Spatial Analysis of Childhood Mortality in West Africa

DHS Geographic Studies 1

MEASURE DHS+

This publication was made possible through support provided by the U.S. Agency for International Development under the terms of Contract No. HRN-C-00-97-00019-00. The opinions expressed herein are those of the authors and do not necessarily reflect the views of the U.S. Agency for International Development.

Spatial Analysis of Childhood Mortality in West Africa

DHS Geographic Studies 1

Deborah Balk
CIESIN (Center for International Earth Science Information Network)
Columbia University

Tom Pullum
Department of Sociology
University of Texas at Austin

Adam Storeygard
CIESIN
Columbia University

Fern Greenwell
ORC Macro

Melissa Neuman
CIESIN
Columbia University

December 2003

MEASURE *DHS+*
ORC Macro

Editor: Sidney Moore
Designer: Katherine Senzee

Suggested citation:

Balk, Deborah, Tom Pullum, Adam Storeygard, Fern Greenwell, and Melissa Neuman. 2003. *Spatial Analysis of Childhood Mortality in West Africa.* Calverton, Maryland, USA: ORC Macro and Center for International Earth Science Information Network (CIESIN), Columbia University.

Contents

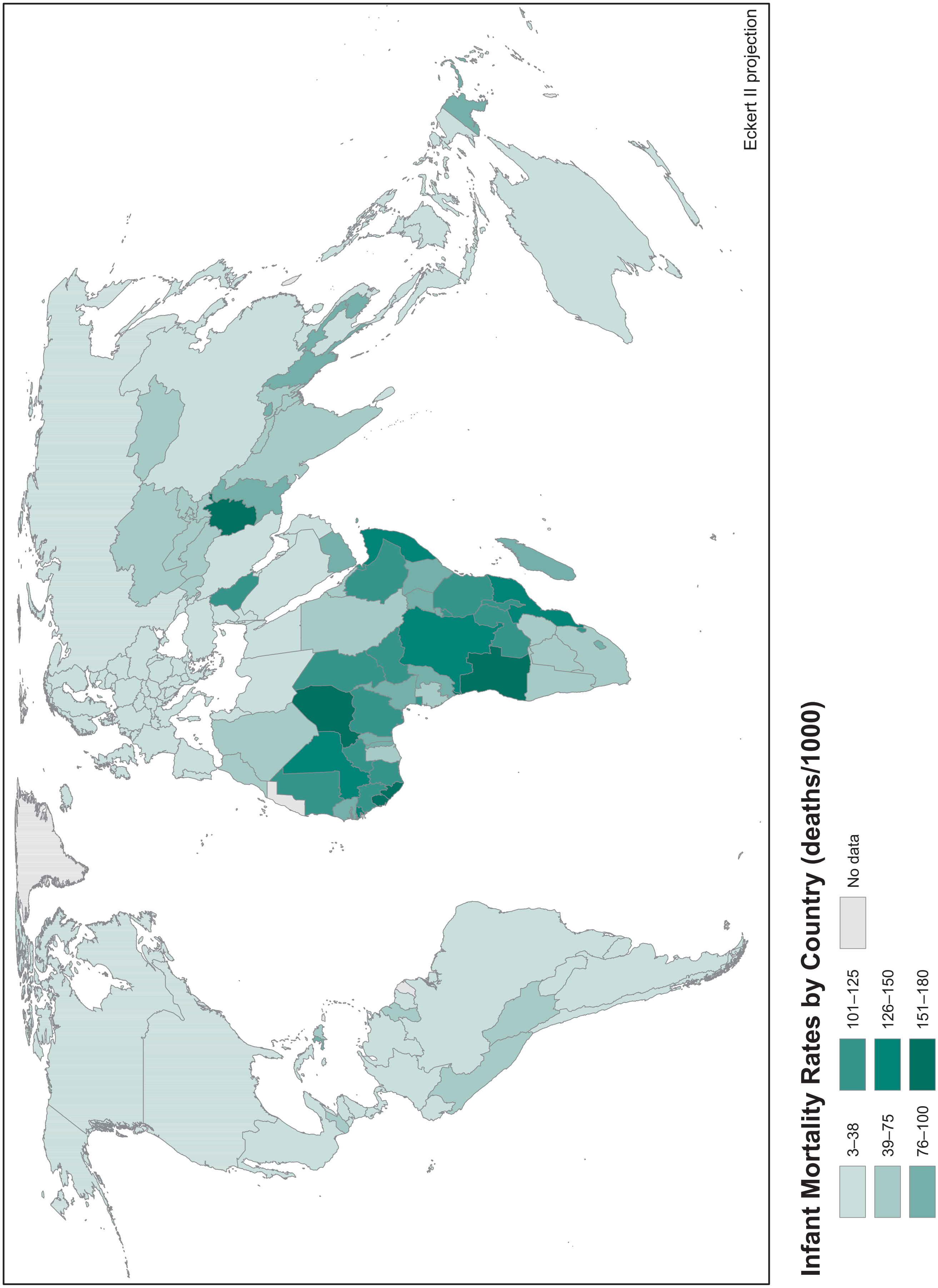
Maps, Tables and Figures

Acknowledgements

The authors thank the following persons: Livia Montana for her support and suggestions at the formative stages of the project; Greg Booma for his assistance in preparing many of the geospatial layers for analysis; David Johnson for his preparation and verification of the study cluster locations; Jerry Sullivan for his guidance on DHS data quality and age differentials with regard to infant and child mortality; Andrew Mellinger for providing the malaria data; and Alex de Sherbinin, Pedro Sanchez, and Stanley Wood for advice on the farming system and growing season variables.

