The Relationship between Fertilizer Using and Children Undernutrition in India

GIS & Spatial Data Analysis

Shaoqian Zhang (A53273384)

December 13, 2019

1 Background

Malnutrition (undernutrition and overnutrition) is an important indicator to represent a country's health (Rumi 2007). India has the highest recorded numbers of undernourished in the world. These malnourished children are at increased risk for death, disease and infection, stunted growth, cognitive deficits, poor academic performance and short school attendance.[1]

Undernutrition of children are usually represented by three nutrition status: stunted, wasted, and underweight for age. More than one third (57.9 million) of the world's stunted children who are under the age of five live in India, and about 48 percent of India's children under 5 are stunted[2]. In India, data from National Family Health Surveys (NFHS) 2015-16 shows that 38 percent of children below five years are stunted, and 36 percent are underweight (Rumi 2007).[3] Therefore, reduction in undernutrition is a major task in India recent years.

India's government applies different policy measures to reduce malnutrition. Among those policy measures, increasing production of food grains is necessary. India has developed many ways to improve food production in the "green revolution" (Potter 2010), which has increased crop yields many times by using new technologies, including new crop varieties, irrigation and fertilization (Borlaug 2007; Tilman 1998).

Therefore, chemical fertilizer application plays an important role in agriculture. And most importantly, the "green revolution" changed the malnutrition status in India. During 1992–2016, the prevalence of underweight among children had declined from 53% to 36%, and stunting had declined from 52% to 38% (Khan 2018).

However, chemical fertilizer will have negative effects on children's health through different ways. First, as we know, fertilizer is used to ensure food security, so it will be related to human's daily food. The main component of chemical fertilizer is nitrogen (N) and phosphorus (P). Excessive N fertilizer inputs will cause infection of diseases and pests, greedy green and late maturity for plants. Excessive P fertilizer using will weaken the effectiveness of soil micronutrients, make the nutritional elements needed by the crops out of balance and even lead to a deficiency of micronutrients, as a result, will affect the crop yield and quality. If children consume the low-quality crops, vegetables and other plant products for a long time, they may get affected for some nutritional endemic diseases (Mascie 192). Second, fertilizer application will influence the water and soil quality, so it indirectly influences child health.

[1] India Nutrition Profile. (2019, November 21). Retrieved from https://globalnutritionreport.org/nutrition-profiles/asia/southern-asia/india/.

[2] India Nutrition Profile. (2019, November 21).

[3] According to Rumi 2007, data are generated by author.

Nutrients applied to farmlands can leach into aquatic systems and alter ecosystem functions (Carpenter et al.1998; Smil 2000). Excessive fertilizer consuming may cause disturbances of the kidneys, lungs and liver and even cause cancer (Potter 2010).

Therefore, I try to figure out the relationship between N and P fertilizer using on child health, especially children undernutrition in this paper. My hypothesis is that as fertilizer application rate (average application per crop) increases, percentage of children stunting and underweight will goes down, since more and more children could intake enough nutrition. However, after excessive usage of chemical fertilizers, these two children's health indicator will go down because the negative effects of chemical fertilizer I mentioned before. I use children stunted rate and children underweight rate at district level as two health indicators (dependent variables) to reflect the impact on children's height-for-age and weight-for-age. I use spatially explicit fertilizer inputs of nitrogen (N) and phosphorus (P) in India as independent variables which fuse statistics on fertilizer use with harvested area for 175 crops (Potter 2010). To avoid omitted variables biases, I choose distance to river as a control variable to verify that the distance from river may influence the fertilizer application and stunting rate or underweight rate. Finally, I run the quadratic regression of N and P fertilizer application rate on children stunted rate and children underweight rate separately.

This paper proceeds as follows. Part 2 explains the data sources and methods to analysis data. Part 3 shows the results. Part 4 draws the conclusion and recommendation.

2 Data and methods

2.1 Data Sources

Three indicators of nutritional status of children always includes stunting (height-for-age), underweight (weight-for-age) and wasted (weight-for-height)[4], I only choose stunting and underweight as the dependent variables, undernutrition, in this project. The dependent variables are calculated from National Family Health Survey round four (*NFHS-4, 2015-16*) by *STATA 16*. *NFHS 4* provides *STATA* data (*IAKR74FL.dta*) on a variety of child nutrition variables for the states and districts of India. It covers 259627 units, 640 districts of India. In order to calculate the underweight rate and stunted rate under age of five, I use the unit data, which could be requested on Demographic Health Survey (DHS) data repository[5], for each child's age in months, height for age standard deviation, weight for age standard deviation of new WHO standard. My analysis is based upon 259627 cases weighted to 640 districts.

