International Journal of Innovation, Creativity and Change. www.ijicc.net
Volume 9, Issue 12, 2019

Hybrid Correspondence with PCA Biplot for Grouping Districts/ Cities of West Java Based on Toddler Nutritional Status and the Causes of Malnutrition

Titi Purwandari[a], Yuyun Hidayat[b], Irlandia Ginanjar[c], Sukma Budi Prasetya[d], Sukono[e], [a,b,c,d]Department of Statistics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Indonesia, [e]Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Indonesia, Email: [a]titipurwandari@yahoo.com, [b]yuyun.hidayat@unpad.ac.id, [c]irlandia_g@unpad.ac.id, [d]Sukmabudiprasetya@rocketmail.com, [e]sukono@unpad.ac.id

Nutritional status of toddlers is divided into four groups: poor nutrition, lack of nutrition, good nutrition, and over nutrition. The province of West Java still has the highest case of malnutrition in Indonesia. Malnutrition is caused by the absence of exclusive breastfeeding; not implementing Clean Healthy Life Behaviors (CHLB); unemployment; lack of health workers; level of parental education; poverty, and other factors. This study aims to describe the Regency / City and the causes of malnutrition in one map to make the process solution easier. The method used to solve this problem is Hybrid Correspondence with PCA Biplot. Using this tool, the West Java Provincial Health Office will be able to create a map for the Regency / City in light of the causes of malnutrition in the region. Based on the results of this study, a map showing regency / city groups detailing the nutritional status of children and the causes of malnutrition was obtained. From these results, West Java Provincial Health Office can deal with the causes of malnutrition optimally.

Key words: *West Java Province Indonesia, Poor nutritional status, hybrid correspondence, PCA Biplot.*

Introduction

Poor nutrition is a condition of lack nutrition in children based on body weight index according to age (WI/A) < -3 Standard Deviation (SD). Lack nutrition is a condition identified in children based on body weight index according to age (WI/A) of -3 elementary school to -2 elementary school. Good nutrition is a state of nutrition in children based on body weight index according to age (WI/A) of -2 elementary school up to +2 elementary school. Over nutrition is a condition of excessive nutrition based on body weight index according to age (WI/A) of > +2 SD (Sutarjo et al., 2014; 2015).

Malnutrition is caused by two factors: direct causes and indirect causes. The direct causes are for example: that toddlers are not fed breast milk exclusively; have unbalanced food intake and or are infected with congenital diseases. The indirect causes of malnutrition included limited food supply, inadequate children care, lack of sanitation and clean water and inadequate basic health services, lack of education and poor parental knowledge, the lack of empowerment of women in families, limited community resources, unemployment and poverty (Sutarjo et al., 2014; 2015 ; Paramita et al., 2017).

The World Health Organization (WHO) in 1999 classified regions based on the prevalence of malnutrition into four groups: low (under 10%), moderate (10-19%), high (20-29%), and very high (in over 30%). Based on the prevalence of malnutrition, in 2004 Indonesia was classified as a country with a high malnutrition status (28.47%) and yearaly, the problem of malnutrition is increasing. The risk of death from children who are malnourished is 13 times greater than for normal children. WHO estimates that 54% of the causes of infant and under-five deaths are based on their poor nutritional conditions (Sutarjo et al., 2014; 2015; Paramita et al., 2017).

Based on the 2010 population census, the total population of Indonesia reached 237.56 million with a population growth rate of 1.49 percent, every day 10,000 babies are birthed in Indonesia (Sutarjo et al., 2014; 2015; Paramita et al., 2017). The Basic Health Research (Riskesdas) in 2010 found that the number of children under five in Indonesia reached 23 million and 4.5 percent or 900,000 of those children suffer malnutrition.