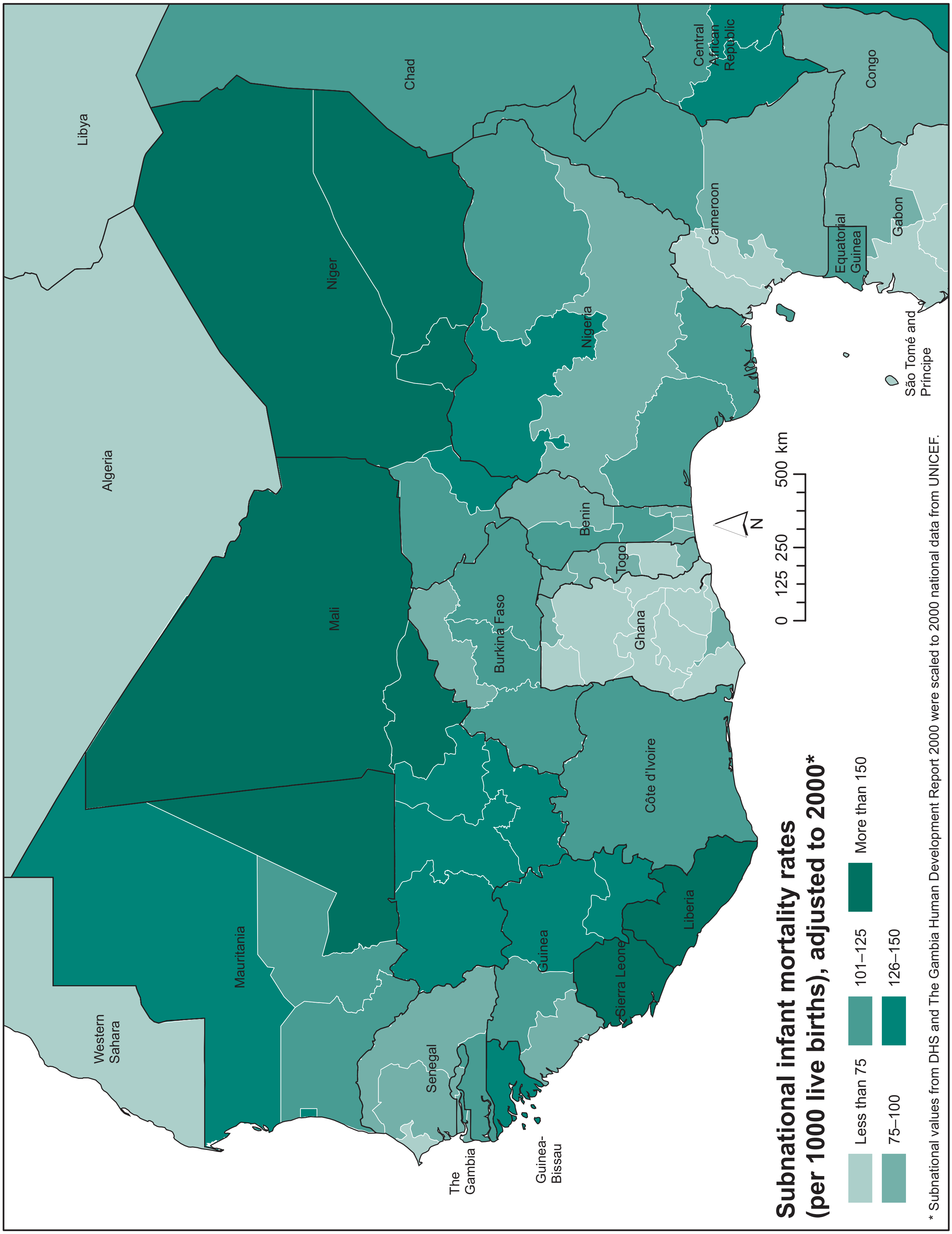
Maps

Map 1 **Infant mortality rates by country**



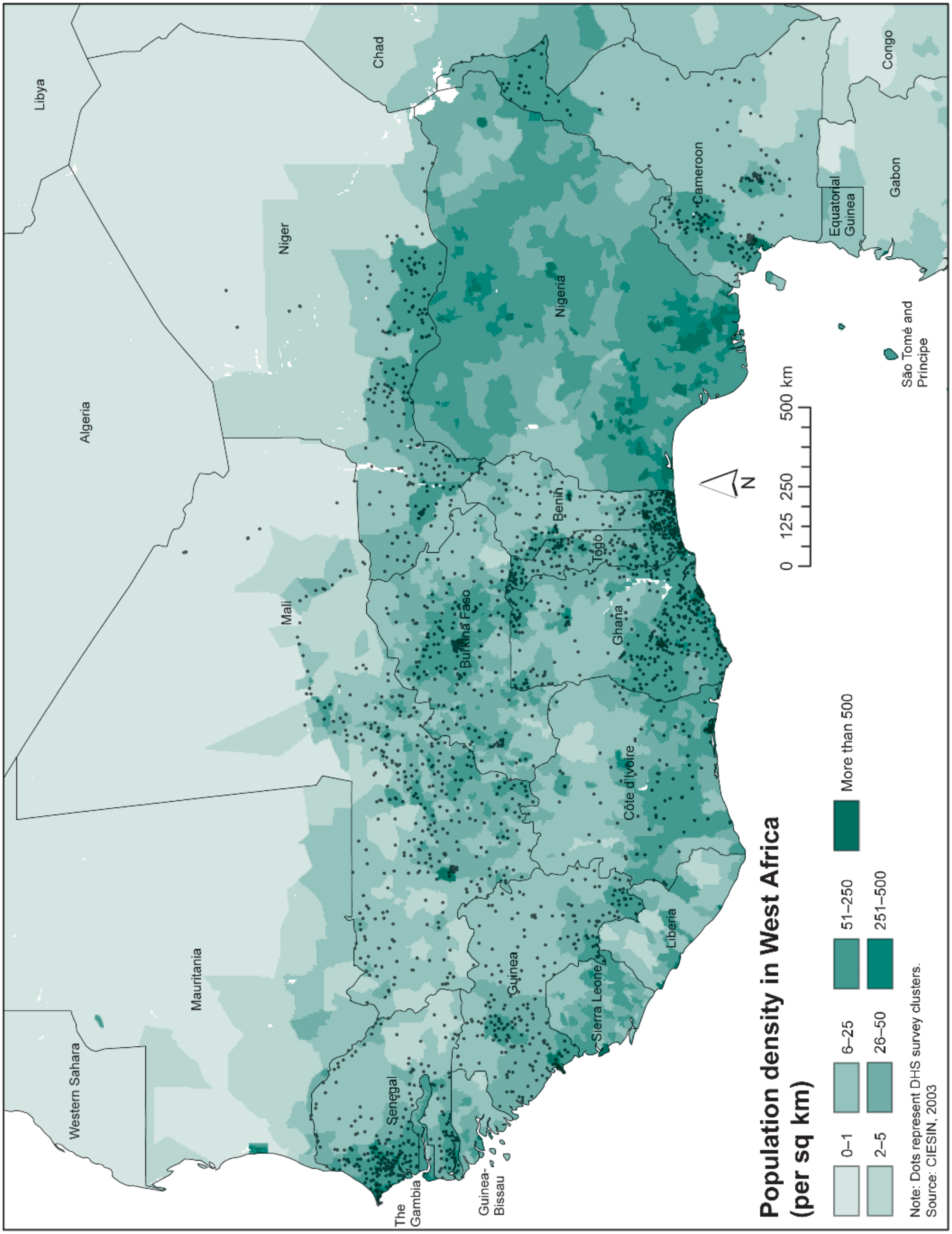
Source: UNICEF, 2000