For independent variables, I use the two major fertilizers (N and P) application rate data in 2002 which are taken from Potter's paper in 2010. The author obtained national-level fertilizer application rates of two major fertilizers (N and P) for crops found in 88 countries from the International Fertilizer Industry Association (IFA). The unit of analysis is average application per crop (kg ha^{-1} yr^{-1}), including both fertilized and unfertilized crop area.

In order to control the distance from river, I use the river distribution geographic data form World-wide Hydrogeological Mapping and Assessment Programme (WHYMAP) of BGR Hannover / UNESCO of Paris 2008[6].

[4] UNICF. unstats | Millennium Indicators. Retrieved from https://millenniumindicators.un.org/unsd/mdg/Metadata.aspx?IndicatorId=0&SeriesId=560.

[5] Database available by request from DHS :https://www.dhsprogram.com/data/

[6] The resources could be found on BGR-WHYMAP website:

The geographic data of India's three level of boundaries come from GADM data website[7]. I *project* the district-level boundary shapefile under *WGS_1984* coordinate system and display the XY geographical coordinates. All the following analysis in *ArcGIS 10.7.1* is under this projection.

2.2 Data Manipulation

2.2.1 Dependent Variables

To calculate the percentage of children under five that are underweight/stunted, I use the following function:

$$\text{Percentage of children under five that are underweight/stunted} = \frac{\text{Number of children under age five that fall below minus three standard deviations from the median weight/ height for age of the NCHS/WHO standard}}{\text{Total number of children under age five that were weighed/ heighted}} * 100$$

For numerators, according to the *Guide to DHS Statistics DHS-7*,[8] the number of children whose height-for-age z-score is under minus 3 (-3.0) standard deviations (SD) below the mean on the WHO Child Growth Standards is defined as severely stunted. Also, the number of children whose weight-for-age z-score is below minus 3 (-3.0) standard deviations (SD) below the mean on the WHO Child Growth Standards (*hw71 < -300*) is defined as severely underweight.

For denominator, to get the number of children under age five that were weighed or heighted, I use the number of living children between ages 0 and 59 months before the survey who have valid non-flagged height for age z-scores (*hw70 < 9990*) for stunting, and valid non-flagged weight for age z-scores (*hw71 < 9990*) for underweight.

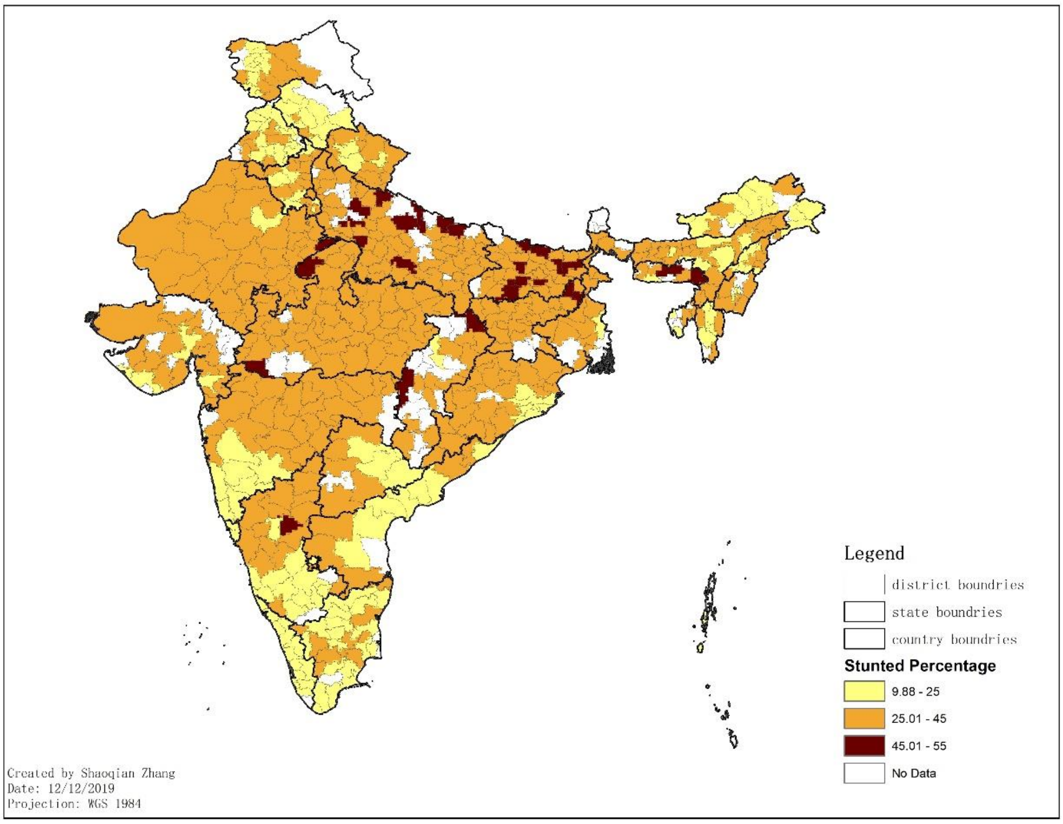
After I tabulate all the statistics of the stunted and underweight percentage in each district, I join this table with the boundary shapefile performed in *ArcGIS 10.7.1*. Then, I use *Feature to Raster* to get *Figure 1 (a) and (b)*, which maps the geographic distribution of children stunted and children underweight. Missing values are generated as *No Data* because the districts in the district level boundary and health indicator data does not perfectly match. Specifically, for both health indicator, I define percentage higher than 45% as high stunting/underweight, percentage between 25% to 45% as medium stunting/underweight, and less than 25% as low stunting/underweight. From *Figure 1*, I find that both high stunted rate area and high underweight rate are concentrated in the middle-northwest area of India. And the most part of India is under medium stunting and underweight.

