Interactive Nutrition Maps of Paraguay 2012

Matteo Cellamare,[a] Vit Bubak,[b] John Newman[c]

[a] Instituto Desarrollo. Asunción, Paraguay.

[b] Instituto Desarrollo. Asunción, Paraguay.

[c] Kenact, EE.UU.

Interactive Nutrition Maps of Paraguay 2012

Matteo Cellamare[*], Vit Bubak[†] & John Newman[‡]

November 27, 2018

In this work We develop a web application in order to provide a first tool to monitor the nutritional situation in Paraguay and represents one of the final products of the research project entitled *La transicin nutricional en Paraguay:En dónde estamos.*

The app contains several maps in which the user can explore different aspects of the nutritional status at the department and district level in Paraguay. Maps included illustrate daily calorie intake for the entire population and the prevalence of malnutrition of children under five years of age.

We consider two measures of child malnutrition based on body measurements and which are not subject to seasonal variation for small area estimation: stunting (low height-for-age) and underweight/overweight (low/high weight-for-age). To produce reliable estimates at the national level we apply the Small Area Estimation (SAE) technique on data from the 2012 Census (*Censo Nacional de Poblacion y Viviendas 2012*) and 2011-2012 Survey of Household Income and Expenditure (*Encuesta de Ingresos y Gastos y de Condiciones de Vida 2011-2012*).

[*]Instituto Desarrollo, Asunción, Paraguay. Email: mcellamare@desarrollo.edu.py

[†]Instituto Desarrollo, Asunción, Paraguay. Email: vbubak@desarrollo.edu.py

[‡]The World Bank, Washington DC, US. Email: jnewman@kenact.com

The authors acknowledge financial support from the Paraguayan National Council for Science and Technology (CONACYT), project PINV15-1304.

Small Area Estiamtion models

Small area estimation (SAE) is a statistical technique that models data collected from one or more data sources, to produce more accurate estimates at small area level than using only data collected from each small area.

The improvement that SAE provides to the estimates is due to the fact that for the estimation of a particular small area it uses information from areas with similar characteristics borrowing strength to the estimation accuracy.

The idea is to combine individual and household (micro) survey data and population (macro) census data with the objective of estimating welfare indicators for specific geographic area as small as district or towns. A statistical model is fitted to survey data collected around the same time as the census, and this model is used to predict a variable not collected in the census, based on variables that are collected in both survey and census.

The underlying model used in SAE is a unit-level mixed model, also known as *nested error linear regression model*, introduce by Battese [1] and extended by Rao and Molina [6] that includes a random area/domain-specific effect and a unit (household or individual)-level error term.

Therefore, let us define by n_i the number of unit sampled from the area $i = 1, \ldots, m$ (department, district,...) and let x_{ij} be the p-dimensional auxiliary values for the unit $j = 1, \ldots, n_i$ in area i. The nested error linear regression model is defined by

$$y_{ij} = x'_{ij}\beta + u_i + e_{ij} \qquad u_i \sim N(0, \sigma_u^2), \quad e_i j \sim N(0, \sigma_e^2)$$

where y_{ij} is the value of the variable of interest for the unit j sampled in are i, β the p-vector of unknown fixed effects, u_i is the area specific random effect and e_{ij} is the individual level random error.

Several approach has been proposed in literature to estimate the model, including the empirical best predictor (EBP) [5], the World Bank method [3] and the M-Quantile method

[2, 7].

We use the EBP approach in which at least two data sources are required. The sample data set, used to fit the nested error linear regression model and a population data set (usually census) used for predicting synthetic values of the outcome for the entire population. The main assumption is that both data sources must share identically defined auxiliary variables but the variable of interest is only present in the sample data set. Note that, measuring the uncertainty of the EBP estimates is done by using bootstrap methods. Theoretical and technical details on the estimation methodology can be find in [5].

The resulting estimates can be illustrated by the use of maps that make interpretation simpler. However, the central point is not the map itself, but that indexes of interested can be assessed at a much finer level at a much lower cost than by increasing the sample size of the survey or rerunning the census with additional variables.