Map 2 **Subnational infant mortality rates for West Africa**

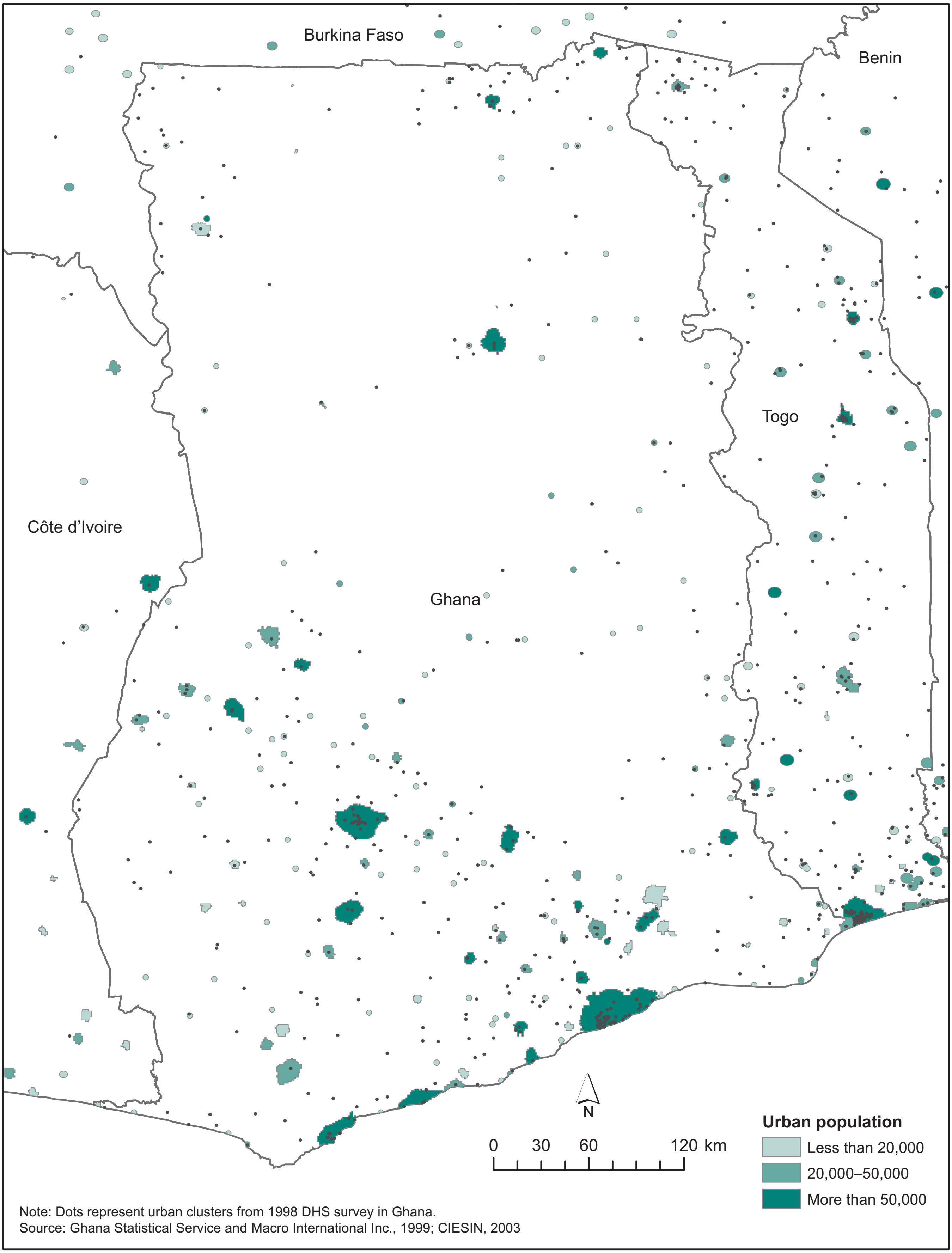


Map 3 **Study region for spatial analysis of childhood