---

https://www.whymap.org/whymap/EN/Maps_Data/Gwr/gwr_node_en.html

[7] Database could be downloaded from GADM data website: https://gadm.org/data.html.

[8] In this analysis, I use the BR file (IAKR74FL.dta) downloaded from DHS. The analysis steps could be found on: https://dhsprogram.com/data/Guide-to-DHS-Statistics/index.htm#t=Nutritional_Status.htm

(a)


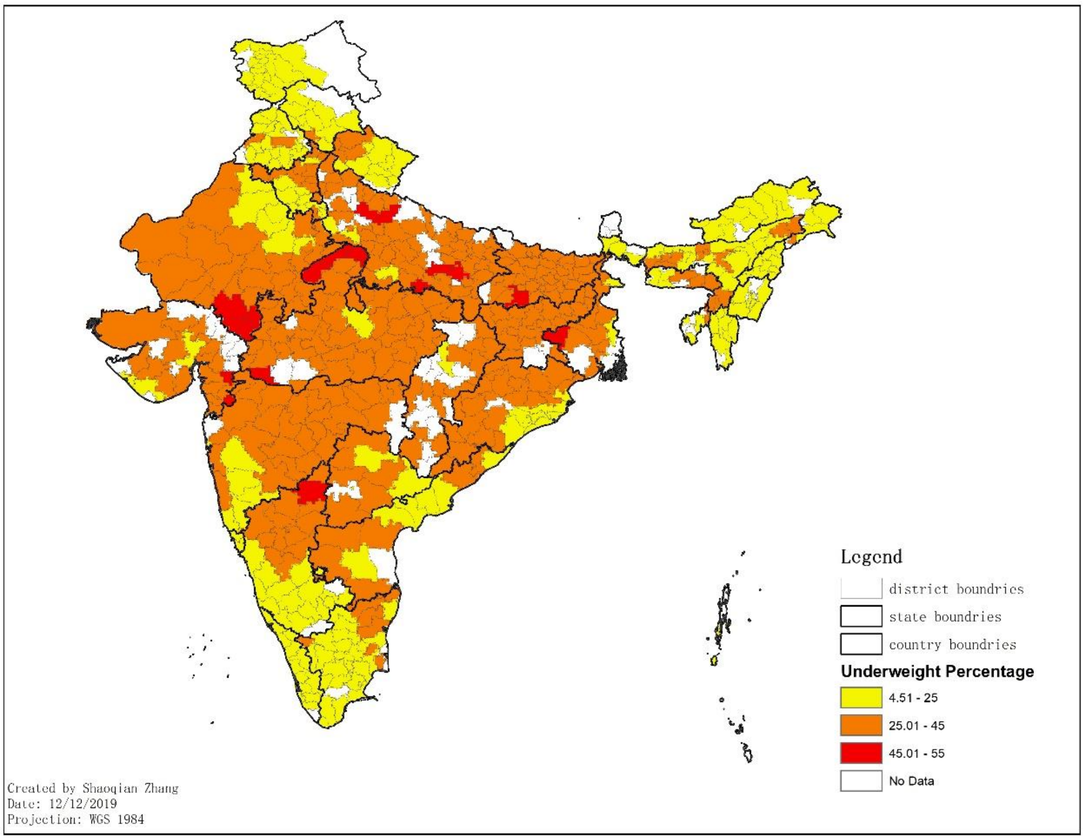
(b)