The health profile of West Java Province, 2010 was that the proportion of malnutrition rates in 2010 was lower than in 2009 with 388 children under five suffering from malnutrition. This is a problem because malnutrition has decreased very slowly. In addition, there are still a number of regencies / cities in West Java that experience increased malnutrition. Basic Health Research (Riskesdas) data in 2010 reveals the prevalence of children under five with malnutrition in 2010 was 4.9% and these results are quite good in progress towards the MDG's target (*Millennium Development Goals)*. It is expected that in 2015 the MDG's target could reach 3.6% (Bappenas, 2010). From the discussion above, it can be concluded that the West

Java Provincial Health Office's efforts in tackling malnutrition have not been optimal (Sutarjo et al., 2014; 2015; Paramita et al., 2017; Jake, 2017). The problem is how to group districts / cities based on the nutritional status of children and the causes of malnutrition. Therefore, this research aims to apply *Hybrid* Correspondence with PCA Biplot to classify Districts / Cities based on nutritional status of children and causes of malnutrition in West Java Province. The benefit of this research is that the results of this mapping can be used as a guideline for the West Java Provincial Health Office in the combat against malnutrition in West Java.

Research Methodology

The analysis used in this study was hybrid correspondence with PCA Biplot. The object of observation in this study is the Regency/City in West Java Province. The column category in this study is the nutritional status of toddlers and their respective poor nutrition, lack of nutrition, good nutrition, and over nutrition. The characteristics of the factors in this study are the causes of direct and indirect malnutrition: non-exclusive breastfeeding, poverty, lack of health care workers, non-CHLB households (Clean Healthy Living Behaviors), unemployment, and the level of education of parents.

Toddler nutritional status was data mapped with correspondence analysis. The first step was to make a contingency matrix as follows:

$$\mathbf{N}_{(I \times J)} = (n_{ij}),\ n_{ij} > 0$$

Then a correspondence matrix was developed as follows (Nieto et al., 2014 ; Bro et al., 2014) :

$$\mathbf{N}(I \times J) \equiv [n_{ij}]\ ;\ n_{ij} \geq 0$$
$$\mathbf{P} \equiv (1/n..)\mathbf{N}\ ;\ n.. = \mathbf{1}^{\mathbf{T}}\mathbf{NI}$$

The number of rows and columns **P** is written as:

$$\mathbf{r} \equiv \mathbf{P1} \text{ and } \mathbf{c} \equiv \mathbf{P}^{\mathbf{T}}\mathbf{1}$$
$$\text{which } r_i > 0\ (i = 1, ..., I),\ c_j > 0\ (j = 1, ..., J)$$
$$\mathbf{Dr} \equiv \text{diag}\ (\mathbf{r}) \text{ and } \mathbf{Dc} \equiv \text{diag}\ (\mathbf{c})$$

Next the row and column of matrix profile (Opeoluwa and Sugnet, 2017; Evgenidis et al., 2011) were determined:

$$R \equiv D_r^{-1} P \equiv \begin{bmatrix} \tilde{r}_1^T \\ \vdots \\ \tilde{r}_i^T \end{bmatrix} \qquad C \equiv D_c^{-1} P^T \equiv \begin{bmatrix} \tilde{c}_1^T \\ \vdots \\ \tilde{c}_j^T \end{bmatrix}$$

Then the center of the row and the center of the column were determined (Qureshi et al., 2017; Akter et al., 2014):

Center row : $\mathbf{c} = \mathbf{R}^T\mathbf{r}$ and center column : $\mathbf{r} = \mathbf{C}^T\mathbf{c}$
Supposed SVD is from $\mathbf{P} - \mathbf{rc}^T$ is
$\mathbf{P} - \mathbf{rcT} = \mathbf{AD}_\mu\mathbf{B}^T$ which $\mathbf{A}^T\mathbf{D}_r^{-1}\mathbf{A} = \mathbf{B}^T\mathbf{D}_c^{-1}\mathbf{B} = \mathbf{I}$

$\mu_1 \geq \ldots \geq \mu_K > 0$, and the columns of matrix **A** and **B** were defined respectively asthe main axis of the column and the main axis of the row, where :

$$A = \left[\left(\frac{1}{\mu_1} \times e_1 \right) \left(\frac{1}{\mu_2} \times e_2 \right) \ldots \left(\frac{1}{\mu_k} \times e_k \right) \right]$$

$$B = [e_1 e_2 \ldots e_k]$$

$$D_\mu = \begin{bmatrix} \mu_1 & 0 & \cdots & 0 \\ 0 & \mu_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mu_k \end{bmatrix}$$

For example, $\underset{I\times K}{F} = \left(D_r^{-1} \underset{I\times J}{P} - 1c^T \right) \underset{J\times J}{D_c^{-1}} \underset{J\times K}{B}$ is the main coordinate of the line profile against the main axis **B**, then $\mathbf{F} = \mathbf{D}_r^{-1}\mathbf{AD}_\mu$.