mortality in West Africa**

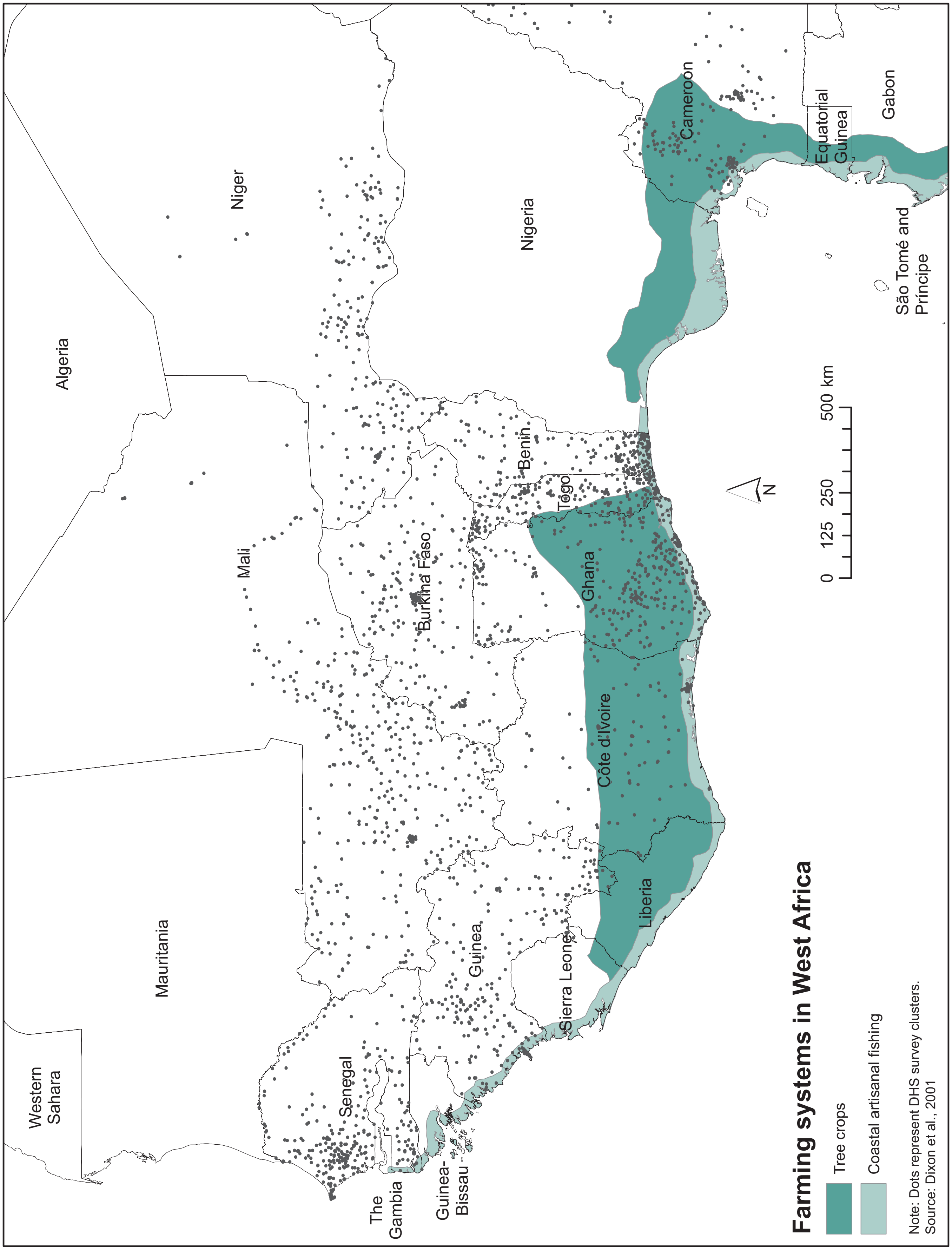
Map 4 **Population density in West Africa based on distribution of DHS survey clusters**