Figure 1 Maps of India showing the geographic distribution of the rates of (a) stunting (b) underweight across districts of India, 2015. Source: Authors generated the maps using ArcGIS version 10.7.1

2.2.2 Independent Variables

I use the raster results of N and P fertilizer application rates from Potter (2010), and average the fertilizer using by districts using *Zonal Statistics* tool in *ArcGIS 10.7.1*, showing in *Figure 2. (a), (b)*. From *Figure 2*, I find that the highest N fertilizer application rate is in the northwest part (Punjab and north part of Haryana) and northeast corner, which is higher than 81 kg ha^{-1}(*Figure 2a*). The highest P fertilizer application rate is located in the northeast corner (north part of Assam), which is higher than 81 kg ha^{-1}(*Figure 2b*).

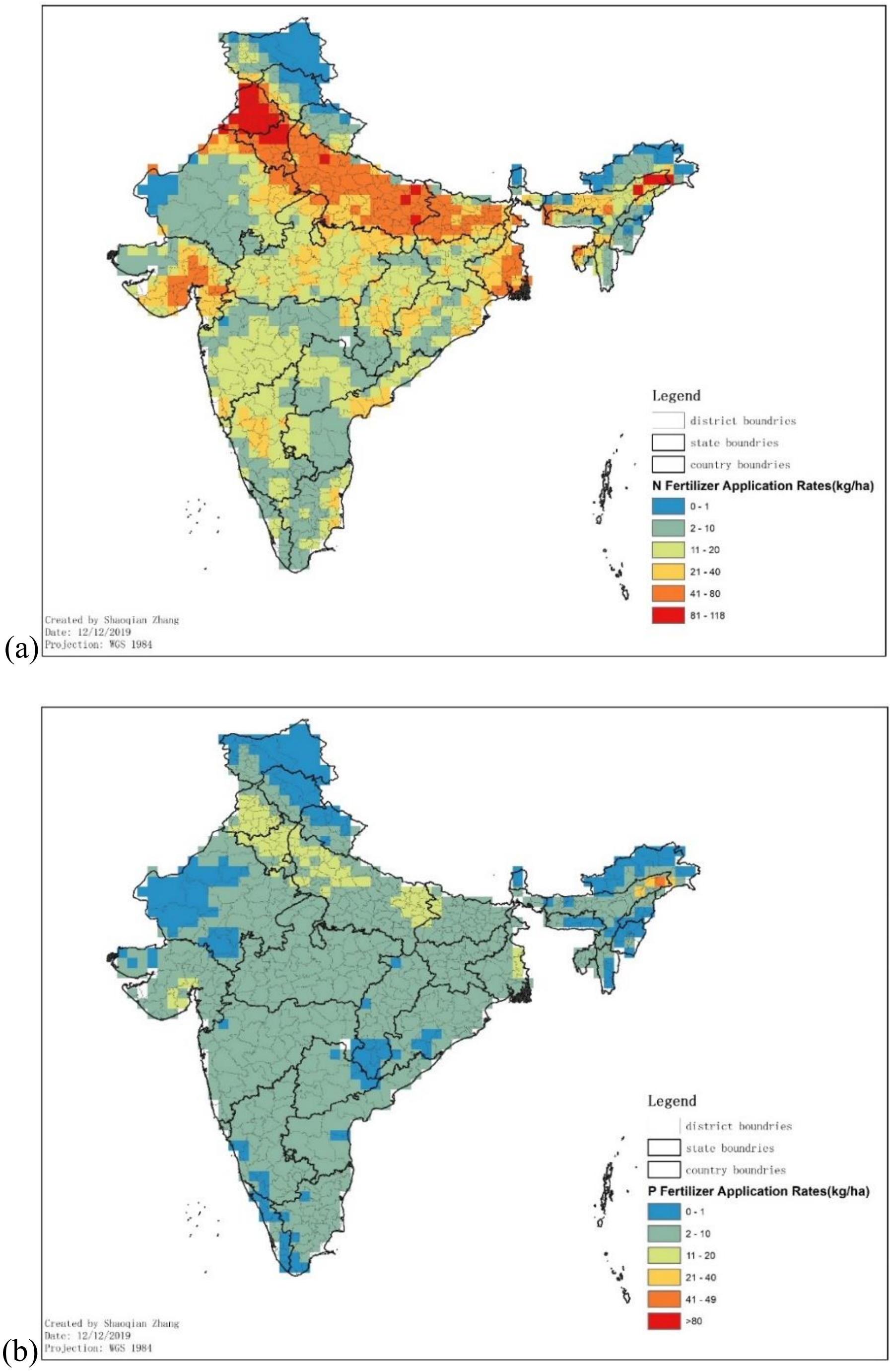


Figure 2. (a) India map of N fertilizer application rates (kg/ha of grid cell area). (b) India map of P fertilizer application rates (kg/ha of grid cell area). Both values represent the average fertilizer inputs in every 0.5x 0.5 grid cell (roughly 55x55 km).

