2020 Asia–Pacific Statistics Week
A decade of action for the 2030 Agenda: Statistics that leaves no one and nowhere behind
15-19 JUNE 2020 | Bangkok, Thailand

A Pilot Project of Area Sampling Frame for Maize Statistics: Indonesia's Experience

Octavia Rizky Prasteyo*[1]; Kadir[2]; Ratna Rizki Amalia[3]
[123] BPS-Statistics Indonesia
Jl. Dr. Sutomo 6-8, Jakarta, Indonesia
Email: [1]octavia.rizky@bps.go.id, [2]kadirsst@bps.go.id, [3]ratna.amalia@bps.go.id

Abstract:
A critical issue in the context of food policy in Indonesia is the accuracy of crops statistics, particularly rice and maize. In 2018, Statistics Indonesia (BPS), in collaboration with the Agency for Assessment and Application of Technology (BPPT), has successfully implemented the area sampling frame (ASF) method to improve the accuracy of paddy harvested area estimation. The result was released in October 2018 and received appreciation from many parties. The success story of the ASF implementation on rice statistics has encouraged BPS to replicate the same technique to estimate the harvested area of maize that suffers from the issue of overestimation for many years. The results of the 2013 Agricultural Census and the 2018 Inter-Censal Agricultural Survey, for instance, strongly indicate that the maize harvested area statistics obtained using the conventional method, which is so-called eye estimate, was overestimated. For that reason, in April 2019, BPS began to conduct a pilot project on the implementation of the ASF for maize. One of the hardest challenges in replicating the ASF method for maize is the frame construction due to the insufficient spatial information about the land specifically dedicated to maize cultivation. To address this challenge, BPS constructed the frame by making use different sources of spatial information, namely the results of the area sampling frame for paddy (paddy field that is not planted by paddy), field identification and information from the Ministry of Agriculture. This paper provides a comprehensive look at the whole process of the development of ASF for maize statistics. The discussion in this paper covers two main issues, namely the methodology applied and the business process of data collection and processing.

Keywords: Area Sampling Frame; methodology; objective measurement

1. Introduction

A critical issue in the context of food policy in Indonesia is the accuracy of crops statistics, particularly rice and maize. In 2018, Statistics Indonesia (BPS), in collaboration with the Agency for Assessment and Application of Technology, has successfully implemented the area sampling frame (ASF) method to improve the accuracy of paddy harvested area estimation. The development of ASF for paddy was presented by Muhlis (2018) in the Asia-Pacific Economic Statistics Week (APES 2018).

The result was released in October 2018 and received appreciation from many parties. The success story of the ASF implementation on rice statistics has encouraged BPS to replicate the same technique to estimate the harvested area of maize that suffers from the issue of overestimation for many years. Indeed, the results of the 2013 Agricultural Census and the 2018 Inter-Censal Agricultural Survey, for instance, strongly indicate that the maize harvested area statistics obtained using the conventional method, which is so-called eye estimate, was overestimated. The harvested area data, which is the main component in the maize production computation besides the yield, has created a prolonged public debate around its accuracy as a policy input, particularly the importation of maize for price stabilization. There are many criticisms from many parties such as academia, economic observers and politicians regarding this issue.

The need for an objective measurement to address the accuracy issue as successfully implemented for paddy has motivated BPS to conduct a pilot project on the implementation of the ASF for maize. The project has been implemented since April 2019. This paper presents a detail description about the project. It covers the methodology applied, the business process of data collection and processing, current progress and preliminary results and some challenges faced as the subject of innovation and improvement in the future.

2. Methodology

A. Construction of Area Sampling Frame for Maize

The construction of ASF starts from collecting spatial information on the distribution of the potential location of maize cultivation as a frame.