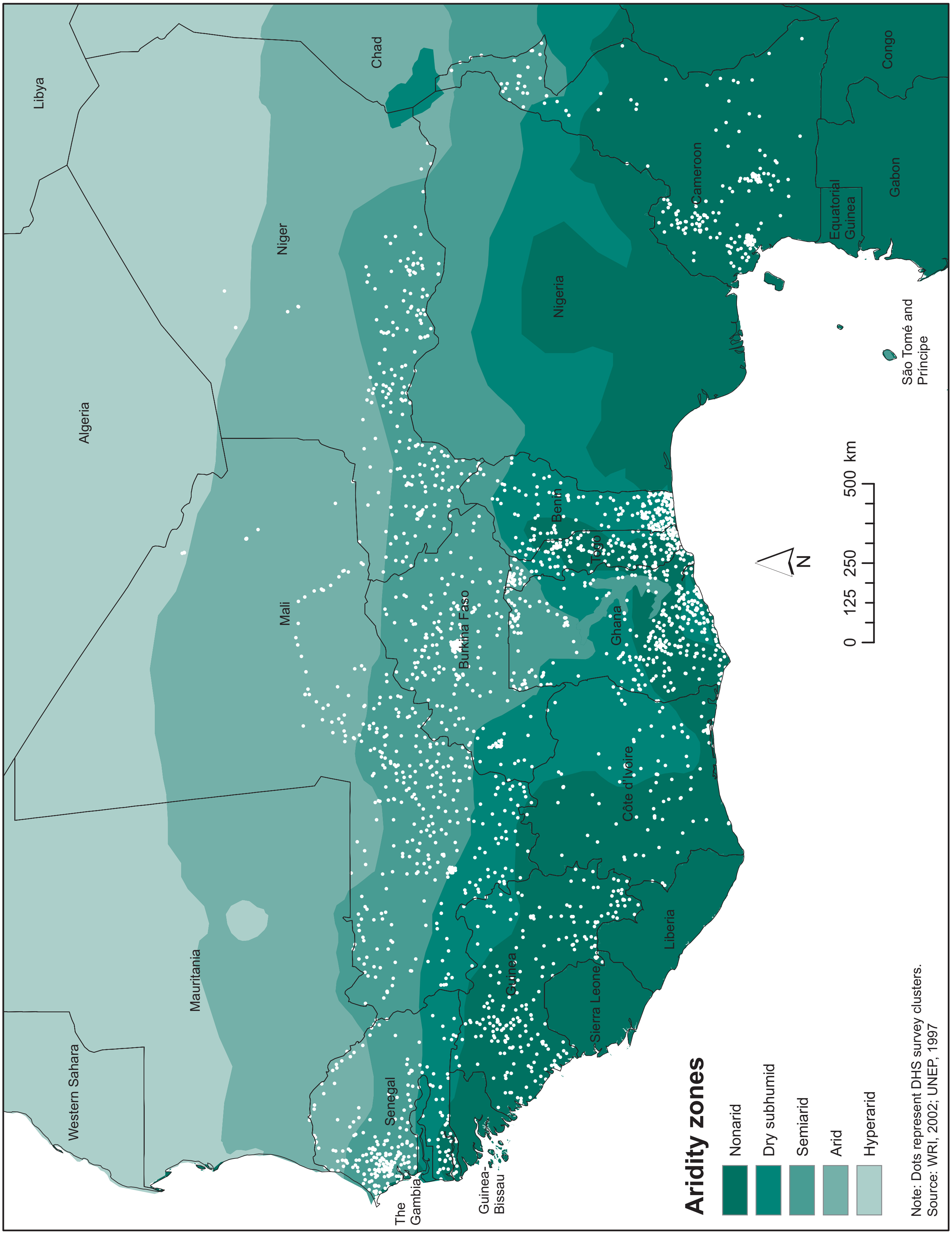
Map 5 **Urban areas and clusters in Ghana**