2.2.3 Control Variables

Considering the omitted variable bias, I try to rule out other factors which may influence the causal relationship between fertilizer using and stunting or underweight rate of children. The reason why I choose this distance to river as control variables is because I want to rule out the effects of distance from the water source on children's health indicator. For example, with close access to river, agricultural land will be more concentrated, and crop production will be more abundant, and may result in the higher reduction in children stunted rate and children underweight rate. Therefore, I use the river distribution in India to perform the impact of the distance to water source.

The distribution of river in India is mapped in *ArcGIS 10.7.1* showing in the *Figure 3*. First, I use *Polyline to Raster* tool to generate the raster of river distribution with the mask of India boundary. Then I use *Reclassify* tool to classify all rivers into one category. In order to calculate the distance to river, I use *Euclidean Distance* tool to calculate the distance from any point of the map to the raster of river, and display the results in *Figure 3*. I notice that most places with high fertilizer inputs are geolocated within 39km of the watershed.

In order to perform regression analysis, I use *Zonal Statistics as Table* tool to generate all the variables into table at district lever. I join tables together and use *Table to Excel*. The regression analysis is explained in Section 3.

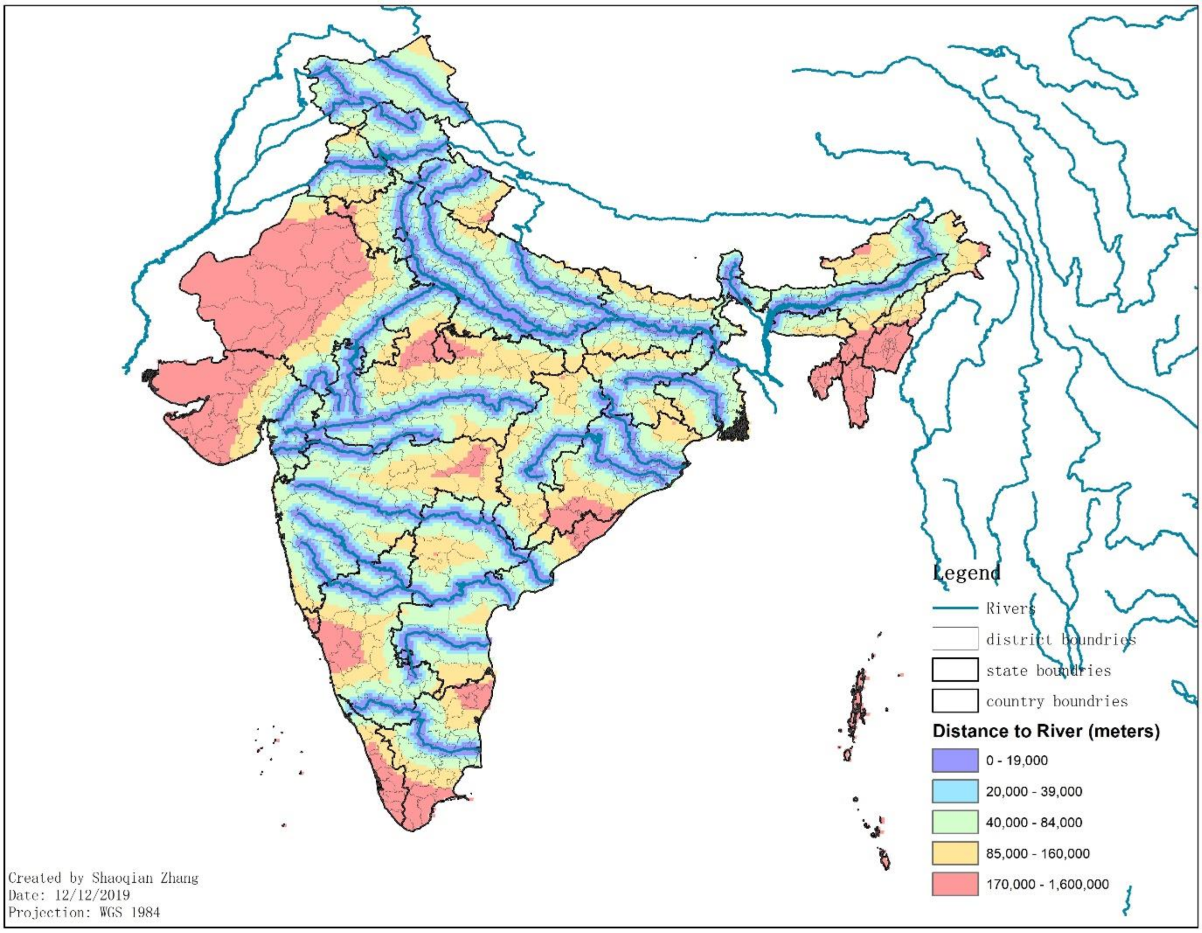


Figure 3. Distance to any river. The value performs in every 14000 x 14000 m as cell size.

3 Results

First, I generate the non-linear relationship between variables using Local polynomial smooth plots with CIs in *STATA*, without control variables. *Lpolyci* could evaluate all the non-missing smoothed value and performs a local regression fit at each specified grid point.[9]

The results show in *Figure 4*. From *Figure 4*, I find that there is common regular pattern that as fertilizer application rate (independent variables) goes up, both of the percentage of children stunted and children underweight (dependent variables) increases at first, and when the independent variable exceeds a certain value, the dependent variable starts to decrease. Therefore, this pattern confirms the hypothesis I suggest before, that fertilizer using may reflects the positive effects for child health but turns to be negative impacts if it's overused.

In addition, from *Graph (a) and (b)*, the turning point of N fertilizer application is at around 65 kg/ha for children stunted, and 45 kg/ha for children underweight; from *Graph (c) and (d)*, the turning point of P fertilizer application is at around 10 kg/ha for children stunted and children underweight. Therefore, the implementation of P fertilizer will accelerate the advent of negative effects than N fertilizer.

