Mapping of Provincial Food Security in Indonesia Using Based Clustering Model

Riski Sayuti Rahayu*, Yunastiti Purwaningsih, Akhmad Daerobi

Faculty of Economics and Business, Sebelas Maret University

Corresponding Author: riski_sr@bps.go.id

Abstract

Indonesia was known as an agrarian and maritime country, should not experience difficulties in fulfill food needs or having high food security. However, it is a formidable challenge for the Indonesia to meeting food needs. The low level of food security was caused more by Indonesia's geographical conditions in the form of islands that cause inequality of food production, distribution and absorption among provinces in Indonesia. To reduce the occurrence of food security inequality between provinces in Indonesia, clusters was formed based on food security indicators. Based clustering technique is chosen to overcome the problem of overlapping and the limited availability problem in food security data. The results of research produce 3 clusters based on the classification of food security levels. Based on Bayesian Information Criterion, the most fitted cluster model is a three-cluster model with diagonal distributions. The first cluster consisted of 19 provinces with a classification of middle food security levels, the second cluster consisted of 10 provinces with the classification of the level of high food security, and the third cluster consisted of 5 provinces with a classification of low food security levels. It is expected that the results of this clustering can provide input to the Indonesian government to focus more on 5 Provinces with low food security classification, which focuses on access to food.

1. Introduction

Food is the most basic human needs, so its availability for community must be guaranteed and it is an obligation of a country, including Indonesia. As mandated in the 1945 Constitution of Republic of Indonesia, the State is obliged to carry out food sovereignty (food right) and fulfill population food need.

Indonesia is well known as an agrarian and maritime country, should not experiance difficulties in fulfill food need for its population or should have high food security. However, fulfillness of food needs for Indonesia which consists of 34 Provinces is a big challenge. Its shown from data of Global Hunger Index that released by IFPRI (International Food Policy Research Institute) in 2017. Based on IFPRI data, Indonesia was catagorized country that experiance serious hunger with value of global hunger index of 22 (Research Institute (IFPRI), 2017)multinational study aimed to confirm in a community practice setting the efficacy and safety of cetuximab plus irinotecan in patients with epidermal growth factor-expressing metastatic colorectal cancer (mCRC. This condition shows that Indonesia's food security is very low.

The serious level of hunger and low food security in Indonesia caused by several factors, one of them is Indonesia's geographical condition which in the form of islands. This condition causes inequality of production, distribution, and absorption of food between provinces in Indonesia due to the differences of potential natural resource, human resource and level of development between provinces in Indonesia (Fagi, 2013).

The first step to reduce the occurrence of food security inequality between provinces in Indonesia is mapping the state of food security each province in Indonesia. By knowing the state of food security each province in Indonesia, it will be known which provinces are categorized priority provinces to have immediate policy to increase the level of food security implemented.

Mapping of food security in Indonesia has not yet gotten a fixed model, because there are many variables that affect food security in an area, and the availability of data is quite complicated and limited, so mapping of food security in Indonesia is using variable approaches that are expected to illustrate food security in an area (Ismawan, 2009). There are two models of food security mapping that have been carried out in Indonesia, the mapping of food security carried out by Statistics Indonesia Agency and food security mapping conducted by the Food Security Agency collaborated with the World Food Program (WFP).

Statistics Indonesia has mapped food security in 2013 by calculating the composite index of agricultural business household food security. Food security mapping was carried out in 2013, because the data used were data on results of the 2103 Agricultural Business Household Income Survey (SPP). The Agricultural Business Household Food Security Index divided 34 provinces in Indonesia into three categories of food security, namely Food Security Category High (IKP ≥83.01), Sufficient Food Security (74.92 P IKP <83.01) and Less Food Security (IKP <74.92). The results is only 6 Provinces in Indonesia that categorized high level food security, 24 Provinces in Indonesia are categorized sufficient level food

security, and 4 Provinces are categorized low level food security (Badan Pusat Statistik, 2013).

