Neighbourhood Socio Economic Disadvantage Index’s Analysis of the Flood Disasters Area at East Jakarta in 1996 and 2016

Christina Ranti Ristiani*, Rokhmatusoh, Revi Hernina
Department of Geography, University of Indonesia, Kampus UI Depok 16424
christina.ranti@gmail.com

Abstract. Flood is one of natural disasters that have often happened in East Jakarta. Flood can give several negative impacts and it can affect all aspects of society lives such as economics, political, cultural, socials and others. East Jakarta is an urban area which continuously grows and establishes to become a rapid area. It can be seen from the highest population density in East Jakarta (BPS, 2016) and categorized into a region prone to flooding based on data Prone Flood Map in 1996 and 2016. The higher population exists in East Jakarta, the bigger possibility of the negative effects of disaster it gets. The negative impacts of flood disaster can affect societies especially with socio-economic disadvantage. One of the index to measure socio-economic disadvantage is NSDI (Neighbourhood socio-economic disadvantage index). However, to adjust indicators used in NSDI with Indonesia statistical data compatibility, it needs further assessment and evaluation. Therefore, this paper evaluates previous main indicators used in previous NSDI studies and improves with indicators which more suitable with statistical records in Indonesia. As a result, there will be improved 19 indicators to be used in NSDI, but the groups of indicators remain the same as previous namely; income, education, occupation, housing, and population.

Keywords: NSDI, flood, Jakarta

1. Introduction

Recently, NSDI has been used to analyse the relationship between the public urban green open space availability and neighbourhood socio-economic disadvantage [1]. When being applied in spatial context, the relationship between NSDI and other phenomenon can be assessed by using Geographically Weighted Regression (GWR). Furthermore, by using Principal Component Analysis (PCA) socio-economic indicators within initial loading can be removed if not correlated with dependent variables [1,2,3].

Therefore, this study will discuss indicators used in previous NSDI application, and evaluate similar indicators which are more suitable with statistical records in Indonesia. The analysis within this paper is purposed to adapt NSDI indicators as possible approach to measure social deprivation in communities located in areas where flood occurred in East Jakarta from 1996-2016.
2. Methods
Firstly, all possible socio-economic indicators exist in East Jakarta statistical records will be explored especially those reflect disadvantage such as number of employment and education level.

The data exploration mainly focus on East Jakarta statistical records published in East Jakarta Statistical Bureau website [4]. The next step is to list all possible indicators within the website and compare it with previously used indicators in NSDI (Table 1). If there are many indicators adaptable with previous studies, it will select only the same number of indicators within one group.

3. Systematic review and discussion
At first, Australia Bureau of Statistics (ABS) in 2011 developed an index that summarizes variables to illustrate relative disadvantage, this index is called Index of Relative Socio-Economic Disadvantage (IRSD) [3]. In form of vector maps, IRSD can be displayed to rank the regions from the most disadvantaged socio-economic condition to least disadvantaged regions.

If an area has a low score of IRSD, it means that this area has high proportion of relatively disadvantaged people, while high score of IRSD reflects that the area has low incident of disadvantage [3].

The dimension used by ABS to compute IRSD:
1. Income variables,
2. Education variables,
3. Employment variables,
4. Occupation variables,
5. Housing variables, and
6. Other miscellaneous indicators of relative advantage or disadvantage.

However, for some Asian countries the indicators used for every variable by ABS seems out of context such as % people who do not speak English well. These differences can be found when Li and Liu (2016) conducted research to observe the relationship between Urban Public Green Spaces (UPGSs) and Neighborhood Socioeconomic Disadvantage Index (NSDI) [1]. Li and Liu (2016) instead used variables from ABS, they used 14 socioeconomic indicators which was grouped into 5 groups to compute NSDI.

Therefore, this paper develops new indicators for socioeconomic disadvantages based on data availability and compatibility with East Jakarta Municipality condition. The comparison of socioeconomic disadvantage indicators can be observed from Table 1.

| Groups     | Indicators, Li and Liu (2016), 14 indicators in 5 groups | Data Compatibility and availability |
|------------|----------------------------------------------------------|-------------------------------------|
|            |                                                          | Indonesian Bureau of Statistics/ Government of Jakarta Province |
|            |                                                          | Proposed 19 indicators              |
| 1. Income  |                                                          |                                     |
|            | Percentage of household with income below the average (%) | Percentage of poverty (%)           |
|            | Percentage of household receiving subsistence allowance (%) | Percentage of household receiving subsidized rice (%) |
|            |                                                          | Percentage of people receiving subsidized health-care (%) |
| 2. Occupation |                                                      |                                     |
|            | Percentage of people living unemployment (%)            | Percentage of people living unemployment (%) |
|            | Percentage of people in temporary work (%)              | Percentage of job seeker (%)        |
| Percentage of blue collar worker (%) | - |
|--------------------------------------|---|
| Percentage of people working in agriculture (%) | - |

3. Education

| Illiteracy rate (%) | Percentage of people never go to school (%) |
|---------------------|-------------------------------------------|
| Percentage of people receiving no fundamental education (%) | Percentage of people with discontinued school (%) |
| - | Percentage of people with no certificate (%) |
| Percentage of people with degree below high school (%) | Percentage of people with primary school degree (%) |

4. Housing

| Percentage of household without kitchens or toilettes (%) | Percentage of household without kitchens or toilettes (%) |
|----------------------------------------------------------|----------------------------------------------------------|
| Percentage of household without clean energy (%) | Percentage of household without access to clean water (%) |
| Percentage of low-rent housing households (%) | Percentage of low-rent housing households (%) |