For example, $\underset{J\times K}{G} = \left(D_c^{-1} \underset{J\times I}{P^t} - 1r^T \right) \underset{I\times I}{D_r^{-1}} \underset{I\times K}{A}$ is the main coordinate of the column profile against the main axis **A**, then $\mathbf{G} = \mathbf{D}_c^{-1}\mathbf{BD}_\mu$ (Geraldin et al., 2014).

The Kolmogorov-Smirnov Test was used as the normality test of the data on the causes of malnutrition. After normality was fulfilled, then the bivariate normality test was conducted between the scores of Regency/City factors and data on the causes of malnutrition using *Q-Q* plots (Paramita et al., 2018; Rat et al., 2013). After the bivariate normality test was fulfilled, the correlation between district/city factor scores and the malnutrition variable was calculated to obtain the main component matrix from *Pearson* correlation.

$$\rho_{ij} = corr\left(z_i, f_j\right)$$

$$\rho_{ij} = \frac{I \times \sum_i \sum_j z_i f_j - \left(\sum_i z_i\right)\left(\sum_j f_j\right)}{\sqrt{\left(I \times \sum_i z_i^2 - \left(\sum_i z_i\right)^2\right)\left(I \times \sum_j f_j^2 - \left(\sum_j f_j\right)^2\right)}}$$

for $i = 1,2,\ldots p$; $j = 1,2,\ldots r$; assuming Z_i and f_j are normal bivariate distributions; therefore, based on that form the main matrix component is:

$$K = \begin{bmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1r} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p1} & \rho_{p2} & \cdots & \rho_{pr} \end{bmatrix}$$

The vector mapping coordinate of causes of malnutrition was calculated using the equation from PCA Biplot which is $H^T = \left(D_\mu^{1/2}\right)^{1/2} K^T$ (Abdi and Mazandarani, 2016; Kundu et al., 2017).

The first two columns of **H** became the coordinates for vector mapping of malnutrition causes. The percentage of variation (inertia) used as a measure of mapping quality was calculated by : $\tau = \left(1'\mu^2\right)^{-1} \times \mu^2$ (Kahriman et al., 2016; Ukalski and Klisz, 2016).

Results and Discussion

Result

Data on Toddler Nutrition Status was analyzed by correspondence analysis. The following results from the correspondence analysis are:

Table 1: Inertia and District / City Coordinate Points and Nutritional Status of Toddlers

Dimension		1	2
Inertia		65.52585	31.47437
Regency/City	Bogor	-0.04316	-0.01384
	Sukabumi	-0.04455	0.005933
	Cianjur	-0.09037	-0.04779
	Bandung	-0.01053	0.012557
	Garut	-0.03411	0.050021
	Tasikmalaya	-5.04E-05	0.029093