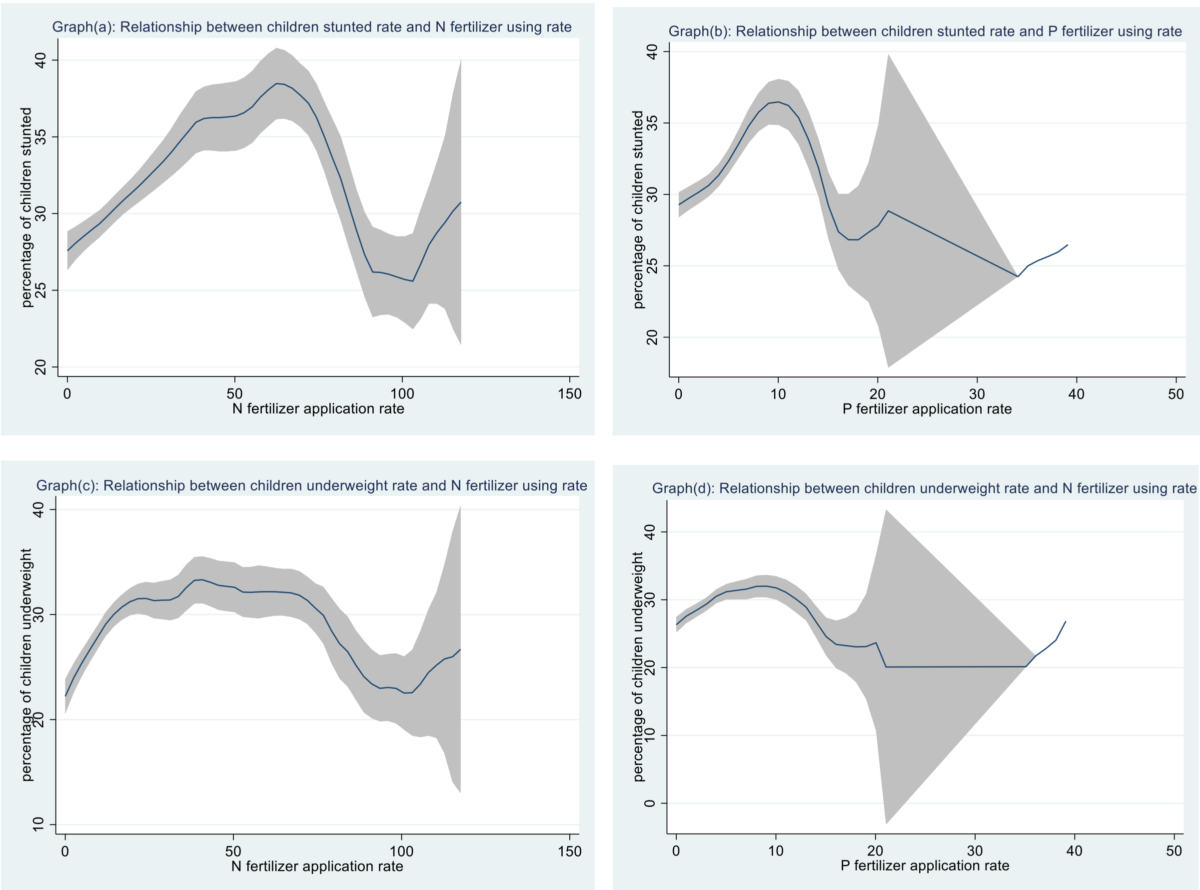


Figure 4: *Graph(a)*: nonlinear relationship between children stunted rate and N fertilizer using rate; *Graph(b)*: nonlinear relationship between children stunted rate and P fertilizer using rate; *Graph(c):* nonlinear relationship between children underweight rate and N fertilizer using rate; *Graph(d)*: nonlinear relationship between children underweight rate and N fertilizer using rate.

[9] https://www.stata.com/manuals13/g-2graphtwowaylpolyci.pdf

Therefore, I set up linear and quadratic relationship functions between variables with consideration of control variables (distance to river), where *i* represents the number of districts:

Linear regression:

$$\text{Percentage } \widehat{\text{of chi}}\text{ldren stunted}_{\mathrm{i}} \left(\text{percentage o}\widehat{\text{f chi}}\text{ldren stunted}_{\mathrm{i}}\right)$$

$$= \beta_0 + \beta_1 \text{ N or P fertili}\widehat{\text{zer app}}\text{lication rate}_{\iota} + \beta_3 \text{ Distance to river}_i + \varepsilon_i$$

Quadratic regression:

$$\text{Percentage o}\widehat{\text{f chi}}\text{ldren stunted}_{\mathrm{i}} \left(\text{percentage o}\widehat{\text{f chi}}\text{ldren stunted}_{\mathrm{i}}\right)$$

$$= \beta_0 + \beta_1 \text{ N or P fertili}\widehat{\text{zer app}}\text{lication rate}_{\iota} + \beta_2 \text{ N or P fertili}\widehat{\text{zer app}}\text{lication rate}_{\iota}^2$$

$$+ \beta_3 \text{ Distance to river}_i + \varepsilon_i$$

First, I conduct the linear regression and the results present in *Appendix Table A*. The coefficient of β_1 for both N and P fertilizer application is positive.