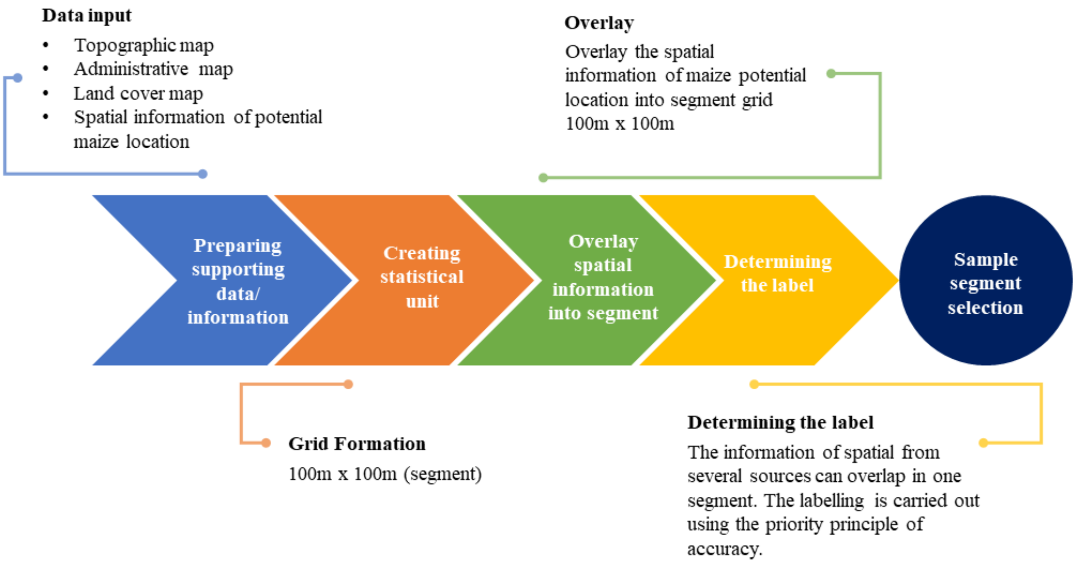


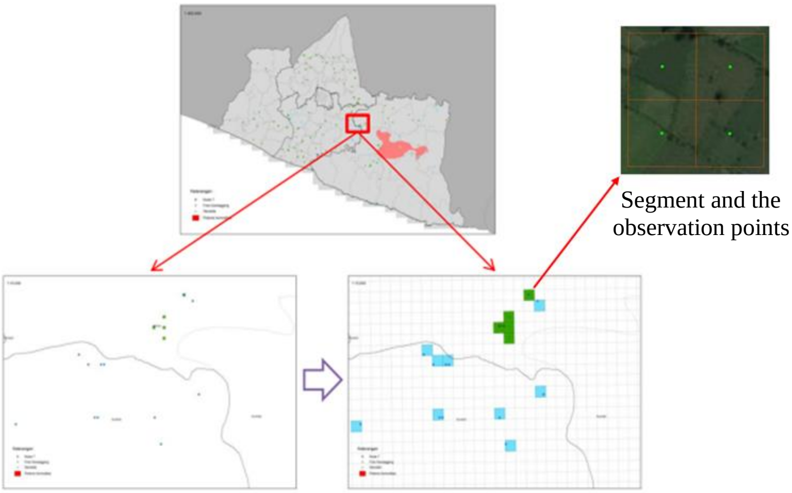
Figure 1. Construction of Area Sampling Frame for Maize

The frame construction for maize is very challenging due to insufficient spatial information about the land specifically dedicated to maize cultivation. This condition was reasonably different with the construction of area sampling frame for paddy because of the availability of information on paddy field area obtained from satellite imagery, which is collected and officially released by Ministry of Land and Spatial Planning/National Land Agency. To address this challenge, BPS in collaboration with the Agency for the Assessment and Implementation of Technology (BPPT) made use different sources of spatial information to construct the frame. There are four information sources of potential maize cultivation locations exploited in the frame construction as follows:

Source 1: Coordinate points of maize cultivation obtained from the implementation results of the area sampling frame for paddy.
Source 2: Field identification. The surveyors took geotagged photographs of maize cultivation location to get the coordinates of land planted by maize and sent them to the headquarters office.
Source 3: Coordinates of maize cultivation location from the Directorate of Cereals, Ministry of Agriculture (MoA).
Source 4: Polygons of potential areas of agricultural commodities, including maize, obtained from the Agricultural Data and Information Center, MoA.

