The Biophysical and Geographical Determinants of Hunger in Africa

Alex de Sherbinin
CIESIN, Columbia University

Presentation in the session "Making Population-Environment Research Relevant to Policy Makers"

2005 Open Meeting of the Human Dimensions of Global Environmental Change Research Community

12 October 2005

CIESIN
Columbia University

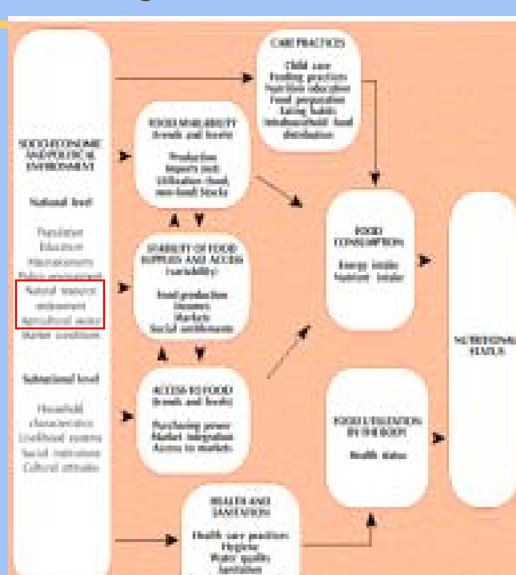
Structure of presentation

1. Presentation of the research on the determinants (or correlates) of child malnutrition in Africa
2. Consideration of the policy implications of this research, and more broadly how this kind of research may contribute to policy development

Research Questions

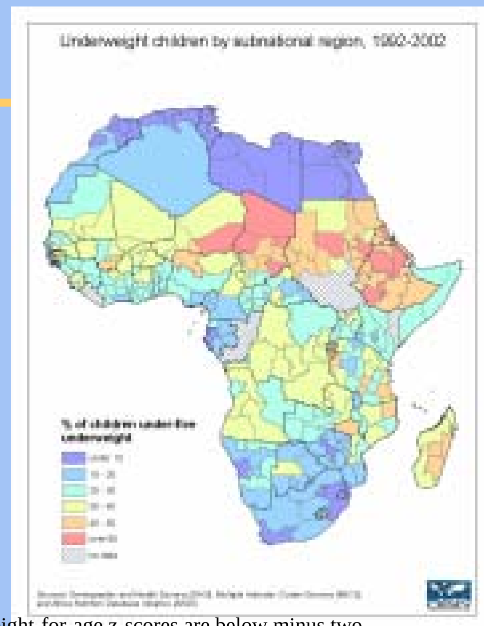
1. Most studies of malnutrition only look at household-level factors (e.g. income, ed. of parents, HH size, access to services)
2. When controlling for income, to what degree do biophysical and geographical variables explain variation in the rates of child malnutrition?
3. How does spatial autocorrelation affect the OLS results, and how can we correct for this?

Conceptual framework for understanding the causes of hunger



Source: FIVIMS http://www.fivims.net/static.jspx?lang=en&page=overview.

Data Set Construction



- Obtained percent of children underweight* from DHS and MICS surveys
- Match survey data to boundary data
- 377 sub-national units (SNUs)

* Children are defined as underweight if their weight-for-age z-scores are below minus two standard deviations (-2 SD) from the median of the NCHS/CDC/WHO International Reference Population.

Independent Variables (1)

