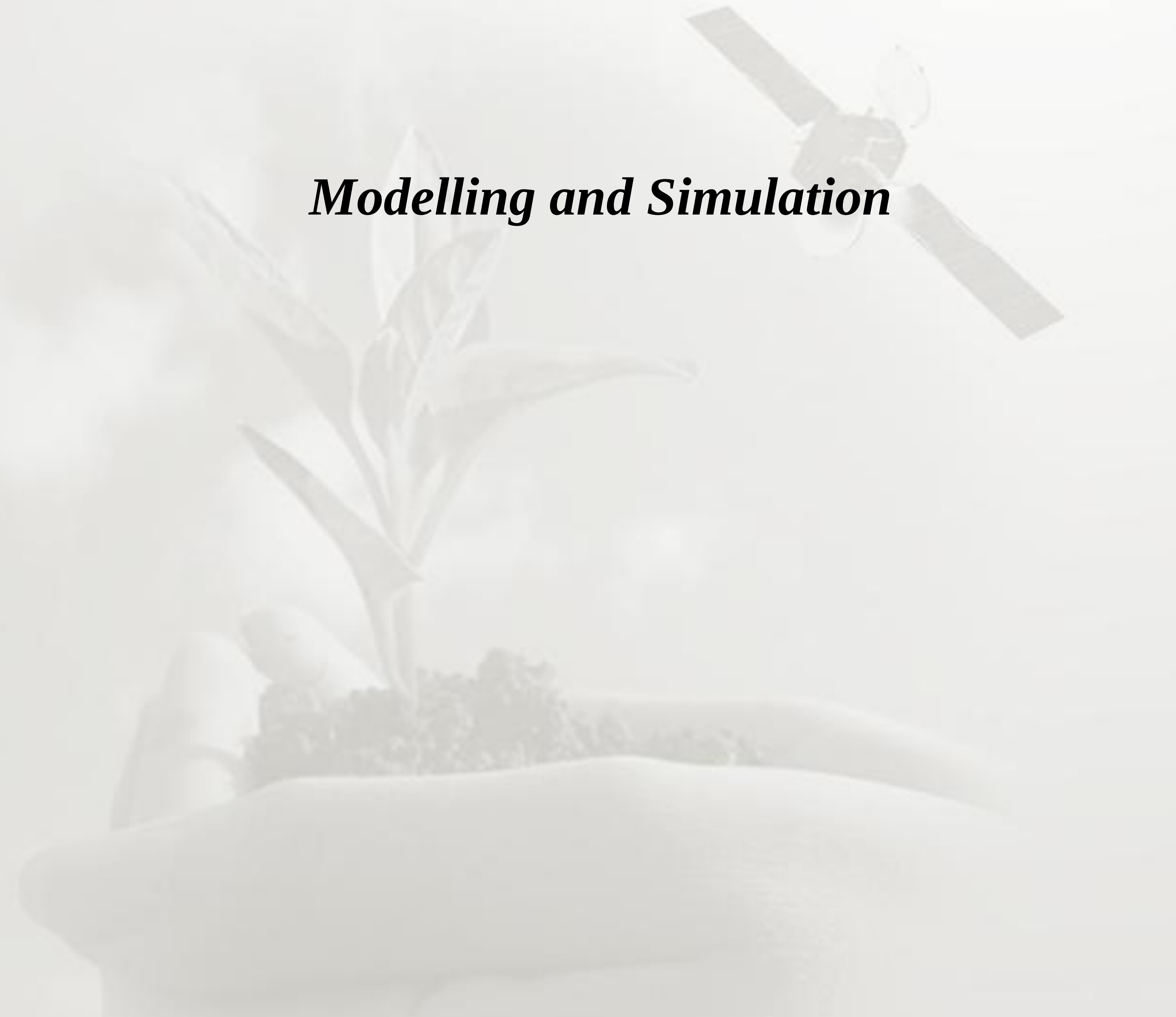
Modelling and Simulation



Sampling of Square Segments by Points for Rice Production Estimate and Forecast

Mubekti
Pusat TISDA, BPPT Jl. M.H. Thamrin No. 8 Jakarta
e-mail: mubekti@webmail.bppt.go.id, Ph: 021-3168914