Map 6 **Farming systems in West Africa**

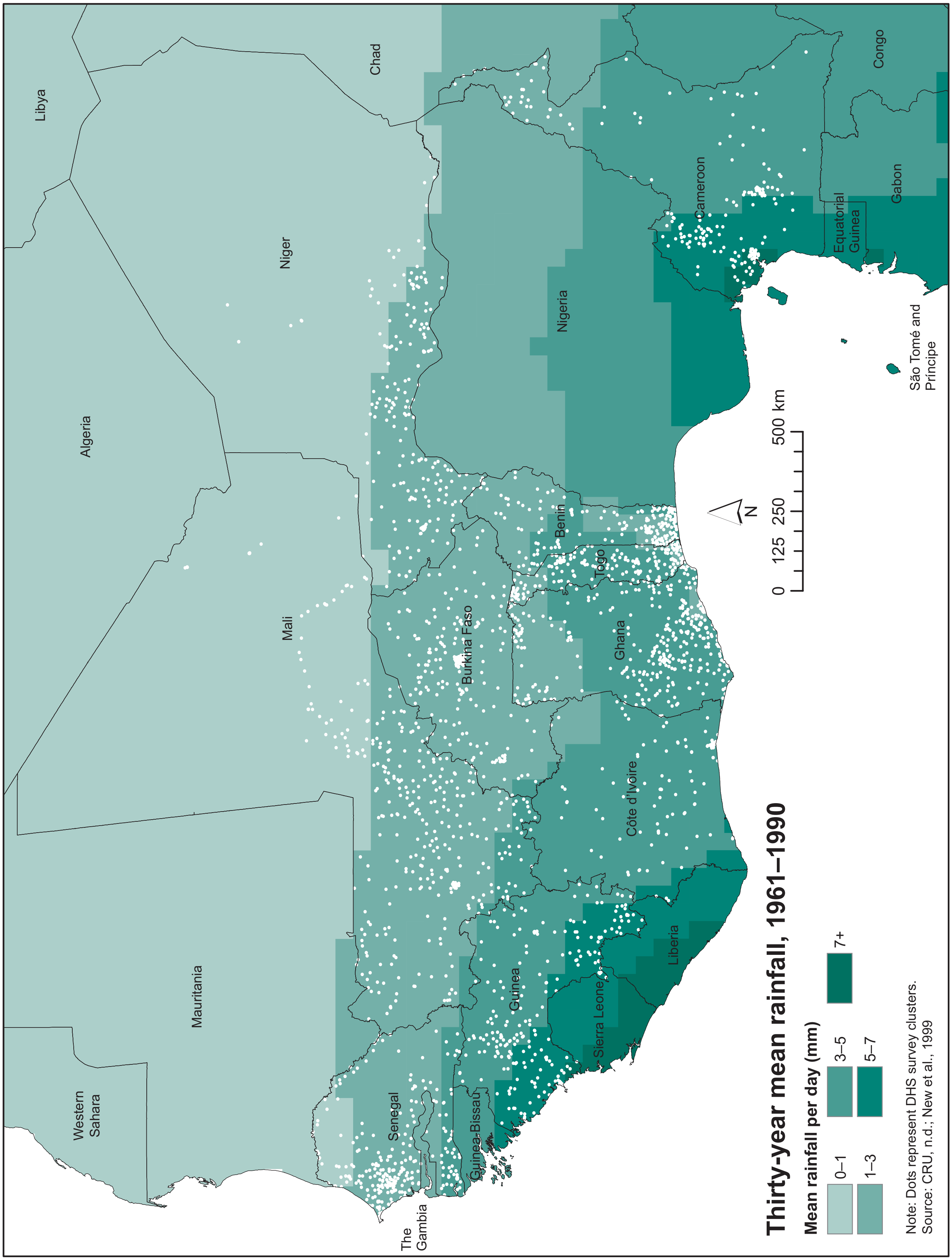


Map 7 **Aridity zones in West Africa**

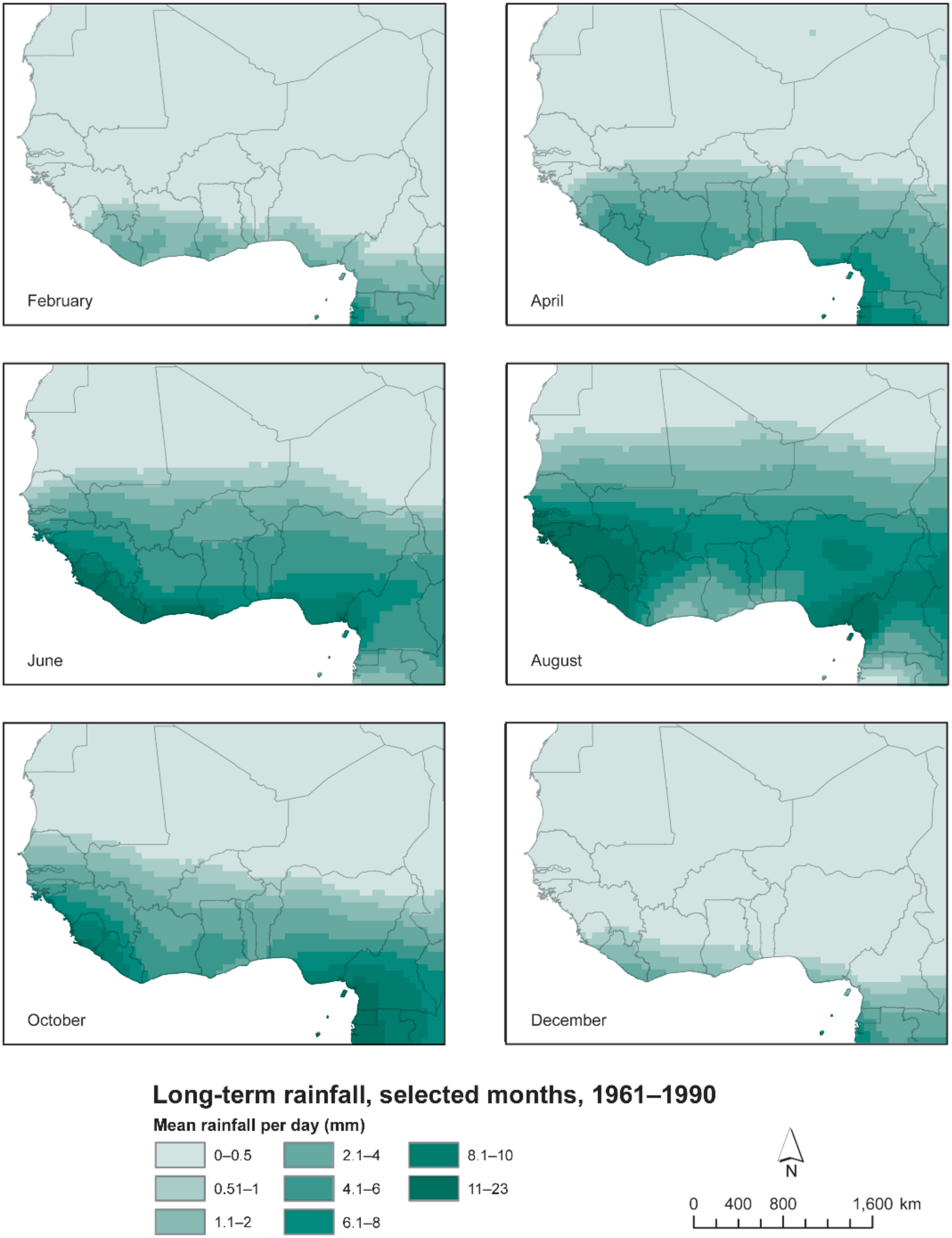