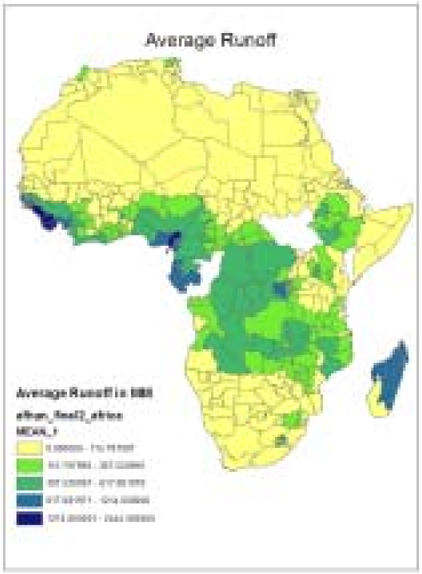
- GDP per capita (at national level)
  - Source: CIA World Factbook
  - Range: $500 to $10,700
- Runoff
  - Runoff is the proportion of precipitation that is left after evapotranspiration and the soil moisture deficit are satisfied
  - Source: GRDC/UNH Composite Runoff Fields v. 1.0
  - Range: 0 to 2.4 m
- Proportion of SNU within 2 km of a road
  - Source: Andy Nelson/UNEP Road Data 2003
  - Range: 0.0003 to 1
- Elevation (mean and standard deviation)
  - Source: SRTM
  - Range: 0 to 2,600 meters mean, 0-700 meters SD

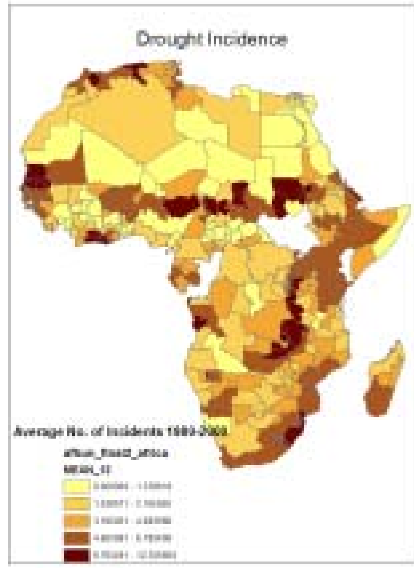
Independent Variables(2)

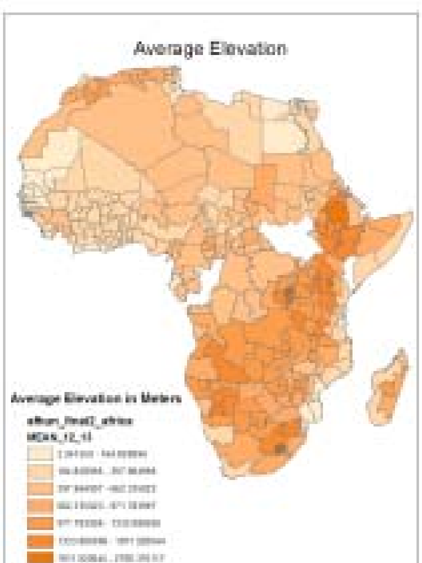
- Number of Drought Incidents (1980-2000)
  - Drought is defined as precipitation less than 75% of the median for 3 months or more
  - Source: International Research Institute for Climate Predictions
  - Range: 0-12.3 incidents (theoretical 0 to 14)
- Agricultural Constraints (soil, terrain, climatic)
  - Source: FAO-IIASA Global Agro-Ecosystem Zone Assessment
  - Range: 0.7 to 7 (min-max) (theoretical 0 to 7)
- Average level of land utilization for crops
  - Source: FAO
  - Range: 1.4 to 4.9 (max-min) (theoretical 1 to 6)
- Malaria Transmission Index
  - Source: Kiszewski, A. *et al.* “A Global Index Representing the Stability of Malaria Transmission.” *Am. J. of Trop. Med. & Hyg.*
  - Range: 0-33.7

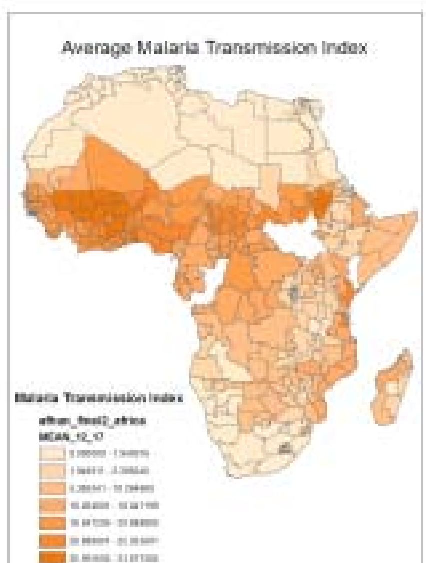
Mean conditions were calculated for populated portions of SNUs

- Utilized CIESIN’s GRUMP 1km population density grid
- Removed those portions of SNUs that were populated at less than 2 persons per sq. km.