The Food Security Agency (BKP) collaborated with the World Food Program (WFP) developed a Food Security and Vulnerability Atlas (FSVA). This map was created to illustrate the conditions of regional food security and vulnerability in Indonesia (Dewan Ketahanan Pangan, 2015). The FSVA was developed using 9 indicators of chronic food insecurity and 4 indicators of temporary food insecurity. In determining the status of regional food security and vulnerability, this map uses a composite index of all indicators of food insecurity by weighting based on Principal Component Analysis. The data used in making FSVA is secondary data obtained from the Food Security Agency and publications from Statistics Indonesia, Ministry of Health, Ministry of Forestry, National Disaster Management Agency, and the Meteorology and Geophysics Agency (Badan Ketahanan Pangan NTT, 2015).

In contrast to the previous mapping conducted by the Statistics Indonesia and the World Food Program, which in classifying regional food security status in Indonesia using a composite index, this study used a Multivariate Normal Based Clustering model in classifying regional food security status in Indonesia with a three-dimensional approach of food security used by Statistics Indonesia and the Food and Agricultural Organization (FAO), they are food availability, food affordability and food use. Multivariate Normal Based Clustering model is chosen because its has feature that can overcome overlapping problems that often accur in other clustering methods and this analysis accommodates the limited availability of food security data in Indonesia compared to the composite index method (Damayanti, 2007).

This study aims to mapping the condition of food security in each province in Indonesia. While the benefits of this research are giving the information for the steakholder that can be used as basic planning of policy in effort to increase the food security or reduce the risk of food insecurity in Indoneisa.

2. Research Method

This study uses secondary data with reference to data year 2017 from literature books, economic and business journals, and publications from Statistics Indonesia. While the analysis method used is different from the previous research that using composite index, this research is using clustering method using normal based clustering model to classify 34 Provinces in Indonesia according to their food security state. This clustring method is based on the dimensions of food security variabels from simplification definition of food security proposed by USAID (United States Agency for International Development). USAID defines food security as a condition when all people at all times have physical and economic access to obtain their consumption needs for a healthy and productive life (Nurhemi & Soekro R., 2014). Based on this definition, food security has a broad and complex concept influenced by interactions of various agro-physical, socio-economic and biological factors. Like the concept of health or social welfare, there is no definite measure in measuring food security. However, the complexity of measuring food security can be simplified by focusing calculations in three dimensions (Rivani, 2011), namely:

Food Availability, it is reached when the amount of food is sufficient and is available consistently for all individuals in a country. This food can be fulfilled through domestic production, imports and food aid. In this study, ratio of food crops availability to normative consumption per capita is used to represent the dimensions of food availability.

Food Access, it is guaranteed when the household and all individuals in it have adequate resources to obtain nutritious food. This access depends on household income, household income distribution and food prices.

Food Utilization, it is the use of biologically appropriate food, which makes a food meet important energy and nutrition needs, drinking water, and adequate sanitation. Effective use of food depends on how broad is household knowledge in storage and good food processing techniques, basic principles of proper nutrition and child care, and disease management.

Prior to grouping provinces based on their food security status the first thing to do is to determine indicators in the formation of provincial clusters in Indonesia according to their food security state. This needs to be done because not all indicators can describe the condition of food security in a region with good and have complete availability of data . Based on the differences in the availability and the completeness of the region's data, this study use six indicators that have been identified and compiled to determine the status of provincial food security in Indonesia. All of these indicators are considered to represent the dimensions of food security according to USAID and FAO (Food and Agriculture Organization) and shown in Table 1.

Table 1. Food Security Indicators Based on USAID and FAO's Food Security Dimensions

No.	Dimensions	Indicators	Denomination
A	1. Food Availability	1. Ratio of Food Crops Availability to Normative Consumption per Capita	Gram/kapita/day
B	Food Access	2. Percentage of population living below the poverty line	Percent
		3. GRDP (Gross Regional Domestic Pruduct) Per Capita	Million Rupiah/capita/ year
C	Food Utilization	4. Life Expectancy	Year
		5. Average consumption of calories per capita per day	Ons/kapita/day
		6. Average consumption of per capita protein per day	Ons/kapita/day

Step to calculate the Ratio of Food Crops Availability to Normative Consumption per Capita by calculating the average net production of rice,corn, cassava and sweet potato at the provincial level using standard conversion factors. For the average net production of cassava and sweet potato divided by 3 (cereal conversion factor) to get an equivalent valuewith cereals. Then the total production of cereals is worth consuming per day. The results are divided by the normative consumption value of cereals / day / capita (300 grams / person / day) and the number of population in the middle of the year so that the Ratio of Food Crops Availability to Normative Consumption per Capita is obtained (Suhartono, 2010).