5. Population

| Proportion of floating population (%) | Percentage of homeless people (%) |
|--------------------------------------|-----------------------------------|
| Percentage of adult female living alone (%) | - |
| Percentage of older people living alone (%) | - |
| - | Percentage of birth assisted by non-medical (%) |
| - | Percentage of divorced people (%) |

In their research, Li and Liu (2016) has demonstrated that there is negative correlation among NSDI and UPGSs’ abundance, quality, and accessibility in most districts of Shanghai City, China. Interestingly, Li and Liu (2016) used Principal Component Analysis (PCA) method to analyze the most significant indicators should be involved in their study. PCA had already been used by ABS (2011) to remove non-significant indicators. The formula of NSDI used [1]:

\[
NSDI = \sum_{i=1}^{n} C_i \left( \sum_{j=1}^{r} L_j \cdot x_j \right)
\]

where \( C_i \) expressed the component I’s eigen value; \( L_j \) is indicator of \( j \)’s loading score; \( x_j \) is indicator \( j \)’s standardized value.

In this case to measure the area with disadvantage it will use the percentage of people living with poverty. The consideration behind this, because Indonesian Bureau of Statistics has measured several indicators to be used in poverty computation, hence no need to compute low dimension per capita income.
People living in poverty are the people whom its daily calorie consumption around 2100 kilocalorie or below, plus can only buy certain commodities in housing, cloth, education and health [4].

Since the growth of non-agricultural population proportion reflects rapid economic development [2], therefore in this study, we include the number of population which works in agriculture sector as indicator of disadvantage.

Since the Indonesian Bureau of Statistics has also counted the employed person whom in sick-leave as temporary unemployment, so in this study we do not include temporary unemployment as indicator of socio-economic deprivation, as suggested by previous research [1,2].

3.1. Principal Component Analysis
Principal component analysis (PCA) is a multivariate technique that analyses a data table in which observations are described by several inter-correlated quantitative dependent variables.

The goals of PCA are to (a) extract the most important information from the data table, (b) compress the size of the data set by keeping only this important information, (c) simplify the description of the data set, and (d) analyze the structure of the observations and the variables [5].

PCA works by ranking the component in the variance-covariance which has the highest percentage of variance (\(\%p\)) among other components. In the multivariate statistics, variance of each variable/factor can be computed by using:

\[
\text{var}_j = \frac{\sum_{i=1}^{n}(x_{ij}-\mu_j)^2}{n-1}
\]

In PCA, variance and covariance can be formed to be a single matrix consists of eigenvalue (\(\lambda_{pn}\)) and eigenfactors. Basically, eigenvalue in j column is the variance, therefore the percentage variance of every column can be computed by using:

\[
\%p = \frac{\lambda_{pn} \cdot 100}{\sum \lambda_{pn}}
\]

where \(\%p\) is the percentage variance in component j, \(\lambda_{pn}\) is variance in component j, and \(\sum \lambda_{pn}\) is total variance [6].

To maintain the factor/variable, it can be observed from the factor loading value, which expresses the correlation (\(R_{kp}\)) between variable \(k\) with component \(p\). The formula to calculate correlation between variable \(k\) with component \(p\) is:

\[
R_{kp} = \frac{a_{kp} \sqrt{\lambda_{pn}}}{\sqrt{\text{var}_k}}
\]

where \(a_{kp}\) = eigenvector (factor loading) for variable \(k\), principal component \(p\); \(\lambda_{pn} = p\)-th eigenvalue (component); \(\text{var}_k\) = variance of variable \(k\) (diagonal values in the covariance matrix) (Jensen,1996).

In previous study, the socioeconomic indicators with factor loading more than 0.75 was selected, and removed the factor with value less than 0.75 [1].

3.2. Proposes framework
Therefore, based on aforementioned systematic review and discussion, it would be best for implementation of NSDI to replace earlier indicators used by Li and Liu (2016) with newly 19 indicators (Table1).
Figure 1. Proposed framework for NSDI computation in flood affected villages.
4. Conclusion
Based on systematic review and discussion, proposed framework (Figure 1) to map villages which was affected by flood events from 1996-2016 in East Jakarta Municipality possible to apply with the condition that all proposed data are available.

References
[1] Li H and Liu Y 2016 Neighborhood socioeconomic disadvantage and urban public green spaces availability: A localized modeling approach to inform land use policy, *Land Use Policy*, 57, p 470
[2] Su S, Gong Y, Tan B, Pi J, Weng M, and Cai Z 2017 Area Social Deprivation and Public Health: Analyzing the Spatial Non-stationary Associations Using Geographically Weighed Regression, *Social Indicators Research*, 133, p 819
[3] Australian Bureau of statistics 2011 Technical Paper Socio-Economic Indexes for Areas (SEIFA) http://www.ausstats.abs.gov.au/ausstats/subscriber.nsf/0/22CEDA8038AF7A0DCA257B3B00116E34/$File/2033.0.55.001%20seifa%202011%20technical%20paper.pdf
[4] Indonesia Statistical Bureau 2017 https://www.bps.go.id/Subjek/view/id/23
[5] AbdiH and Williams LJ 2010 Principal component analysis, *WIREs Comp Stat*, 2, p 433
[6] Jensen J. R. 2005 *Introductory Digital Image Processing: A Remote Sensing Perspective* (Prentice Hall)