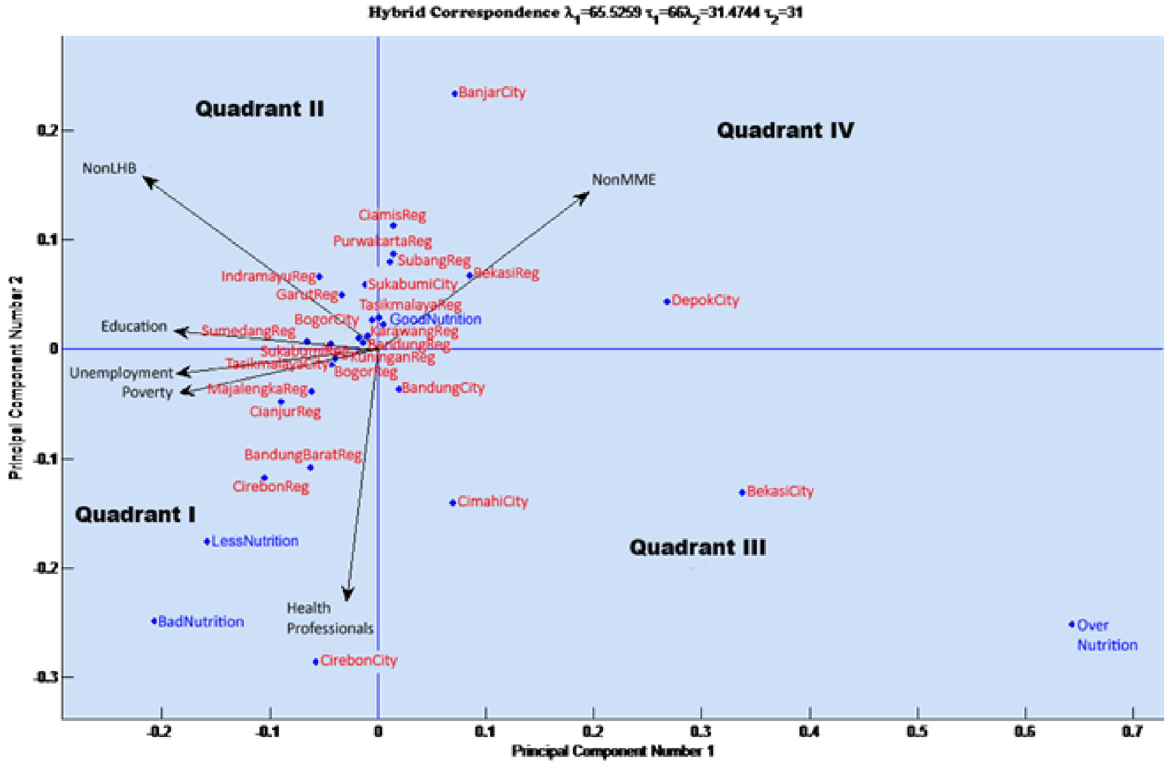
	Ciamis	0.013235	0.113499
	Kuningan	-0.01438	0.006789
	Cirebon	-0.10599	-0.11717
	Majalengka	-0.06261	-0.03762
	Sumedang	-0.06624	0.007928
	Indramayu	-0.05509	0.067225
	Subang	0.011057	0.080865
	Purwakarta	0.013834	0.087273
	Karawang	-0.01828	0.010409
	Bekasi	0,083916	0.067979
	Bandung Barat	-0.0634	-0.10727
	Bogor City	-0.0058	0.027486
	Sukabumi City	-0.01192	0.059734
	Bandung City	0.018704	-0.03574
	Cirebon CIty	-0.05812	-0.28582
	Bekasi CIty	0.336643	-0.1301
	Depok City	0.267597	0.043722
	Cimahi City	0.0686	-0.13973
	Tasikmalaya City	-0.04077	-0.00828
	Banjar CIty	0.071059	0.23459
Toddler Nutrition Status	Poor Nutrition	-0.20746	-0.24769
	Lack Nutrition	-0.15861	-0.17515
	Good Nutrition	0.003945	0.022965
	Over Nutrition	0.642244	-0.25097

Following formation of a correspondence map, the coordinates of the vector mapping that cause malnutrition was calculated.

Dimension	1	2
Non-Breast feed	0.368014	0.263312
Non-CHLBs	-1.099859	0.503297
Health Workers	-0.047946	-0.54404
Poverty	-1.651932	-0.13742
Unemployment	-1.039667	-0.04615
Education	-0.653259	0.045259

Results of Hybrid Correspondence mapping with PCA Biplot are as follows:

Figure 1. Hybrid Correspondence map with PCA Biplot



Discussion

Based on mapping results, the increasing Principal Component I (PC I) value and Principal Component II (PC II) value will make toddler nutritional status better. Based on that grouping Regency /City the map quadrant was formed, the following is the quadrant in question:

1. Quadrant I is quadrant with value (PC I, PC II) <0.
2. Quadrant II is quadrant with the PC I value <0 and PC I I> 0.
3. Quadrant III is quadrant with PC I value > 0 and PC II <0.
4. Quadrant IV is quadrant with value (PC I, PC II)> 0.