Gatot Hendrarto
Pusat TISDA, BPPT Jl. M.H. Thamrin No. 8 Jakarta
e-mail: gatoth@webmail.bppt.go.id , Ph: 021-3168913

Abstract-- Rice is the staple food of most Indonesian peoples. This commodity plays an important role in the economy of Indonesia. Because of this the government of Indonesia should be careful in drawing a certain policy, especially the policy which is related to the context of rice crop development planning. To support this rice crop development planning, an accurate data i.e. a rice production estimate is needed. Statistical services on agriculture are based on different approaches, resulting from both historical reason and available techniques, to provide reliable figures. Most often, the applied systems are based on: village statistics, census, sampling survey based on list frame or area frame, and administrative by products. The Area Frame Sampling is one of the methodologies to obtain the rice crop area estimation as agricultural statistic data. In this method/system, crop area estimates are generated through area frame sampling data. All the components are integrated into the system to support timely generation of reliable information on the crop area of the major commodity in the study area. In the present paper, we mainly try to review the approach of area frame sampling for rice estimate that has been developed in Indonesia for several years. The discussion would be focused on area frame construction, field survey, data communication and some results.

Keywords: Estimation, forecast, rice production, area frame sampling, GIS, remote sensing.

Introduction

Agricultural sector is renewable resources which has an important role in the national economics structure, especially in developing countries like Indonesia. At present, food commodity especially rice still has an important role regarding to daily life of the people and also as a part of economic activity which can accommodate a lot of labors to employ on it. Rice is a staple food for Indonesian people and frequently called as a political commodity due to its vital role, it is to say that unstable of the national rice stock would be followed by unstable of the national politics condition. Improvement of the rice production has been conducted by government through various efforts in order to increase not only harvested area but yield as well. Becouse of the rapid growing cycle of rice dynamic, the accurate and timeline information are needed as an input for decision making to define appropriate management from farm level up to national level.

At the present, the data of rice area is collected manually from the sub-district level by using SP-IA list and assimilated up to the national level. The accuracy of the derived data has not satisfied yet and contain errors due to the subjective factor and many iterations of combining statistics. For that reason, the methodology development is needed in order to get more accurate and objective figure of rice statistics. One of the methods in deriving agricultural statistics is sampling technique in which a partial observation of agriculture is conducted for extrapolating the whole population.

Area frame sampling of square segment is a statistical procedure to measure the quantity of land use area in a geographic region of interest by observing square segments as a small part of the population. The procedure was adopted in technical cooperation between EU, BPPT, MoA from 1998 to 2000 through SARI Project for rice statistic approach (BPPT-EU, 2001), then the improvement of survey method and data communication from farm level to central level is still continued. In the present paper discuses the principle of area frame sampling and its implantation.

I. SOME APPROACHES TO AGRICULTURAL STATISTICS

Different methods are applied on agricultural statistics, resulting from both historical reason and available techniques, to provide reliable figures of data. Meyer-Roux (in Gallego, 1995) explains that, most often, the applied system are based on: village statistics, census, sampling surveys based on list frame or area frame and administrative by-products. But recently, these likely be changed or modified in forthcoming years by the impact of rapid development of new technologies, such us computer, Geographic Information System, Remote Sensing, and the availability of digital maps.

2.1. Village Statistics

The basic unit is an administrative region (e.g. village, sub-district), and the task of data collection is usually performed by local administrators or agricultural officers (Mantri Tani). The data collectors give subjective estimates of area and possibly of production per crop type refer to data collected before. Data processing is relatively simple, since no sampling is involved and the number of units is small in comparison to the number of farmers or the number of agricultural fields. The whole national territory can be covered, if a stable structure of local correspondents has been set up. Indonesia employs this system to collect agricultural data from sub-district cumulative up to national level. Village statistics are useful as an alternative, when it is difficult to collect data from farmer or from other means, such us direct ground observation. But, estimating agricultural magnitudes by the village statistics approach gives results of rather poor quality. The main problem is that often the bias due to subjectiveness is systematically in the same direction.