Then, I run the non linear regression presenting in the *Table 1*. After I include the quadratic term into the function, the coefficient of quadratic N fertilizer application rate in the *column (1)* on children stunted rate is -0.00388 with 95% statistically significance, which means that a 1 unit increase in N fertilizer application rate results in a 0.388% point decrease in the percentage of children stunted; also, β_2=-0.00444 in the *column (2)* means that a 1 unit increase in N fertilizer application rate results in a 0.444% point decrease in the percentage of children underweight, with 95% statistically significance.

The coefficient of quadratic P fertilizer application rate in the *column (3)* on children stunted rate is -0.0218 with 95% statistically significance, which means that a 1 unit increase in N fertilizer application rate results in a 2.18% point decrease in the percentage of children stunted; also, β_2= -0.0184 in the *column (4)* means that a 1 unit increase in P fertilizer application rate results in a 1.84% point decrease in the percentage of children underweight, with 95% statistically significance.

The coefficient of *Distance to River* represents that as the distance form river increases, the percentage of children malnutrition decrease slightly. This coefficient is significant between the N fertilizer application on children underweight and P application on both underweight and stunted children.

In all, there is a negative non-linear relationship between fertilizer usage rate and percentage of children stunted and children underweight. The impact of P fertilizer usage is stronger than P fertilizer usage. The further distance from river, the less likely the stunting and underweight happen.