The second step is the formation of a statistical unit so-called a square grid on the area of study. The area of study boundary used in the construction of ASF for maize is the same as implemented in the construction ASF for paddy. Following Gallego (2015), the units of an area frame can be points, transects (lines of a certain length), or pieces of territory. In this project, the statistical unit is a square sized 100m x 100m, called a segment. The segment size for maize is smaller than that of for paddy as maize cultivation in Indonesia mostly conducted on small pieces of land and is not as massive as paddy cultivation. Segment boundaries are determined based on geographical coordinates with a fixed location.

In the final stage, all potential location information is overlaid into the grid segment. If there is overlapping information, the segment labelling is conducted by selecting the smallest code from these four possible codes: (1) results of paddy area sampling frame; (2) coordinates of geotagged photos; (3) information from the Directorate of Cereals, MoA; (4) information from the Agricultural Data and Information Center, MoA. Figure 2 illustrates the overlay process of maize potential location coordinates and segment grid of 100m x 100m.



Source: BPS-Statistics Indonesia, 2019

Figure 2. Illustration of Overlay Process of Maize Locations Coordinates Information into Segment

In order to obtain the representation of observation points in each statistical unit, each segment is divided into a 50m x 50m square grid so-called a subsegment. Each centre point of a subsegment then regularly observed for the maize growth phase. In total, for each segment, four observation points are representing the characteristics of the segment.

B. Sample Segment Selection

The determination of the sample size, which is the number of sample segments, takes into consideration the accuracy of area estimation at the sub-district level. The computation of the segment population makes use of a particular approach, which is the maximum maize planting area in 2017 obtained from the previous administrative report. The use of this information is due to the unavailability of the maize field area as the paddy has. The maximum planting area then is divided by the segment size of 100m x 100m (1 Hectare) to obtain the number of segment population. The equation is as follow:

$$N_k = roundup\left(\frac{Planting\ area\ (Ha)}{1\ Ha}\right)$$

In this project, the number of segment samples for each sub-district is 1 per cent of the segment population. The selection of the segment samples applies a simple random sampling technique with a distance threshold of 250 meters. It means that a selected segment will be removed if it is drawn with a distance of fewer than 250 meters to any of the previously selected segment. Following Gallego (1995), the distance threshold approach aims to avoid the concentration of the segment samples in some areas.

For the sake of data management, a unique identity was assigned for each selected segment as implemented in ASF for paddy (Muhlis, 2018), which consists of nine digits. The first two digits are the provincial code followed by two digits of municipality code, and the next three digits are the sub-district code, and the last two digits are the serial number of a segment in a sub-district.

C. Implementation

In conducting the field observation, the enumerators use their smartphone equipped with an Android-based application, which is specially developed for this project. In practice, the enumerators will observe and identify the growth phase of maize. Besides, enumerators have to take a picture of the centre point of all subsegments in each selected segment. The aim is to reduce the moral hazard and support in the monitoring process. Direct observation at the observation points will provide objective and unbiased information (Mubekti & Sumargana, 2016). The collected information on the growth phase of maize includes early and final vegetative, early and final generative, maize harvest for livestock, harvest in the form of sweet and baby maize, as well as standard maize harvest (dry shelled maize). Besides, other information collected covers land preparation, land for other crops, non-agricultural field, and damaged.

Enumerators need to know the physical boundaries of the sample segments in the field according to selected segments coordinates. Following Wigton & Bormann (1978), boundaries which can easily be located by the enumerator and supervisor is the key to quality enumeration. Physical boundaries in the field can be determined using features in the application. Figure 3 shows the illustration of field observation and the application display. The blue dot on the first display represents the position of the enumerator while the red dots show the point observation, and the green rectangle is the boundary of the segment, which is divided into four subsegments. To make the blue dot function work properly, the enumerators must activate the Global Positioning System (GPS) in their device before visiting a selected segment.