2.2. Agricultural Census

The census of agriculture is usually a Government-sponsored large scale operation for the collection and derivation of quantitative statistical information on the current structure of agriculture, using household as the unit of enumeration. Producing an exhaustive list of farms in a country or region is an obvious way to compute crop area, yield, or other information. Most countries carry out this operation with an interval of 5 to 10 years, whereas, Indonesia has four types of 10-year census, i.e. population census (SP), Economical census (SE), Village Potential census (Podes), and Agricultural census (ST). The last agricultural census of Indonesia was conducted in 200, and the next census would be in 2013. The general objective of the census is to provide up-to-date and reliable statistics on agricultural structure, and it will also help to update sampling frame as a reference for future agriculture surveys.

2.3. Sampling Survey

Since a census is made every 5 or 10 years, the results obtained directly are clearly not valid for the whole period between censuses. Yearly or shorter periods surveys can be carried out by a partial observation of the agricultural sector. Sampling theory provides tools to make an objective extrapolation to the whole population by just observing a small part (sample) of the population. The common sampling frames used on agricultural statistics are list frame, area frame and the combination of both (multiple) frames (Cochran, 1977, and Carfagna et.al, 1994). A list frame is derived from a list of farms from which a sample is drawn. One main limitation of this type of survey is that the sampling frame is usually not updated at the time of the sample. Some farms or household of farms have disappeared, other farms have been created, split or merged, and many has changed size or specialization. Area frame sampling (AFS) is a statistical procedure to measure the quantity of materials of interest in a geographic region of interest. The "area frame" bounds the region of interest, or study area. It may be specified by an arbitrary boundary (like a square or rectangle), a natural boundary (like an ecosystem), or an administrative unit (like a district). The main advantage of area frame compared to list frame is that the sampling frame is mostly under control. The only possible changes may regard the size of units or stratification, that can be modified by the manager. Farms being created, split, modified will not change the set of segments into which a geographic area have been divided.

II. AREA FRAME SAMPLING

Area frames provide a different approach to agricultural statistics. The units of an area frame are directly bound to a geographical area. If the limits of the region are known, then the elements of the population will be exactly known as well. These elements can be of two main types: points or pieces of land, often known as 'segment'. This approach is primarily use for crop area estimation, although yield or production estimates can be obtained if the surveyors are able to give yield evaluation by using a crop cutting measurement or a subjective estimate (Gallego, 1995, Houseman, 1975). The concepts of AF sampling are simple: i.e. divide the total land area to be surveyed into N small parcels of land, without overlap or omission; select a random sample of n parcels". There are two main types of area frame sampling, namely area frames of physical boundary (cadastral segments) and area frame of non-physical boundary (geometrical segments). Next discussion in this paper will be mainly focused on area frame on non-physical boundary.

3.1. Area Frame of Segments

Area frame of sample segment surveys are designed for specific purposes and are commonly used in agricultural statistics as known by FAO (Cotter & Nealon, 1987). The approach of segment sampling consists of dividing the region into grid of 10 km x 10 km, then pieces of land or segments with regular or irregular shapes are defined. The size of the segments varies, but it size used in Indonesian pilot are 100 ha and 25 ha.

The size of frame units (segments) depend on the average field size and homogeneity of the land use. Trials carried out up to now suggest that, for crop estimation, smaller segments can be more convenient in difficult, mountainous area or when the size of the fields is small. Previous pilot study of area frame in Cianjur District by using segment size of 100 ha showed that, the delineation of agricultural fields in one segment need more than 8 hours due to the complexity of the land use. Then, the size of the segments was modified to be 25 ha to make more efficient time and cost.

3.2. Area Frames of Points

Point samples are often employed in environmental, forestry or mining studies (Ripley, 1981; Cressie, 1993). In theory, a point frame is the set of dimensionless points in a certain region and, therefore, infinite. In practice, points can be given a dimension taking into account the precision of the graphic material used to locate the point. If aerial photographs with a scale 1:5000 are used, a point may be given a dimension of 3 m x 3 m, since 3 m correspond to 0,6 mm in the photographs, which is a reasonable choice to make two points distinguishable. The discussion below is mainly in relation with area frame of points for rice production estimate.

III. AREA FRAME OF SQUARE SEGMENT BY POINTS

The technique of area frame by points which has been operational application is the French TER-UTI survey (Porchier, 1990). The construction of area frame of square segment by points is essentially the same as normal area frame, except that only a set sample of points inside a segment is visited instead of the whole segment and no fields must be delineated. Based on the observation of sample points, area estimates are computed and used as a valid generalisation without studying the entire area under investigation. The discussion in this paper would be focused on area frame construction, field survey, data communication and some results of pilot study in Indramayu District, West Java..