Based clustering model was used for grouping objects in a population that describes a clustering approach, where a group of populations is identified based on probability distributions and the entire population is modeled as a distribution mixture (Melnykov & Maitra, 2010). Multivariate normal based clustering model uses the assumption that in a population, sub-populations can be taken which have normal probability distribution and each sub-population has different parameters. In the multivariate normal based clustering model there are at least two main processes that need to be done, there are parameter estimation and selection of the best model that describes the data structure. The selection of the best model can be done using the bayesian approach, the best model that describes the data structure will be obtained (McLachlan & Peel, 2017). Bayesian Information Criterion (BIC) is a measure that can be used in the selection of models in this study. The use of Bayesian Information Criterion (BIC) in this study is because BIC has a more robust nature

(Agusta, 2007). In general, the formula used is as follows:

$$BIC = 2 \ln LM (\Psi) - m_p \ln (n) \qquad (1)$$

Where is the maximum value of the ln likelihood function for the model and m_p is the number of observation objects. The best model is the model that produces the largest BIC value. After obtaining the best model, parameter estimation and clustering of research objects were carried out, research objects in this study were 34 provinces in Indonesia. To calculate multivariate normal based clustering, this research uses R software package McLust 5.4.1 (Fraley, Raftery, Murphy, & Scrucca, 2012).

The measurement of food security classification is done by comparing the cluster mean with confidence interval (95%) of the variables used in the study. The confidence interval is calculated using the following formula (Arifiliana, 2017):

$$\bar{X} - Z\alpha_{/2}\sqrt{\frac{\sigma^2}{n}} < \mu < \bar{X} + Z\alpha_{/2}\sqrt{\frac{\sigma^2}{n}} \qquad (2)$$

With $\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n}$, n = number of object observation and α = significant level, used 5%.

If the mean cluster is below the lower limit of the confidence interval then the classification of food security achievement is low, if the mean cluster between the confidence intervals then the classification of food security achievement is intermediate, and if the mean cluster is upper limited to the confidence interval then the classification of the class is high food security achievements.

3. Results and Discussion

3.1 Results

Table 2. Three Best Models by BIC Value

Model	Distribution	Number of Cluster	BIC Value
VVE	Ellipsoidal	2	-538,9258
EVE	Ellipsoidal	2	-540,2353
EVI	Diagonal	3	-555,8095

Table 3. Log Likelihood Value, Number of Samples, Degree of Freedom and BIC Value

Log Like lihood	**n**	**Degree of Freedom**	**BIC**
-214,4303	34	36	-555,8095

Table 4. Mean and Variant Clusters and Mixing Probabilities

	Mean Cluster			*Varians Cluster*		
	1	**2**	**3**	**1**	**2**	**3**
X_1	6,0339	1,5955	0,7309	7,5574	1,3393	0,2755
X_2	11,6532	5,9573	19,8740	12,3303	2,0230	73,3221
X_3	38,8533	100,8553	40,7793	57,7628	3.803,9590	746,5741
X_4	69,4948	71,1424	65,8940	5,6000	2,1483	1,0308
X_5	2,1623	2,1298	1,9002	0,0055	0,0086	0,0086
X_6	0,0616	0,0635	0,0518	0,0000	0,0000	0,0000
Mixing Probabilities			0,5814	0,2715	0,1471	

Table 5. Classification of Food Security Status Between Clusters

	Confidence Interval		Mean Cluster			Food Security			The difference relative to the overall mean		
	Lower Limit	Upper Limit	1	2	3	1	2	3	1	2	3
X_1	2,93	5,16	6,03	1,60	0,73	H	L	L	1,98	-2,45	-3,32
X_2	9,30	13,34	11,65	5,96	19,87	M	H	L	0,34	-5,36	8,56
X_3	41,28	70,67	38,85	100,86	40,78	L	H	L	-17,12	44,88	-15,19
X_4	68,52	70,30	69,49	71,14	65,89	M	H	L	0,08	1,73	-3,52
X_5	2,07	2,16	2,16	2,13	1,90	H	M	L	0,05	0,01	-0,21
X_6	0,06	0,06	0,06	0,06	0,05	M	H	L	0,00	0,00	-0,01
Conclusion						M	H	L			