Source: BPS-Statistics Indonesia, 2019

Figure 3. Illustration of Field Observation and the Application Display

After enumerators have completed data recording in each subsegment within a selected segment, then they must send the results to the server through the application. The results consist of the information of maize growing phase and the images of vegetation at the centre point of each subsegment taken by the enumerators. However, if internet access is not available in the sample location, they can continue to observe other selected segments and send the results all together after having internet access as long as it is conducted during the observation period.

The observation is carried out during the last seven days every month. Then, in the first five days of the following month, the observation results will be validated by the supervisors based on an image of maize growing phase sent by the enumerator. The survey can be considered as a panel study since the same segment samples are going to be observed every month. The final target output is the estimation of the maize field area according to the growth phases. It is also expected that the observation results for a certain month could provide predictions of the maize harvested area for the next four months.

One of the main advantages of the ASF implementation on maize data collection is the exploitation of so-called cost-effective technology and tool for data collection as recommended by the Food and Agriculture Organization (FAO) in the Global Strategy to improve Agricultural and Rural Statistics (GSARS). The implementation could be considered as the use of Bring Your Own Device (BYOD) system in data collection since the enumerators only use their Android-based mobile phones to collect data sent them to the server. This innovation allows the business process of data collection becomes more efficient than before. Moreover, the availability of the data is faster than that of the previous method. The last one took about as much as one month to provide useful information due to reliance on the paper-based data collection process. Furthermore, the use of the geospatial information technology (GIT) provides spatial information related to the maize cultivation area so that the policy can be more targeted when it comes to the spatial context. This spatial information can be used by the policymakers to monitor the dynamic of maize cultivation regularly.

3. Current progress and preliminary results

This project has been carried out in all provinces in Indonesia since April 2019, except in DKI Jakarta Province. In 2019, the total target of samples was 21,965 segments observed every month at the fixed locations. In so far, the realisation of the project is delightful. As shown in Figure 4, above 99 per cent of the total samples have been observed every month since June 2019.

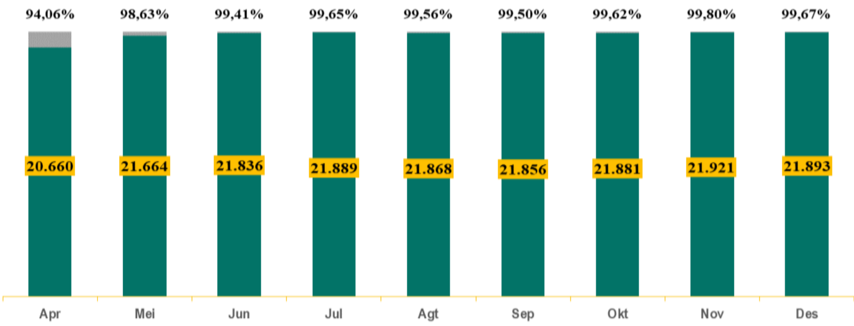


Figure 4. The Realisation of a Segment Sample in 2019

As mentioned before, the implementation of ASF on maize data collection can provide spatial information on maize cultivation all over Indonesia. Figure 5 shows an example of the spatial information produced based on the observation result in February 2020 in West Nusa Tenggara Province. The map provides us with some useful information regarding the spatial distribution of maize cultivation in the province.

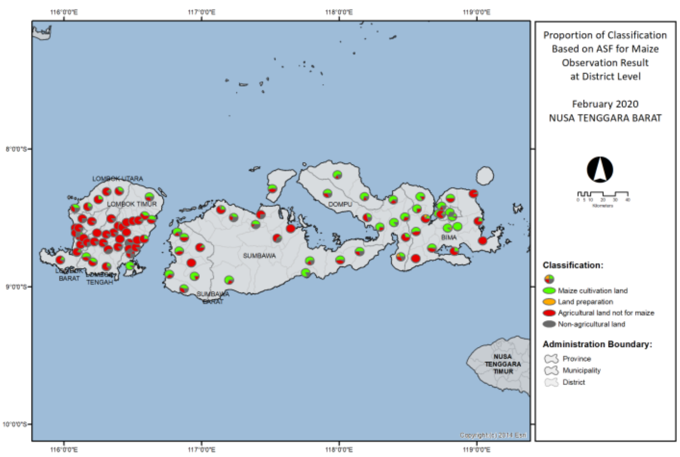


Figure 5. Example of Spatial Information Provided for West Nusa Tenggara Province, February 2020