4.1. Area Frame Construction

In the area frame construction not only Indramayu District was built but also other districts of West Java Province were built in the coincidence works. Then the administrative boundary was used to cut-off study area of Indramayu District. Land use map was used to delimitate and stratify the study area.

Stratification aims to divide a population (Ω) of size N into H non-overlaping sub-population (Ω_h-strata) of size N_h *in order to get more eficience in both level of accuration and cost* (Taylor, at al, 1997). Considering the classes of present land use on the maps and with the intention to obtain only two strata for each District - a bigger number of strata would in fact give us a too big sample dimension, the study area is divided into 3 katagories (Mubekti, et. al., 2010)):

- Strata-0 or "no-rice" is the polygons with no rice field content, such as forest, settlement, water body etc.
- Strata-1 or "Rice" is the polygons which have high expected percentage of rice field content, i.e. irrigated rice field and rainfed rice field
- Strata-2 or "possibly Rice" is the polygons which have low expected percentage of rice content, i.e. up-land arable.

Figure 1. shows the spatial stratification in the study area.

The result of the stratification shows that the total area of Indramayu district is 207,675 hectares consisted of 142,950 hectares strata-1, 2,800 hectares strata-2, and 61,925 hectares strata-0. The area of sampling frame is the addition between strata-1 and strata-2, while strata-0 is not included from sampling frame.

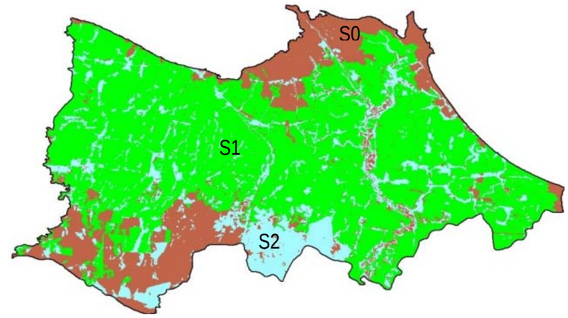


Figure 1. Stratification map of study area

Since the area frame refers to square segment, then a fishnet of elementary unit sized 500 m x 500 m is set by converting the stratification map so as to make easier for sample sampling (Figure 2.). The result of the set fishnet shows that the total of the area frame is 5830 cells consisting of 5718 cells strata-1 and 112 cells strata-2.

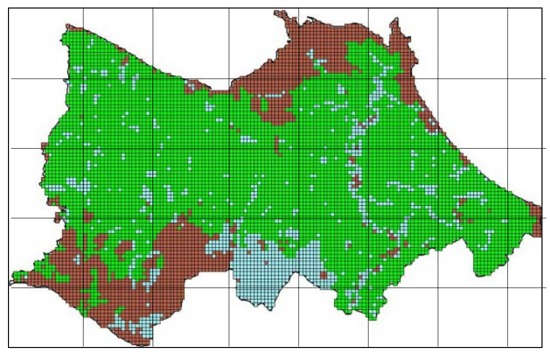


Figure 2. Grid cell map of study area

In the statistical inference estimation process the sample dimension is always a crucial point. The statistician has to find the best solution to fit the cost of the survey with acceptable data estimation accuracy. In that cases are also to be considered the difficulties arising from the operative point of view since it is certainly a huge task to make the process starting from planning, training, field survey, up to data analysis. Due to these operative problems the sample dimension was set considering the minimum acceptable dimension. In this study the dimension of sample segments is approximately 1 % of the area frame. Because of very small sample segment in strata-2, it is decided to use only one stratum where strata-1 and strata-2 are combined all together.

After setting the sample dimension it was analyzed the methodology to randomly extract the elementary units. In spatial statistics it is very important to manage a sample geographically distributed and several approaches have been developed. The selected method is known as "Systematic Aligned Sampling with a distance threshold" (Gallego, 1995). The GIS Arc-View software was employed to extract sample segments.

This method involve the division of area frame into blocks of segments and the selection of a fixed number of segments in each block repeating the same “replicate”. The replicate is a fixed pattern or location of selected segments in one block. The same selected segment positions in each block are used.