Malnutrition Measures

In this work we focus on two main aspect of the nutrition in Paraguay.

As first, we focus on estimate the daily calories intake per capita of the entire population using household as unit-level. We follow the guidelines provide by FAO in [4] and we collect all the information on food items buy from each household (in sections Gastos fuera del hogar and Gastos del hogar of the Survey).

For each food item, we re-scale its amount on daily basis consumption and we calculate the daily calories intake using the conversion table provided by INCAP [1].

Finally, the Food calories intake per capita of each household can be obtain by summing the food basket daily calories intake and dividing it by the household size.

In the second part of this work, we consider two central measures of undernutrition, both based on measurements of a childs height, weight and age. Differently from the previous case, here our unit-level are child with less than 5 years old.

[1] link to **Tabla de composicion de alimento de Centroamerica** of *INCAP*

Stunting or low height-for-age is defined as having a height at least two standard deviations below the median height for a reference population. Underweight or low weight-forage is similarly defined while overweight is defined as as having a weight at least two standard deviations above the median height for a reference population.

The data used as a reference standard in these definitions was established in 1975 by the National Center for Health Statistics / Centers for Disease Control in the USA and updated by WHO in 2006.

Within a particular area, stunting is defined as the proportion of such children with a standardized height-for-age (HAZ) value below 2:

$$E(S_d) = \frac{1}{B}\sum_b \left[\frac{\sum_{i\in d}(HAZ_{id}^b < -2)}{N_d}\right]$$

Similarly underweight is the proportion with a standardized weight-for-age (WAZ) value below -2,

$$E(U_d) = \frac{1}{B}\sum_b \left[\frac{\sum_{i\in d}(WAZ_{id}^b < -2)}{N_d}\right]$$

and overweight

$$E(U_d) = \frac{1}{B}\sum_b \left[\frac{\sum_{i\in d}(WAZ_{id}^b > 2)}{N_d}\right]$$

In all cases above, N_d is the number of children in district d and $b = 1, ..., B$ bootstrap estimates.

Data source and variables selection

As stated before, we consider data from the 2012 Census (*Censo Nacional de Poblacion y Viviendas 2012*) and 2011-2012 Survey of Household Income and Expenditure (*Encuesta de Ingresos y Gastos y de Condiciones de Vida 2011-2012*) provided by the *Dirección General de Estadistica, Encuestas y Censos* (DGEEC) of Paraguay.

As noted earlier, it is important that any auxiliary variables used in modeling and

predicting should be comparable in the estimation (survey) data set and the prediction (census) data set. Based on their significance and comparability we used the following variables

Household level	**Head of household level**	**Children level** $age < 5$
Daily Energy Intake per capita	Age	Weight-for-age Z (WAZ)
department	Sex	Height-for-age Z (WAZ)
district	Marital status	Age
Household size	Years of Education	Sex
Prop. of Females	Employment status	
Prop. of children $age < 6$		
Prop. of children $6 < age < 14$		
Prop. of adult $age > 65$		
Female with at least a 2^{nd} degree		
Electricity connection		
Unsatisfied Basic Need: Water		
Unsatisfied Basic Need: House		

Note that, same availability for Census microdata except for the variables of interest (underlined in the table).

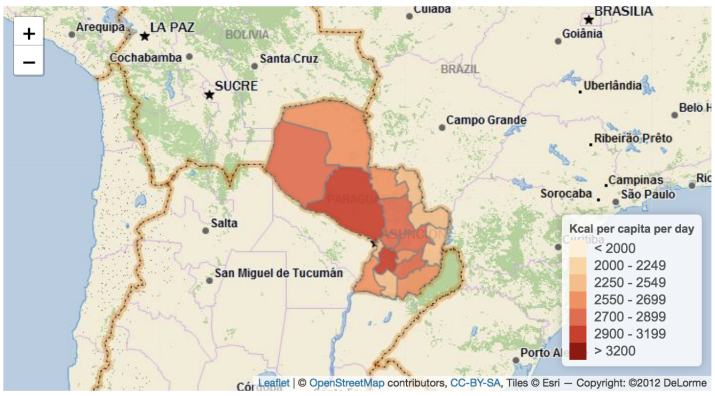
Interactive Nutrition Maps of Paraguay

In order to show the results of our study, we develop a web application that illustrates different aspects of the nutritional status at the department and district level in Paraguay. Maps included illustrate daily calorie intake for the entire population and the prevalence of malnutrition of children under five years of age.