Average Runoff
Average Runoff in MM

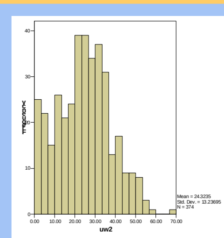

Drought Incidence
Average No. of Incidents 1980-2000

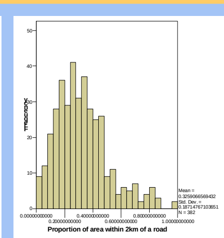

Average Elevation
Average Elevation in Meters

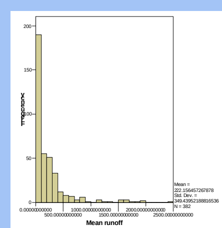

Average Malaria Transmission Index
Malaria Transmission Index

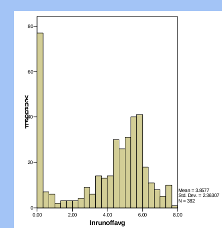
Data transformations

- Most variables approximated a normal distribution
- Took the log of highly skewed variables: runoff, elevation, and malaria transmission index
- Created dummy variables for North Africa, Ethiopia, and High Agricultural Constraints









Bivariate relationships mostly in the expected direction

Correlations

		Percentage Children Underweight	lnrunoffavg	Mean drought frequency	lnelevavg	Mean agricultural constraints	Average crop suitability index	Proportion of area within 2km of a road	lnmalavg	GDP per capita (CIA)
Percentage Children Underweight	Pearson Correlation	1	.237**	.189**	.256**	.016	.223**	-.392**	.409**	-.560**
	Sig. (2-tailed)	.	.000	.000	.000	.759	.000	.000	.000	.000
	N	374	374	374	374	374	324	374	374	374
lnrunoffavg	Pearson Correlation	.237**	1	.127*	.137**	-.412**	-.502**	-.254**	.330**	-.461**
	Sig. (2-tailed)	.000	.	.014	.008	.000	.000	.000	.000	.000
	N	374	374	374	374	374	324	374	374	374
Mean drought frequency	Pearson Correlation	.189**	.127*	1	.149**	-.202**	-.263**	-.091	.061	-.052
	Sig. (2-tailed)	.000	.014	.	.004	.000	.000	.079	.237	.315
	N	374	374	374	374	374	324	374	374	374
lnelevavg	Pearson Correlation	.256**	.137**	.149**	1	.004	.033	-.415**	-.218**	.049
	Sig. (2-tailed)	.000	.008	.004	.	.931	.555	.000	.000	.347
	N	374	374	374	374	374	324	374	374	374
Mean agricultural constraints	Pearson Correlation	.016	-.412**	-.202**	.004	1	.455**	.093	-.087	.095
	Sig. (2-tailed)	.759	.000	.000	.931	.	.000	.073	.093	.067
	N	374	374	374	374	374	324	374	374	374
Average crop suitability index	Pearson Correlation	.223**	-.502**	-.263**	.033	.455**	1	-.078	-.043	.066
	Sig. (2-tailed)	.000	.000	.000	.555	.000	.	.159	.446	.235
	N	324	324	324	324	324	324	324	324	324
Proportion of area within 2km of a road	Pearson Correlation	-.392**	-.254**	-.091	-.415**	.093	-.078	1	-.109*	.274**
	Sig. (2-tailed)	.000	.000	.079	.000	.073	.159	.	.036	.000
	N	374	374	374	374	374	324	374	374	374
lnmalavg	Pearson Correlation	.409**	.330**	.061	-.218**	-.087	-.043	-.109*	1	-.446**
	Sig. (2-tailed)	.000	.000	.237	.000	.093	.446	.036	.	.000
	N	374	374	374	374	374	324	374	374	374
GDP per capita (CIA)	Pearson Correlation	-.560**	-.461**	-.052	.049	.095	.066	.274**	-.446**	1
	Sig. (2-tailed)	.000	.000	.315	.347	.067	.235	.000	.000	.
	N	374	374	374	374	374	324	374	374	374