Note: Dots represent DHS survey clusters.
Source: WRI, 2002; UNEP, 1997

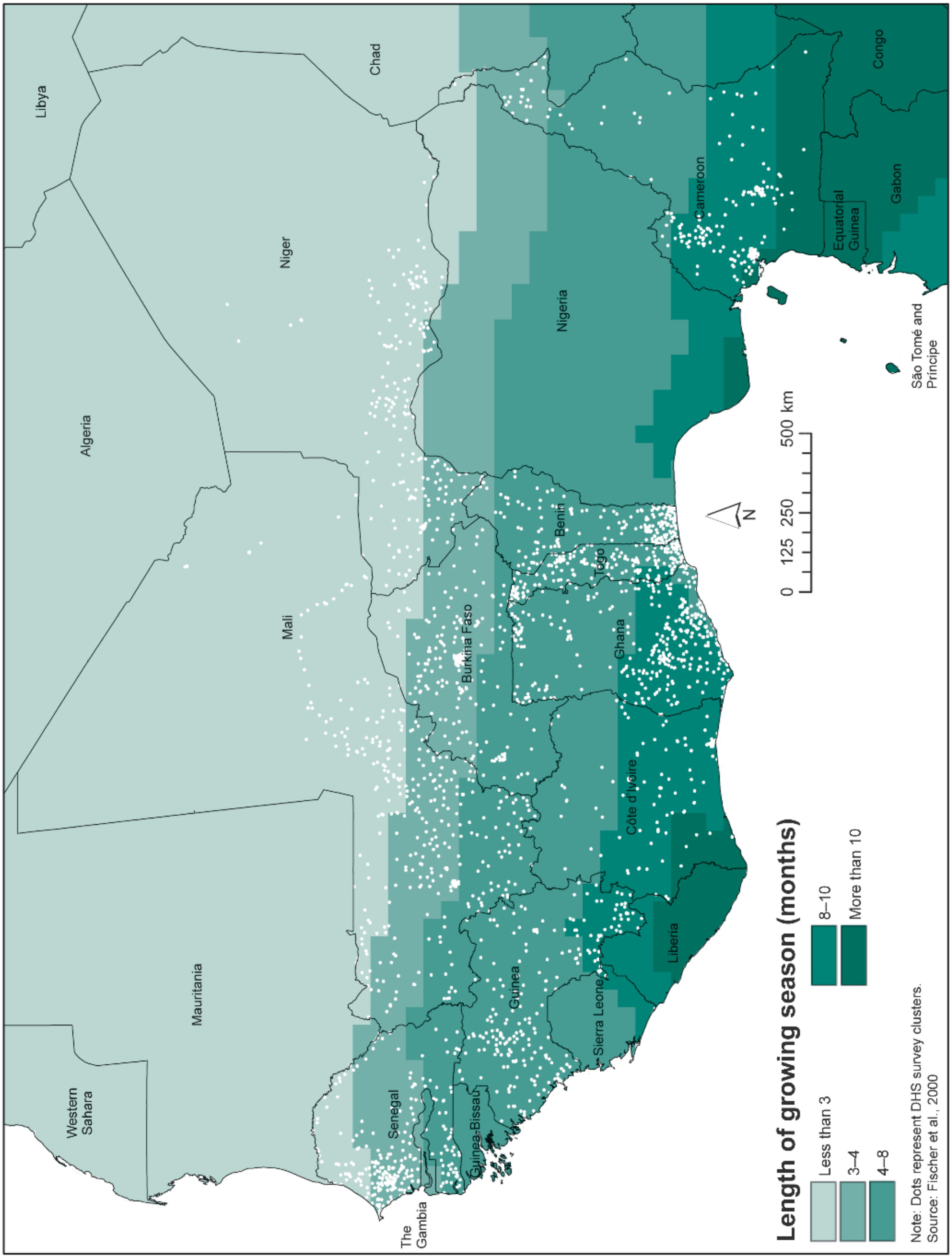
Map 8 **Thirty-year mean rainfall in West Africa**