The used block was 10 x 10 Km wide containing 400 segments of 500 m x 500 m and in each block 10 segments were drawn. Only the segments that exceed a distance threshold between each other were kept and sequentially numerated. In this extraction 3 segments were deleted and 7 segments were obtained for each block. The general rule is to use the first 4 positions to obtain the desired sample dimension. If it is need modification the location number 5, number 6 can be added and so on. Figure 3. illustrates the distribution of extracted sample segments on land use map. Total number of extracted sample (n) are 52 segments spread over 29 sub-districts out of the total 31 sub-districts available in Indramayu district. Each extracted sample segment is attributed with code number and geographical co-ordinate at bottom-left corner.

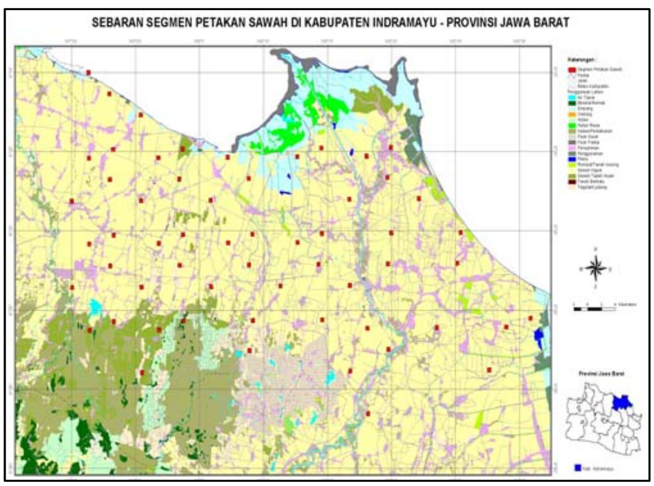


Figure 3. Distribution map of sampel segment

Next step is to locate sample points which 25 sample points will be allocated inside each segment. The implementation is to make grid cell of 100 m x 100 m size on the extracted sample segments. The center of each grid cell is to be sample point for field observation. The surveyor will only write down the land use in each of the 25 points instead of drawing and further digitizing all the field in the segment. For the whole of Indramayu district there are 52 (segments) x 25 (points) = 1300 sample points.

4.2. Field Survey

In the study, the limit of the segments is not physical feature but based on geographical boundary, so that, it need supporting material to identify location of sample segments and observation points when the field survey done. Land use map (scale 1:25000) and aerial photograph (scale 1:2500) are used as supporting material where all sampel segments and observation points are plotted on it. The result can be used as a field map for survey guidance Figure 5. illustrates how to plot sample segments on land use map and observation points on aerial photograph regarding to their geographical coordinates.

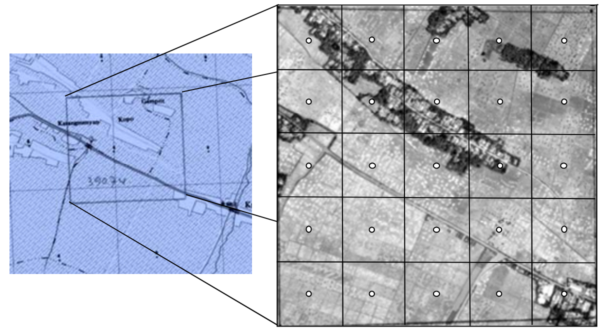


Figure 5. Illustration of sampel plot on land use map

By using the above guidance the surveyors can look for the location of sample segments and observation points based on the features available on the maps, such as road, settlement, irrigation canal etc. The aim of the survey is to collect rice data directly from the field by observing rice growing stage in each point inside the sample segment.

The field survey activitiy conducted by agricultural officers come to each observation point and write down the stage of rice crop, as follows:

- Vegetative-1 (V-1), which the approximate stage of rice is 1-35 days after transplanting
- Vegetative-2 (V-2), which the approximate stage of rice is 35-55 days after transplanting
- Generative (G), which the approximate stage of rice is 55 days up to harvest
- Harvested (P) which is the stage of rice from harvested up to land preparation
- Land Preparation (PL), which is the rice field being prepared
- Others (LL) if the land is not for rice crop cultivation

Knowing the above rice growing stages is important because the calculation of rice area and the prediction of rice harvested area will refer to those stages.