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Bi-variate Relationships

- A number of significant ones in the expected direction between underweight status and:
  - Drought incidence
  - Elevation
  - Crop suitability index
  - Accessibility to roads
  - Malaria transmission index
  - GDP per cap
- Malaria & GDP pc most highly correlated
- Surprisingly, runoff was *positively* related to percent underweight at the .01 level, and there was no significant relationship between agricultural constraints and percent underweight
- No bi-variate correlations exceeded .70

OLS Model Results

Dependent Variable: % of Children Underweight	**Unstandardized Betas**	**Standardized Betas**
Constant	16.136 ***	
GDP per capita	-0.002 ***	-.441
Log of Average Runoff	-0.875 **	-.158
Log of Average Elevation	2.292 ***	.244
Log of Average Malaria Transmission	2.808 ***	.271
Average No. of Drought Incidents	0.691 **	.122
Proportion of SNU <2km from road	-10.82 ***	-.154
North Africa Dummy	-4.185 **	-.122
Ethiopia Dummy	8.845 **	.113
High Agricultural Constraints Dummy	3.17 *	.098

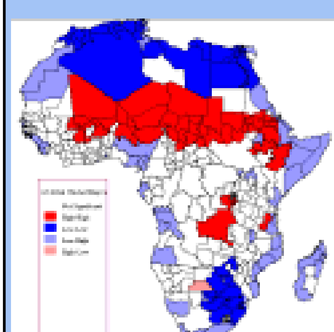
* p <.05, ** p <.01, *** p <.001
Adjusted R^2 = .524
N = 374

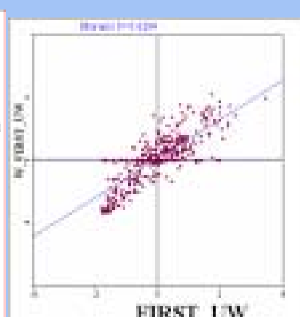
Spatial Autocorrelation (SA)

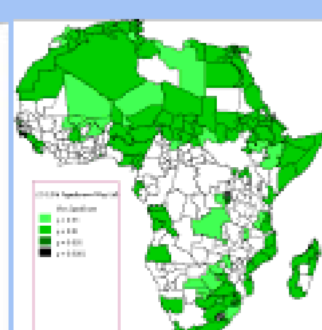
- The extent to which an occurrence of an event constrains or makes more likely an event in a a neighboring unit
- Like serial autocorrelation (in time series data), the events are not independent, and thus violates Gauss-Markov assumptions*
- Estimated coefficients are biased and inconsistent
- Residuals/Standard Errors are artificially deflated leading to type I errors (inproper rejection of null hypothesis)

* According to Lembo (undated): “If the observations… are spatially clustered in some way, the estimates obtained from the correlation coefficient or OLS estimator will be biased and overly precise. They are biased because the areas with higher concentration of events will have a greater impact on the model estimate and they will overestimate precision because, since events tend to be concentrated, there is actually a fewer number of independent observations than are being assumed.”

