadeh M., Taghipour H., and Aslani H., Estimation of Plastic Waste Inputs from Land into the Persian Gulf and the Gulf of Oman: An Environmental Disaster, Scientific and Social Concerns, Science of the Total Environment, Vol. 733, 2020, pp. 1-10.
[52] Ghayebzadeh M., Aslani H., Taghipour H., and Mousavi S., Estimation of Plastic Waste Inputs from Land into the Caspian Sea: A Significant Unseen Marine Pollution, Marine Pollution Bulletin, Vol. 151, 2020, pp. 1-8.
[53] Law K. L., Starr N., Siegler T. R., Jambeck J. R., Mallos N. J., and Leonard G. H., The United States' Contribution of Plastic Waste to Land and Ocean, Science Advances, Vol. 6, Issue 44, 2020, pp.1-7.
[54] Liu B., Zhang L., and Wang Q., Demand Gap Analysis of Municipal Solid Waste Landfill in Beijing: Based on the Municipal Solid Waste Generation, Waste Management, Vol. 134, 2021, pp. 42-51.
[55] He R., Sandoval-Reyes M., Scott I., Semeano R., Ferrao P., Matthews S., and Small M. J., Global Knowledge Base for Municipal Solid Waste Management: Framework Development and Application in Waste Generation Prediction, Journal of Cleaner Production, 2022, pp. 1-11.
[56] Intharathirat R., Salam P. A., Kumar S., and Untong A., Forecasting of Municipal Solid Waste Quantity in a Developing Country Using Multivariate Grey Models, Waste Management, Vol. 39, 2015, pp. 3-14.
[57] Elvidge C. D., Baugh K., Zhizhin M., Hsu F. C., and Ghosh T., VIIRS Night-Time Lights, International Journal of Remote Sensing, Vol. 38, Issue 21, 2017, pp. 5860-5879.
[58] Gaughan A. E., Stevens F. R., Linard C., Jia P., and Tatem A. J., High Resolution Population Distribution Maps for Southeast Asia in 2010 And 2015, PLOS One, Vol. 8, Issue 2, 2013, pp. 1-11.
[59] Gong P., Li X., Wang J., Bai Y., Chen B., Hu T., Liu X., Xu B., Yang J., Zhang W., and Zhou Y., Annual Maps of Global Artificial Impervious Area (GAIA) between 1985 and 2018, Remote Sensing of Environment, Vol. 236, 2020, pp. 1-12.
[60] Black K., Business Statistics: For Contemporary Decision Making, John Wiley and Sons, 2023, pp. 1-809.
[61] Shmueli G., Bruce P. C., and Patel N. R., Data Mining for Business Analytics: Concepts, Techniques, and Applications with XLMiner. John Wiley and Sons, 2016, pp. 3-527.
[62] Chang C. J., Li D. C., Huang Y. H., and Chen C. C., A Novel Gray Forecasting Model Based on the Box Plot for Small Manufacturing Data Sets, Applied Mathematics and Computation, Vol. 265, 2015, pp. 400-408.
[63] Alshahrani T., Jethave G., Nemade A., Khairnar Y., Fegade U., Khachane M., Al-Ahmed A., and Khan F., Application of ANN, Hypothesis Testing and Statistics to the Adsorptive Removal of Toxic Dye by Nanocomposite, Chemometrics and Intelligent Laboratory Systems, Vol. 249, 2024, pp. 1-10.