Remarks :

L=Low, M=Medium, H=High

Blok shows negative variables

Table 6. Provinces in Indonesia by Food Security Status

Food Security Status		
High Level of Food Security	**Middle Level of Food Security**	**Low Level of Food Security**
Riau Province, Jambi Province, Bangka Beliung Province, Kepulauan Riau Province, Jakarta Province, Banten Province, Bali Province, Kalimantan Tengah Province, Kalimantan Utara and Kalimantan Timur Province	Aceh Province, Sumatera Utara Province, Sumatera Barat Province, Sumatera Selatan Province, Bengkulu Province, Lampung Province, Jawa Barat Province, Jawa Tengah Province, Yogyakarta Province, Nusa Tenggara Barat Province, Kalimantan Barat Province, Kalimantan Selatan Province, Sulawesi Utara Province, Sulawesi Tengah Province, Sulawesi Tenggara Province, Gorontalo Province, and Sulawesi Barat Province	Nusa Tenggara Timur Province, Maluku Province, Maluku Utara Province, Papua Barat Province, and Papua Province

3.2 Discussion

The first step that needs to be done to form a cluster with based clustering model is choosing the best model. The selection of the best model in this study is done by Bayesian Information Criterion.Using Mclust 5.4.1 software, this calculation produced 9 models, which each model was tried using the number of clusters ranging from 1-9 clusters. The largest BIC value indicates that the model used with a certain number of clusters is the most feasible model or the best model.The processing results recommend three best models with optimal number of clusters with each BIC value, as presented in Table 2.

The best model according to BIC criteria is the EVI model with the number of clusters as much as 3. the EVI model is a cluster model with diagonal distribution, the volume is Equal (same), the shape is variable (different) and its orientation is coordinate axes. BIC criteria also becomes basis for obtaining the number of optimal culsters and members of each clusters. The summary of the EVI model with 3 clusters is shown in Table 3.

After getting the best model, the next step to do is estimating the parameter. Parameter estimation using EM algorithm. Model based clustering produces several parameters including mixing probabilities (π), mean vectors (μ) and matrix variance covariance (Σ). Table 4 presents the parameter estimation results with the EM algorithm.Mixing probabilities show cluster membership degree where cluster 1 is 56% or 19 provinces, cluster 2 is 29% or as many as 10 provinces and cluster 3 is 15% or as many as 5 provinces.

To measure the classification of provincial food security status in Indonesia in each cluster, it is done by comparing the group mean formed with a 95% confidence interval calculated from the mean and variable standard deviations. There are two types of variables used in this study, namely positive and negative variables. A positive variable is a variable that, if the value is greater, indicates a good condition, while a negative variable is a variable which, if the value is greater, indicates an unfavorable state. For positive variables, if the mean cluster is below the confidence interval, the classification of food security achievement is low, the mean cluster between the confidence interval of medium food security achievement and if the mean cluster is above the confidence interval, the food security achievement is high. The opposite applies to negative variables.

Positive variables used in this study were ratio of food crops availability to normative consumption per capita (X_1), GRDP (Gross Regional Domestic Product) per capita (X_3), life expectancy (X_4), average consumption of calories per capita per day (X_5), average consumption of

per capita protein per day (X_6). While the negative variables used are the percentage of people living below the poverty line (X_2). The results of the classification of the provincial food security achievement of each cluster can be seen in Table 5.

Based on Table 5, in Cluster 1 most of the variables have a classification of food security in the middle level, except X_1 (ratio of food crops availability to normative consumption per capita) and X_5 (average consumption of calories per capita per day) which have a high level classification, while X_3 (GRDP per capita) have low level classification. This means that the province in the classification of middle food security have good ratio of food crops availability to normative consumption per capita and average consumption of calories per capita per day, but still low in terms of the purchasing power of represented by GRDP per capita which is an approach of per capita income.