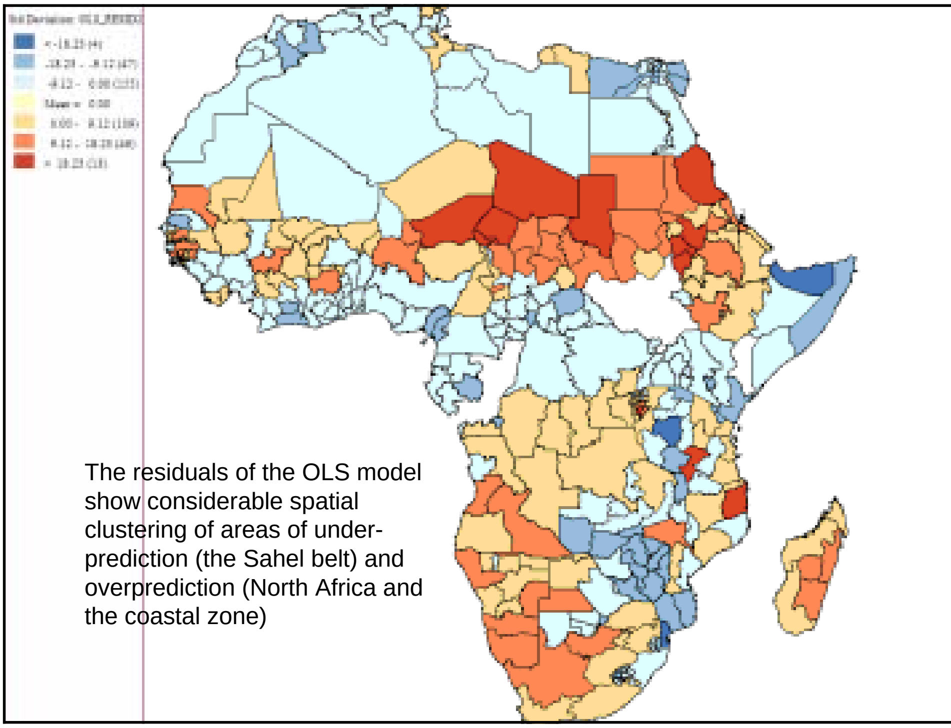
Evidence of Spatial Autocorrelation





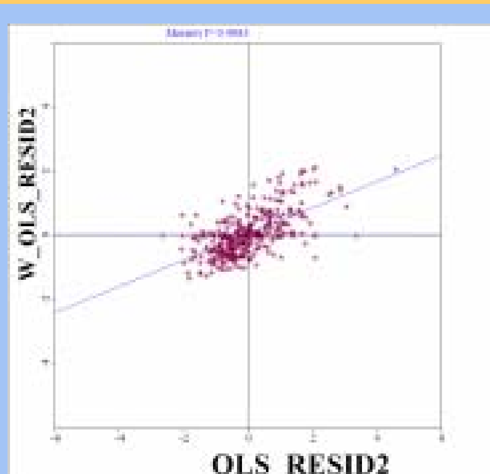


Moran’s I is similar to correlation coefficient, varying between –1.0 and + 1.0. When autocorrelation is high, the coefficient is high. A positive I value indicates positive autocorrelation.



The residuals of the OLS model show considerable spatial clustering of areas of under-prediction (the Sahel belt) and overprediction (North Africa and the coastal zone)

Moran's scatter plot for residuals of the OLS model



Correcting for SA

1. Identify any potential regimes that were not included in the model
   - Ethiopia dummy
   - North Africa dummy
2. Determine if a spatial lag or spatial error model is most appropriate
3. Fit an error model:

   "Under this specification, spatial autocorrelation in the dependent variable results from exogenous influences. Portions of the spatial autocorrelation may be 'explained' by the included independent variables (themselves spatially autocorrelated) and the remainder is specified to derive from spatial autocorrelation among the disturbance terms. The latter is assumed to occur because of one or more relevant spatially autocorrelated variables omitted from the design matrix, X." –Voss *et al.* 2005