Figures and show how the web application displays the results of this study, in particular referring to Daily calories intake at department and district level. The user can access detailed information by clicking on a specific department/district or in the table positioned under the map.

To see all results please visit https://instituto-desarrollo.shinyapps.io/nutrition_paraguay/

Monitoring nutrition in Paraguay (2012)



Choose a dataset: Daily Calorie Intake

Method
- Direct Estimation
- Small Area Estimation

Level of disaggregation
- Department
- District

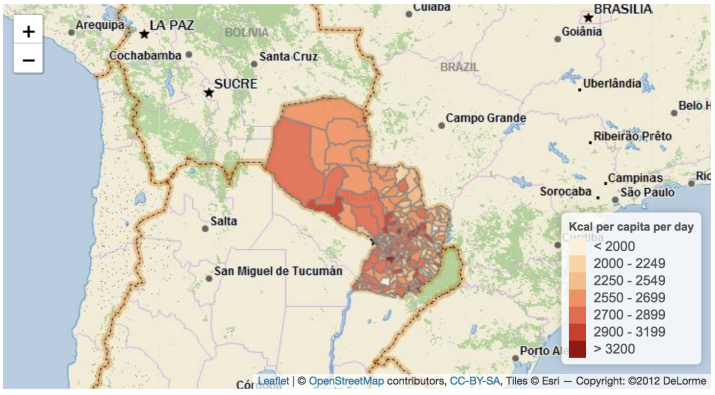
Threshold

Show entries — Search:

	Department	people	hogares	Est. Kcal per capita
1	Ñeembucú	69757	21854	2633
2	Alto Paraguay	10249	2860	2602
3	Alto Paraná	499748	128718	2324
4	Amambay	95000	24125	2287
5	Asunción	375463	98620	2261

Showing 1 to 5 of 18 entries — Previous 1 2 3 4 Next

Monitoring nutrition in Paraguay (2012)



Choose a dataset: Daily Calorie Intake

Method
- Direct Estimation
- Small Area Estimation

Level of disaggregation
- Department
- District

Threshold

Show entries — Search:

	District	Department	Population	Households	Est. Kcal per capita
1	Pedro juan caballero	Amambay	68375	17190	2315
2	Bella vista	Amambay	8831	2257	2219
3	Capitan bado	Amambay	13502	3499	2319
4	Zanja pyta	Amambay	4292	1179	2603
5	Saltos del guaira	Canindeyu	20798	6231	2313

Showing 1 to 5 of 246 entries — Previous 1 2 3 4 5 … 50 Next

References

[1] G. E. Battese, R. M. Harter, and W. A. Fuller. An error-components model for prediction of county crop areas using survey and satellite data. *Journal of the American Statistical Association*, 83(401):28–36, 1988.

[2] R. Chambers and N. Tzavidis. M-quantile models for small area estimation. *Biometrika*, 93(2):255–268, 2006.

[3] C. Elbers, J. O. Lanjouw, and P. Lanjouw. Micro–level estimation of poverty and inequality. *Econometrica*, 71(1):355–364, 2003.

[4] K. Grünberger. Estimating food consumption patterns by reconciling food balance sheets and household budget surveys. *Rome: Food and Agriculture Organization of the United Nations*, 2014.

[5] I. Molina and J. Rao. Small area estimation of poverty indicators. *Canadian Journal of Statistics*, 38(3):369–385, 2010.

[6] J. N. Rao and I. Molina. *Small area estimation*. John Wiley & Sons, 2015.

[7] N. Tzavidis, S. Marchetti, and R. Chambers. Robust estimation of small-area means and quantiles. *Australian & New Zealand Journal of Statistics*, 52(2):167–186, 2010.

Este Proyecto es cofinanciado por el Consejo Nacional de Ciencia y Tecnología - CONACYT con recursos del FEEI