In cluster 2 almost all variables have a classification of high food security levels, except X_1 (ratio of food crops availability to normative consumption per capita) has a classification of low food security level. So that it can be concluded that the provinces included in cluster 2 are provinces that have high food security classification. Condition ratio of food crops availability to normative consumption per capita that have low level, means that provinces in Indonesia which have high food security have a low food availability. This condition is happen because food availability not only comes from the food corps but it also comes from province's domestic food productions and inter-provincial food imports. However, data on the number of domestic food productions and food imports between provinces is difficult to obtain.

While in cluster 3 all variables have a low level of food security classification. So it can be concluded that the provinces included in cluster 3 have a classification of low food security levels.

To improve the occurrence of food security inequality between provinces in Indonesia, the government can focus more on provinces with a classification of low levels of food security in implementing development programs. The variables that need to get more attention are the variables whose conditions are far below the overall average, namely the GRDP Per Capita (X_3) variable which is an indicator of food access. GRDP Per Capita is commonly used as a per capita income approachment that can reflect people's purchasing power, the higher GRDP Per Capita the higher people's purchasing power the easier people's food access in the Province.

Provinces that fall into the category of middle food security (cluster 1), are Aceh Province, Sumatera Utara Province, Sumatera Barat Province, Sumatera Selatan Province, Bengkulu Province, Lampung Province, Jawa Barat Province, Jawa Tengah Province, Yogyakarta Province, Nusa Tenggara Barat Province, Kalimantan Barat Province, Kalimantan Selatan Province, Sulawesi Utara Province, Sulawesi Tengah Province, Sulawesi Tenggara Province, Gorontalo Province, and Sulawesi Barat Province. While Provinces that fall into the category of high food security (cluster 2), are Riau Province, Jambi Province, Bangka Beliung Province, Kepulauan Riau Province, Jakarta Province, Banten Province, Bali Province, Kalimantan Tengah Province, Kalimantan Utara and Kalimantan Timur Province. Almost all are provinces that fall into categoriy high food security, have a large potential economic. It can be seen on the value of GRDP of those Provinces that have higer value than other Provinces that fall into cluster 1 nor cluster 3.

Provinces that fall into the category of low food security are Nusa Tenggara Timur Province, Maluku Province, Maluku Utara Province, Papua Barat Province, and Papua Province. Those Provinces low level of all variables of food security that used on this study. This means that those Provinces are reasonable to be priority provinces for immediate implementation of food insecurity prevention policies to improve their achievement of food security and in the long term it will improve indonesian food security state.

The provincial group according to the food

security status produced using the normal based clustering method is almost the same as using the composite index method carried out by BPS and WFP, Where the eastern Indonesian province is a province that has low food security or is in a high food insecurity area. This shows that the normal based clustering method is a method that can be used to mapping regional food security in Indonesia. this method can be used as a recommendation to mapping regional food security in Indonesia every year because this method does not require complete data availability to approach the dimension of measuring food security. while the mapping of food security using the composite index method cannot be done every year because this method requires complete data in approaching the dimensions of measuring food security.

4. Conclusions

The results of the study using based clustering model, produced 3 clusters based on the classification of food security levels. The first cluster consists of 19 provinces with the classification of the level of middle food security, the second cluster consists of 10 provinces with a classification of high food security levels, and the third cluster consists of 5 provinces with a classification of low food security levels. The variable that needs to get more attention is the GRDP Per Capita (X_3) which is an indicator of food access.

To improve the occurrence of food security inequality between provinces in Indonesia, the government can focus more on provinces with a low level of welfare level in implementing a program to strengthen food security. The program to strengthen food security that is designed and implemented should be more focused on programs to increase food access.

5. Acknowledgement

6. References

Agusta, Y. (2007). K-Means – Penerapan, Permasalahan dan Metode Terkait. *Jurnal Sistem Dan Informatika*, *3*(11), 47–60.

Arifiliana, E. (2017). *Multivariate Normal Based Clustering Kabupaten / Kota Di Jawa Timur Berdasarkan Indikator Kesejahteraan*. Universitas Padjadjaran.

Badan Ketahanan Pangan NTT. (2015). *Peta Ketahanan dan Kerentanan Pangan Nusa Tenggara Timur 2015*. Kupang: Pemerintah Provinsi Nusa Tenggara Timur, Dewan Ketahanan Pangan, Kementerian Pertanian and World Food Programme (WFP).