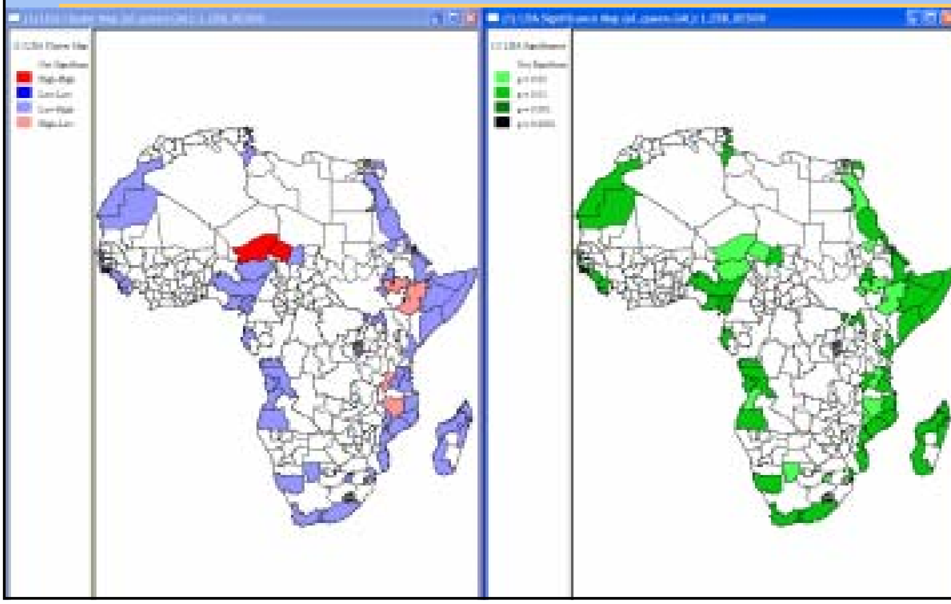
Voss, P.R., D.D. Long, R.B. Hammer, and S. Friedman (in press). "County Child Poverty Rates in the U.S.: A Spatial Regression Approach." Based on a paper presented at the 2003 Annual Meeting of the Population Association of America.

Spatial Error Model Results

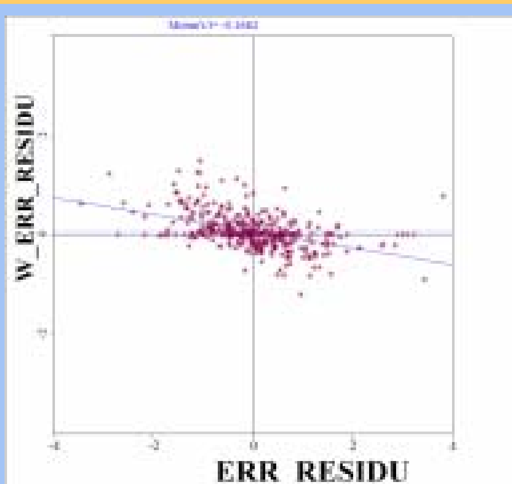
Dependent Variable: % of Children Underweight	**Unstandardized Betas**
Constant	22.132 ***
GDP per capita	-0.002 ***
Log of Average Runoff	0.348
Log of Average Elevation	1.05 *
Log of Average Malaria Transmission	0.246
Average No. of Drought Incidents (1980-2000)	0.684 ***
Proportion of SNU <2km from road	-13.436 ***
North Africa Dummy	-4.807 *
Ethiopia Dummy	10.943 **
High Agricultural Constraints Dummy	3.22 **
Lambda (autoregressive error term)	1.005 ***

* p <.05, ** p <.01, *** p <.001
Pseudo R^2 = .74
N = 374

Spatial clustering of residuals for error model



Moran’s I for the error model residuals



This low Moran’s I indicates that including the spatially autoregressive error term (Lamda) in the model has largely eliminated spatial autocorrelation.

Conclusions

- **What does all this mean?**
  - **Higher elevation areas tend to have higher levels of child malnutrition (even when controlling for the "Ethiopia effect"). This may reflect greater isolation, or constrained agricultural systems due to high slopes**
  - **Overall water availability is less important that the perturbations to agricultural systems from frequent drought (deviations from the mean)**
  - **High road density means greater access to markets, but may also be a proxy for wealth and accessibility to health and other services**
  - **SNUs that face the highest climate, soil and slope constraints to agriculture have significantly higher child malnutrition**
- **Limitations: scale dependence, coarse spatial resolution, error in the measures, lack of other household variables as controls**

Policy relevance

- **Potential policy responses:**
  - **build/improve roads into isolated areas**
  - **promote irrigated agriculture or bunds to trap rainwater**
  - **integrated soil fertility management (increase soil organic matter)**
- **Population-environment research in the past has been largely descriptive**
- **Importance of describing the specific set of geographical and biophysical constraints experienced by the poor**
- **Great potential for using geospatial databases to test relationships between demographic and biophysical variables in both directions, and to provide policy recommendations based on quantitative methods**
- **But, we must avoid the *ecological fallacy* of some past studies and control for *spatial autocorrelation***

Thank you very much!