Table 1: the quadratic relationship between fertilizer application rate and percentage of children stunted or children underweight				
	(1) Stunted Percentage	(2)Underweight Percentage	(1) Stunted Percentage	(2)Underweight Percentage
N Fertilizer application rate	0.415***	0.432***		
	(0.0485)	(0.0578)		
N Fertilizer Application Rate^2	-0.00388***	-0.00444***		
	(0.000527)	(0.000648)		
Distance to River	-0.00000751	-0.0000170**	-0.0000129**	-0.0000230***
	(0.00000460)	(0.00000567)	(0.00000450)	(0.00000558)
P fertilizer Application Rate			0.821***	0.602**
			(0.169)	(0.185)
P Fertilizer Application Rate^2			-0.0218***	-0.0184**
			(0.00639)	(0.00686)
Constant	26.21***	24.65***	29.74***	28.77***
	(0.896)	(1.110)	(0.793)	(0.977)
Observations	533	533	533	533
Adjusted R^2	0.174	0.151	0.086	0.066

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

4 Conclusion

In India, fertilizer application is obviously an essential part to improve crop production in agriculture, and agriculture improvement plays an important role on tackling malnutrition problem. The main results of this paper verify the hypothesis that there is a quadratic or non-linear relationship between N or P fertilizer application rate and malnutrition status, which is represented as stunting and underweight rate of children under 5. As N or P fertilizer application rate increases, the children stunted percentage and children underweight percentage will increases as well, until a turning point where fertilizer application exceed a certain amount, the percentage of children malnutrition falls down. It shows an overall slightly negative relationship between fertilizer usage and children malnutrition. The greater the distance from the river, the less likely the stunting and underweight happen. In addition, from the regression results, P fertilization application has more obvious negative relationship on children malnutrition than N.

Facing this pattern, I suggest that fertilizer using should be more cautious and strictly controlled in India. Even if fertilizer using could help to improve food production and reduce the malnutrition, we could find other efficient ways. For instance, first, we could strengthen and improve technical measures in optimizing the formula of fertilization, like adjusting fertilizer composition according to crop type and growing environment. Second, implementing environmental condition analysis on fertilizer application, such as analysis of water quality, harmful substance content in soil, etc. Third, increasing the proportion of organic fertilizer in fertilizer using. Using micro-fertilizer and biological fertilizer is also a good way to improve the production, at the same time, do no harm to environment and human's health.

5 Reference

Rumi Ailjaz, 2007: Preventing Hunger and Malnutrition in India. ORF. https://www.orfonline.org/research/preventing-hunger-and-malnutrition-in-india/.

India Nutrition Profile. Retrieved from https://globalnutritionreport.org/nutrition-profiles/asia/southern-asia/india/.

Mascie-Taylor, C. G. N. (1992). Endemic disease, nutrition and fertility in developing countries. *Journal of Biosocial Science*, *24*(3), 355–365. doi: 10.1017/s002193200001991x

Khan, J., & Mohanty, S. K. (2018). Spatial heterogeneity and correlates of child malnutrition in districts of India. *BMC Public Health*, *18*(1). doi: 10.1186/s12889-018-5873-z

Potter, P., Ramankutty, N., Bennett, E. M., & Donner, S. D. (2010). Characterizing the Spatial Patterns of Global Fertilizer Application and Manure Production. *Earth Interactions*, *14*(2), 1–22. doi: 10.1175/2009ei288.1

Borlaug, N., 2007: Feeding a hungry world. Science, 318, 359, doi:10.1126/science.1151062.

Tilman, D., 1998: The greening of the green revolution. Nature, 396, 211–212.

Carpenter, S. R., N. F. Caraco, D. L. Correll, R. W. Howarth, A. N. Sharpley, and V. H. Smith, 1998: Nonpoint pollution of surface waters with phosphorus and nitrogen. Ecol. Appl., 8, 559– 568.

Smil, V., 2000: Phosphorus in the environment: Natural flows and human interferences. Annu. Rev. Energy Environ., 25, 53–88.

Brown, D. (n.d.). unstats | Millennium Indicators. Retrieved from https://millenniumindicators.un.org/unsd/mdg/Metadata.aspx?IndicatorId=0&SeriesId=560.

Guide to DHS Statistics (English). (n.d.). Retrieved from https://www.dhsprogram.com/publications/publication-dhsg1-dhs-questionnaires-and-manuals.cfm.

Appendix:

Table A: the linear relationship between fertilizer application rate and percentage of children stunted or children underweight

	(1) Stunted percentage	(2)Underweight percentage	(3) Stunted percentage	(3) Underweight percentage
N fertilizer application rate	0.0692***	0.0381*		
	(0.0173)	(0.0182)		
Distance to river	-0.0000141**	-0.0000244***	-0.0000164***	-0.0000259***
	(0.00000445)	(0.00000562)	(0.00000444)	(0.00000554)
P fertilizer application rate			0.245*	0.115
			(0.104)	(0.0957)
Constant	30.76***	29.79***	31.65***	30.39***
	(0.710)	(0.876)	(0.717)	(0.820)
Observations	533	533	533	533
Adjusted R^2	0.068	0.053	0.047	0.047

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$