It can be seen from the map that most of the eastern region of the province was cultivating maize in the observation month period. For the sake of simplification of the presentation, the map only provides the spatial distribution of maize cultivation. More detail information consisting of the growing phase of maize also can be provided.

4. Discussion, Conclusion and Recommendations:

The critical factor for the successful implementation of the ASF for maize is a well-built collaboration between BPS and other government institutions, especially those with expertise in the field of geospatial, namely Geospatial Information Agency (BIG), Space Agency (Lapan), Ministry of Agrarian Affairs and Spatial Planning/National Land Agency (ATR/BPN) and BPPT. In that regard, a useful lesson learnt from the implementation is a strong collaboration between many parties, particularly line ministries and the National Statistical Office (NSO) is needed to ensure that a significant initiative with massive coverage such as this project can be implemented successfully.

The implementation of ASF for maize has changed the business process of data collection dramatically due to the use of both an objective measurement and cost-effective technology, particularly the BYOD system and GIT. More abundant information on maize cultivation, specifically in the form of spatial information, could also be provided regularly faster than before to policymakers. It allows them to be more well-informed in formulating effective policies on maize commodities in Indonesia.

However, there are still several challenges in the development of ASF for maize that would be the subject of innovation and improvement in the future. Some of them can be summarized as follows:

1. The availability of spatial information on maize cultivation land area due to the characteristic of maize cultivation in Indonesia, which has been a significant challenge since the beginning. For the time being, BPS is still conducting studies related to the computation of potential maize area, which is the area of land that dedicated explicitly to maize cultivation at least once a year. One of the measures in progress is the use of remote sensing technology for maize cultivation land mapping. Information on maize land area is necessary to produce an accurate estimation of maize harvested area.
2. The second challenge is the possibility of misidentification of maize growth phase by the enumerator when doing the field observation. To address the issue, BPS is continuing to improve the business process by upgrading the web-based monitoring system as well as creating an Android-based monitoring system. In the near future, BPS also has a plan to develop a machine learning model to identify the maize growth phase automatically.
3. Another challenge is the geographical difficulties in some parts of each region. Therefore, sometimes enumerators require more effort in reaching observation points in some areas, especially in mountainous areas. Moreover, the coverage of the implementation is quite massive since Indonesia is a vast archipelago. As a result, a massive number of enumerators is deployed for field observation.

5. References:

1. BPS-Statistics Indonesia. (2019). *Pedoman Teknis Pendataan Statistik Pertanian Tanaman Pangan Terintegrasi dengan Metode Kerangka Sampel Area untuk Komoditas Jagung 2019*. Jakarta: BPS.
2. FAO, World Bank, and UNSC. (2012). Action Plan of The Global Strategy to Improve Agricultural and Rural Statistics. Rome.
3. Gallego, F. J. (1995). Sampling frames of square segments. Luxembourg: European Commission.
4. Gallego, F. J. (2015). Area Sampling Frames for Agricultural and Environmental Statistics. EUR 27595 EN.
5. Mubekti, M., & Sumargana, L. (2016). *Pendekatan kerangka sampel area untuk estimasi dan peramalan produksi padi*. Jurnal Pangan, *25*(2), 71-146.
6. Muhlis. (2018). Area Sampling Frame: A New Approach to Reform Agricultural Data Collection. Asia-Pacific Economic Statistics Week.
7. Wigton, W. H., & Bormann, P. (1978). A guide to area sampling frame construction utilizing satellite imagery. Rome: *Submitted at the Second International Training Course in Remote Sensing Applications for Agriculture: Crop Statistics and Agricultural Census.*