Therefore, if sorted from the highest malnutrition status to the poorest nutritional status, based on the Regencies/ Cities location in quadrant I to quadrant IV, the results are as follows: the Regencies/ Cities located in quadrant I are Cirebon, West Bandung, Cianjur, Majalengka, Bogor regency, Tasikmalaya, and Cirebon; Regencies/ Cities located in quadrant II are Sumedang, Sukabumi, Indramayu, Garut, Kuningan, Bandung, Karawang, Bogor City, and Sukabumi; Regencies/ Cities located in quadrant III are Bandung, Cimahi, and Bekasi; Regencies/ Cities located in quadrant IV are Tasikmalaya, Subang, Purwakarta, Ciamis, Bekasi, Banjar, and Depok.

Based on the causes of malnutrition as previously discussed, the findings are that non-breast feeding is positively correlated with PC I and PC II. Education and non-CLHBs are negatively correlated with PC I and positively correlated with PC II. Unemployment, poverty and lack of health workers are negatively correlated with PC I and PC II. Based on this, the rating of the causes of malnutrition from the highest to the lowest is unemployment, poverty, education, and non- CLHBs. Meanwhile the lack of health workers does not have a high correlation with PC I and non-breast feeding has a positive correlation with malnutrition. This might occur because non-breast fed toddlers are given formula milk so that the nutritional necessities of these toddlers is fulfilled.

From the two-dimensional map produced, the inertia value of 97.000222% is obtained, in other words the percent diversity explained by the map is 97.000222% and the map can be used as a basis for obtaining information about malnutrition.

Conclusion

In this paper, research has been conducted on hybrid correspondence with PCA Biplot for grouping districts/ cities of West Java based on toddler nutritional status and the causes of malnutrition. Based on the results of the analysis the following conclusions can be drawn: Regencies/ cities with the highest malnutritional status are Cirebon regency, West Bandung, Cianjur, Majalengka, Bogor Regency, Tasikmalaya city, and Cirebon City. Regencies/ cities with high malnutritional status are Sumedang, Sukabumi, Indramayu, Garut, Kuningan, Bandung Regency, Karawang, Bogor City, and Sukabumi. Regencies/ cities with middle malnutritional status are Bandung City, Cimahi, and Bekasi. Regencies/ cities with low malnutritional status are Tasikmalaya, Subang, Purwakarta, Ciamis, Bekasi, Banjar, and Depok. The rating of malnutritional causes from the highest to the lowest are unemployment, poverty, education, and non-CLHBs. The lack of health worker is not included because it is neither related nor correlated with PC 1 value. Non-breast feeding also has positive correlation with malnutrition and this might occur because the non-breast fed toddlers access formula milk and so their nutritional needs are fulfilled.

Acknowledgements

The authors acknowledge the financial support provided by Universitas Padjadjaran Fundamental Research (RFU), number: 2408/UN6.D/KS/2018.

References

Abdi, H. and Mazandarani, M.T. 2016. Study of Drought Tolerance in Bread Wheat Cultivars Using Biplot. *International Journal of Life-Sciences Scientific Research*, Vol. 2, Issue 6, (2016), pp. 651-657.

Akter A, Jamil Hassan M, Umma Kulsum M, Islam MR, Hossain K, et al. (2014) AMMI Biplot Analysis for Stability of Grain Yield in Hybrid Rice *(Oryza sativa L.)*. *Journal of* Rice Research, 2: 126, pp. 1-4. doi: 10.4172/jrr.1000126.

Bro, R. and Smilde, A.K. 2014. Principal Component Analysis. *Analytical Methods*, 2014,6, pp. 2812–2831.