While rice productivity survey is conducted referring to the method developed by National Bureau of Statistic (BPS), where plots 2.5 m X 2.5 m are placed in an area segment (Ubinan System). Productivity value would be recorded each month on the segment that rice crop is being harvested. The production of rice crop is derived from multiplication between harvested area and productivity.

4.3. Data Cummunication and Analysis

In order to accelerate the delivery of data from the field to the central (Jakarta), a communications systems have been developed by using short message service (sms). Surveyors send field data via sms corresponding to the result of recording the amount of each stage per segment as soon the observation of all points finished. All field data will be stored in the sms server for advanced processing. The address of the system can be visited in: **http:neonet.bppt.go.id/padi/**

The extrapolation formula of sample to the population is also developed in the sms server system. Therefore, the field data submitted by the surveyors would be immediatly calculated by the system to estimate rice crop production. The mathematical formulation for rice production estimate is given below:

- **Production (P)**

P = A x Y dimana,

P is production (Ton)
A is harvest area (Ha)
Y is average yied/productivity (Ton/Ha)

- **Luas Panen (A)**

D_j = Total area of strata j
n_j = sample dimension in strata j
m = number of strata
p_{ij} = proporsion of rice area in segment i strata j
A_j = Total area of rice in strata j j
D, Total area of the frame in the study area

The average proporsion of rice area in strata j and its variance are calculated by the formulation, as folow:

$$\overline{p}_j = \frac{1}{n_j}\sum_{i=1}^{n_j} p_{ij} \qquad \delta^2_{\overline{p}_j} = \frac{1}{n_j(n_j-1)}\sum_{i=1}^{n_j}(p_{ij}-\overline{p}_j)^2$$

Total rice area in strata j and its variance are calculated by the formulation, as follow:

$$A_j = D_j \overline{p}_j \qquad Var\,(A_j) = D_j^2 \delta^2_{\overline{p}_j}$$

The estimation of total rice area in the study area is the addition of rice area in all strata. The mathematical formulation of rice area and its variace, as follow:

$$A = \sum_{j=1}^{m} A_j \qquad Var(A) = \sum Var(A_j)$$

The standard error is calcualted from the root of the variace, then the coefisien variation is stated as the percentage between standard error and estimated area of rice.

- **Productivity (Y)**

The value of productivity is derived by measuring crop cutting (ubinan) of 2.5 m x 2.5 m size adopted from BPS, then those value is converted into ton/hectare.

Harvest prediction for next 2-months is calculated from the addition between the area of vegetative-2 and generative stages (V-2 + G), while for next 4-months is calculated from the addition between the area of land preparation, vegetative-1, vegetative-2, and generative (LP + V-1 + V-2 +G). Those calculation refer to the average rice growth cycle is approximately 105 days. Production value is derived from multiplication between harvested area and productivity.

4. 4. Results

Each time the field survey is conducted and field data are sent to the sms server by the surveyors, then the results of rice estimate can immediately be seen in the web-site **(http:neonet.bppt.go.id/padi/).** In this year, field survey has been conducted four times with the schedule below:

- Survey-1: 26 April – 04 May 2010
- Survey-2: 14 June – 21 June 2010
- Survey-3: 13 July – 19 July 2010
- Survey-4: 20 August – 29 August 2010

Tabel 1. The results of rice area estimate based on area frame of square segments by points

Growing Stage	Estimated Area (Ha)			
	Survey-1	Survey-2	Survey-3	Survey-4
Land Preparation	33,634	14,194	0.00	0.00
Veg-1	23,544	38,552	22,087	3,725
Veg-2	5,382	33,971	19,732	19,271
Generative	14,575	1,638	54,712	16,356
Harvested	43,612	3,887	14,911	77,175
Harvest bet. 2-survey	0.00	5,677	0.00	558
Cum. harvest		9,564	14,911	77,733
Others	24,777	27,458	27,468	25,749
Total rice field	120,748	112,241	111,443	116,527
Coef. of Var. (%)	4.1	5.0	4.64	4.01

Harvest next 2-month	19,957	55,609	74,445	35,627.78
Harvest next 4-month	77,135	108,355	96,531	39,352.50
Cum. Prod. (Ton)		74,763	116,565	607,656.94
Prod.next 2-month (Ton)	156,004	434,708	581,948	278,509.46
Prod.next 4-month (Ton)	602,983	847,030	754,605	307,626.36

Most of the farmers in Indramayu District have two rice crop season a year then followed by bare soil, secondary crop or third rice crop. The first planting season is generally started early up to mid November, while the second panting season is generally started at mid up to end of April.