Badan Pusat Statistik. (2013). *Analisis Sosial Ekonomi Petani di Indonesia*. Jakarta: Badan Pusat Statistik.

Damayanti, L. (2007). Faktor yang Mempengaruhi Tingkat Ketahanan Pangan Desa (Studi Kasus di Kabupaten Malang). *Agroland*, *14*(3), 217–222.

Dewan Ketahanan Pangan. (2015). Food Security and Vulnerability Atlas Indonesia 2015 : Summary version Food Security and Vulnerability Atlas of Indonesia 2015. In *World Food Programme*.

Fagi, A. M. (2013). Ketahanan Pangan Indonesia dalam Ancaman: Strategi dan Kebijakan Pemantapan dan Pengembangan Threatened Indonesia Food Security : Strategy and Policy for Stabilization and Development angin pasat Barat Laut yang mendukung intensifikasi pertanian tetapi k. *Jurnal Analisis Kebijakan Pertanian*, *11*(1), 11–25.

Fraley, C., Raftery, A. E., Murphy, T. B., & Scrucca, L. (2012). mclust Version 4 for R: Normal Mixture Modeling for Model-Based

Clustering, Classification, and Density Estimation. *Technical Report*, 1–50. https://doi.org/Technical Report No. 504

Ismawan, S. H. (2009). Analisis Ketahanan Pangan Indonesia Tahun 2007. Institut Teknologi Sepuluh Nopember.

McLachlan, G. J., & Peel, S. D. N. (2017). On Clustering by Mixture Models. *Department of Mathematic University of Queensland*, *91*, 399–404.

Melnykov, V., & Maitra, R. (2010). Finite mixture models and model-based clustering. *Statistics Surveys*, *4*(0), 80–116. https://doi.org/10.1214/09-SS053

Nurhemi, S. R. I., & Soekro R., G. S. (2014). Pemetaan Ketahanan Pangan di Indonesia: Pendekatan TFP dan Indeks Ketahanan Pangan. In *Bank Indonesia*.

Research Institute (IFPRI), I. F. P. (2017). *2017 global hunger index: The inequalities of hunger*. https://doi.org/10.2499/9780896292710

Rivani, E. (2011). Determination of the Dimensions and Indicators of Food Security in Indonesia: the National Food Security Board-World Food Program Methodology Revisited. *Widyariset*, *14*(1), 151–162.

Suhartono. (2010). Indikator dan pemetaan daerah rawan pangan dalam mendeteksi kerawanan pangan di kecamatan tanjung bumi kabupaten bangkalan. *Embryo*, *7*(2), 97–110.