Geraldin B. Dela Cruz, *Member, IACSIT*, Bobby D. Gerardo, and Bartolome T. Tanguilig III. 2014. Agricultural Crops Classification Models Based on PCA-GA Implementation in Data Mining. *International Journal of Modeling and Optimization,* Vol. 4, No. 5, October 2014, 375-382.

Evgenidis, G., Traka-Mavrona, E., and Koutsika-Sotiriou, M. 2011. Principal Component and Cluster Analysis as a Tool in the Assessment of Tomato Hybrids and Cultivars. *Hindawi Publishing Corporation International Journal of Agronomy,* Volume 2011, Article ID 697879, 7 pages: 1-7.

Jake, M., 2017. Four pillars to building a positive school culture. International Journal of Innovation, Creativity and Change, 3(2): 33-38.

Kahriman, F., Egesel, C.O., Orhun, G.E., Alaca, B., and Avci, F. 2016. Comparison of Graphical Analyses for Maizegenetic Experiments: Application of Biplots and Polar Plot to Line Tester Design. *Chilean Journal of Agricultural Research,* 76(3), July –September 2016, pp. 285-293.

Kundu, A., Drineas, P., and Magdon-Ismail, M. 2017. Recovering PCA and Sparse PCA via Hybrid-(l_1, l_2) Sparse Sampling of Data Elements. *Journal of Machine Learning Research*, 18 (2017), pp. 1-34.

Nieto, A.B., Galindo, M.P., Leiva, V., and Vicente-Galindo, P. 2014. A Methodology for Biplots Based on Bootstrapping with R. *Revista Colombiana de Estadística Current Topics in Statistical Graphics,* Diciembre 2014, volumen 37, no. 2, pp. 367-397. DOI: http://dx.doi.org/10.15446/rce.v37n2spe.47944.

Opeoluwa, F.O. and Sugnet, L 2017. Biplots in Covariance Analysis. *Biostat Biometrics Open Access Journal,* 3(4): BBOAJ.MS.ID.555523 (2017), pp. 1-8.

Paramita, S.A., Yamazaki, C., Setiawati, E.P., and Koyama, H. 2017. 2018. Distribution Trends of Indonesia's Health Care Resources in the Decentralization Era. *International Journal of Health Planning and Management*, March 2018, pp. 1-11. DOI: 10.1002/hpm.2506.

Qureshi, N.A., Suthar, V., Magsi, H., Sheikh, M.J., Pathan, M., and Qureshi, B. 2017. Application of Principal Component Analysis (PCA) to Medical Data. *Indian Journal of Science and Technology*, Vol 10(20), pp. 1-10. DOI: 10.17485/ijst/2017/v10i20/91294, May 2017.

Rad, M.R.N., M. Kadir, M.A., Rafii, M.Y., Jaafar, H.Z.E., Naghavi, M.R. and Ahmadi, F. 2013. Genotype Environment Interaction by AMMI and GGE Biplot Analysis in Three Consecutive Generations of Wheat (*Triticum Aestivum*) Under normal and Drought Stress Conditions. *Australian Journal of Crop Science* (AJCS), 7(7):956-961 (2013), pp. 956-961.

Sutarjo, U.S., Primadi, O., Yudianto, Budijanto, D., Hardhana, B., and Soenardi, T.A. 2014. *Indonesia Health Profile 2014*. Ministry of Health. General Secretariat Indonesia Health Profile 2014, Jakarta : Ministry of Health RI. 2015. ISBN 978-602-235-949-4.

Sutarjo, U.S., Primadi, O., Yudianto, Budijanto, D., Hardhana, B., and Soenardi, T.A. 2015. *Indonesia Health Profile 2015*. Ministry of Health. General Secretariat Indonesia Health Profile 2014, Jakarta : Ministry of Health RI. 2016. ISBN 978-602-235-911-1.

Ukalski, K. and Klisz, M. 2016. Application of GGE Biplot Graphs in Multi-Environment Trials on Selection of Forest Trees. *Folia Forestalia Polonica,* Series A – Forestry, 2016, Vol. 58 (4), pp. 228–239.