Every survey of area frame sampling is done the related data of rice are derived, such us rice growing stages, harvested area, total rice area, predicted harvest area and production. Tabel …, shows the results of four time field surveys in Indramayu district with the time interval around 1 month each other.

Results on Survey-1 show that rice fields are dominated by Harvested, land preparation and vegetative-1 stages, i.e. 43,612.88 hectares, 33,634.62 hectares, and 23,544.23 hectares respectively indicating that the first harvest season has mostly finished and the second planting season has been started. Cumulative harvest is the sum of harvested area between two-survey and present survey. So that there is no cumulative harvest for survey-1 but it can be calculated for survey-2 onward. Total rice field explains the rice field which is utilized by farmers for rice cultivation, and it exclude non rice cultivation, even though most of farmers cultivate rice crop during April or May. It is to say that 120,748.27 hectares of total rice field in survey-1 give a figure of whole total rice field in Indramayu District. The precision of total rice field estimation in survey-1 is indicated by coefficient of variation, that is 4.06 %.

The prediction or forecast of harvest area for next two-month and four-month in each survey is based on the area of rice growing stages when the survey is conducted. Foe example, the prediction of rice harvest area for next two-month in survey-1 is the sum between the area of vegetative-2 and generative stages, that is 19,956.54 hectares . Furthermore, the area of rice harvest area for next four-month is the sum between the area of land preparation, vegetative-1, vegetative-2, and generative stages, that is 77,135.38 hectares.

Calculation of rice production is based on the multiplication between harvested area and yield, where harvested area is derived from field survey of area frame system and yield is derived from plot cutting (ubinan) system conducted in each segment. The production data include actual production and prediction.

IV. CONCLUSION

1. There are various types of approaches on agricultural statistics, which one of them is area frame sampling
2. Area frame of square segment by points is a simple approach to implement on rice statistics
3. it is a scientifically sound approach based on statistical analysis and unbiased by subjectivity
4. By using sms system for data communication allow to get near real time results
5. It is based on low technology and does not need high investment costs
6. implementation costs are very low once it has become a routine activity
7. With slight adaptations it can be applied to other crops

REFERENCES

.
BPPT-EU, 2001. SARI (satellite Assessment of Rice in Indonesia. Final Phese A Technical Report. Jakarta.

Carfagna, E., and F.J. Gallego, 1994. On the optimum size of segments in the two-step survey on area frame: Possible improvements for rapid European estimates. Conf. on MARS Project overview and perspectives Office for Publications of the E.C. Luxemberg.

Cochran, W., 1977. Sampling Techniques: John Wiley & Son, New York.

Cotter, J., J. Nealson, 1987. Area Frame Design for Agricultural Surveys. U.S. Dept. fo Agric. Nat. Agric. Serv. Washington

Cressie, N., 1993. Statistics for Spatial Data. Revised Eddition. John Weley & Son, New York.

F.J. Gallego, 1995. Sampling Frames of Square Segments. Report EUR 1631 EN. Joint Research Centre, European Commission, Luxembourg

Mubekti and G. Hendrarto, 2010. Sistem estimasi dan peralamalan produksi pertanian dengan kerangka sampel area. Pros. Pertemual Ilmiah Tahunan XVII dan Konggres Mapin V. Bogor, 8-9 Agustus 2010.

Porchier, J.C., 1990. La Teledetection dans dans le Programme d'Enguetes du SCEES. Conference on the Appl. Of Remote Sensing to Agricultural Statistics. Office for Publication of the E.C. Luxemberg.

Ripley, B.d., 1981. Spatial Statistics, John Weley & Sons. New York.

Taylor, C. Sannier, J. Delince, FJ. Gallego, 1997. Regional Crop Inventories in Europe Assisted by Remote Sensing : 1988-1993, Synthesis Report of the MARS Project – Action I, Joint Research Centre, Luxembourg.