7. Appendixes.
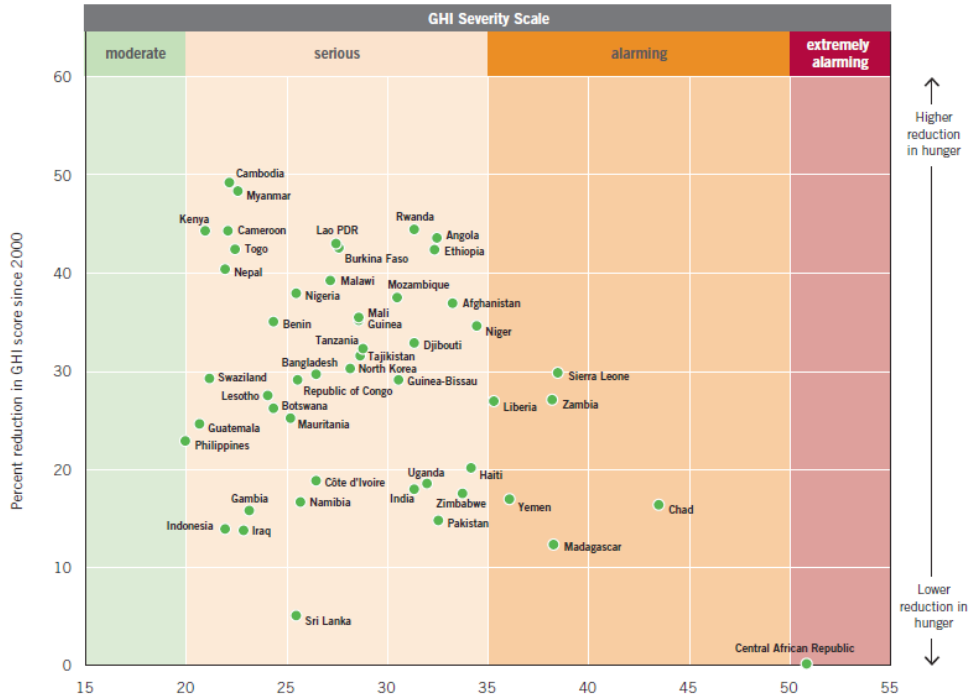
**Appendix A. 2017 Global Hunger Index**

Appendix B
Indeks Ketahanan Pangan Rumah Tangga Usaha Pertanian
Menurut Provinsi di Indonesia Tahun 2013

Provinsi	Indeks Ketahanan Pangan
Aceh	76,69
Sumatera utara	76,69
Sumatera Barat	81,67
Riau	76,45
Jambi	78,21
Sumatera Selatan	77,54
Bengkulu	77,41
Lampung	79,97
Kep. Bangka Belitung	73,58
Kep. Riau	76,20
DKI Jakarta	84,78
Jawa Barat	81,22
Jawa Tengah	85,85
DI Yogyakarta	86,03
Jawa Timur	85,61
Banten	80,60

Provinsi	Indeks Ketahanan Pangan
Bali	83,09
Nusa Tenggara Barat	79,31
Nusa Tenggara Timur	69,62
Kalimantan Barat	71,91
Kalimantan Tengah	75,76
Kalimantan Selatan	84,94
Kalimantan Timur	81,98
Kalimantan Utara	73,79
Sulawesi Utara	77,70
Sulawesi Tengah	81,69
Sulawesi Selatan	82,82
Sulawesi Tenggara	80,04
Gorontalo	76,07
Sulawesi Barat	75,61
Maluku	79,23
Maluku Utara	76,33
Papua Barat	79,19
Papua	77,17
Indonesia	81,37

Sumber : Badan Pusat Statistik, 2013
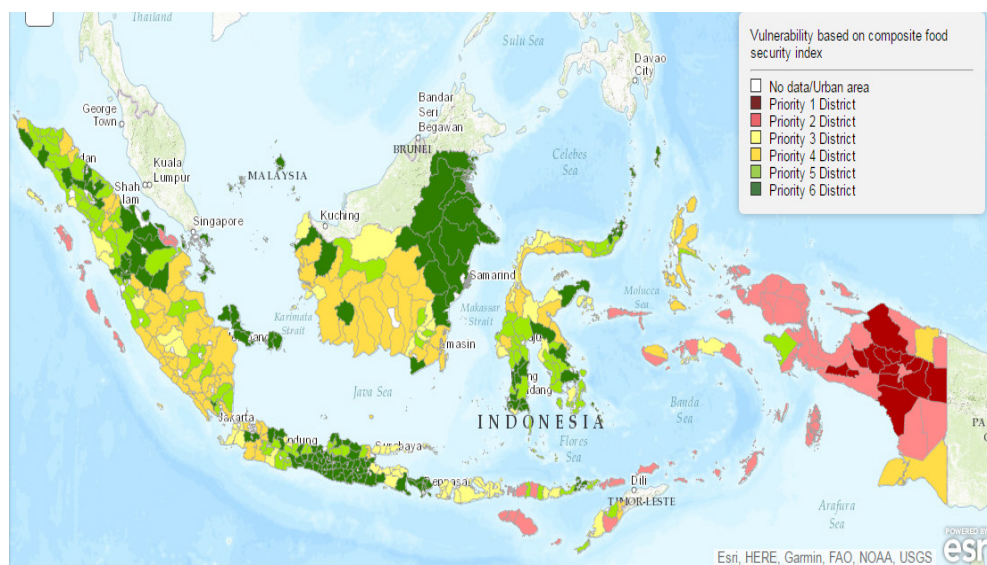
Appendix C. Food Security and Vulnerability Atlas 2015