Map 9 **Long-term rainfall in West Africa**



Map 10 **Length of growing season in West Africa**



1 Introduction

Infant and child mortality in Africa is higher than in any other continent (see Map 1). In particular, West African countries experience mortality two to three times higher than neighboring countries in northern Africa and in much of southern Africa. Still, there is considerable heterogeneity within the region. For example, Niger's infant mortality rate is more than double that of Ghana. Subnationally, even when mapped at a coarse resolution, rates differ by as much as a factor of four (see Map 2). The countries also show differential trends in levels and age patterns of childhood mortality. Further, while it appears that some countries have experienced significant declines in recent mortality (e.g., Niger), others appear to have experienced a reversal in long-term downward trend (e.g., Burkina Faso).[1] Because of the inherent complexities associated with analyzing trends from cross-sectional data, this report will focus on the major determinants of mortality in the ten years preceding the period 1997–2001. Its contribution is a consideration of a broad class of spatial covariates.

Several individual- and household-level factors have been identified as key determinants of infant and child survival. These include maternal education (Agha, 2000; McMurray, 1997; Rao et al., 1997; Root, 1997; Trussell and Hammerslough, 1983) and the pace of childbearing (Agha, 2000; Boerma and Bicego, 1992; Gupta and Baghel, 1999; Rao et al., 1997; Root, 1997; Whitworth and Stephenson, 2002). Many studies indicate that environmental or geographic factors also play an important role. These include, for example, population density (Root, 1997), climate (Curtis and Hossain, 1998; Patz et al., 2000; Pitt and Sigle, 1997; Ronsmans, 1995), disease environment (Root, 1999), and urban residence (Woods, 2003). However, few studies have been able to incorporate potential environmental factors that are explicitly spatial, that is, derived from geographic databases. Spatial variables include simple constructs, such as distances from households or communities (e.g., to the nearest clinic or city) and environmental characteristics that have their own geographic boundaries (e.g., types of farming system or land cover). Geographic databases often provide information (via station measurements, satellites, and other sources) that would otherwise be too costly to obtain through the survey mechanism. This study makes further inroads by incorporating several new or previously hard-to-integrate sources of spatial data.

Until recently, environmental and other geographic data were not readily applicable to analyses of childhood mortality. However, significant improvements are starting to take place. First, spatial data are generally becoming more available, with improved coverage, quality, and variety. Second, since late 1996, the Demographic and Health Surveys (DHS) has consistently recorded the geographical location of each cluster of surveyed households with handheld Global Positioning System (GPS) units. This information at the cluster level permits linkage between DHS determinants of infant and child mortality and information from other data sets.

Information from DHS surveys at the cluster level permits linkage between DHS determinants of infant and child mortality and information from other data sets.

The primary objective of this report is to explore and draw attention to the effects of

[1] Short-term variation in mortality based on survey data needs to be evaluated with caution. This is the subject of an ongoing study by Korenromp and colleagues (forthcoming).

a largely unexplored body of environmental information on infant and child mortality. The underlying motivation is to account for some portion of the variance that has not been explained by the traditional set of socioeconomic and biodemographic determinants of childhood mortality.

1.1. Rationale

Mosley and Chen's (1984) widely accepted analytical framework is based on the assertion that socioeconomic factors influence mortality through biological mechanisms called proximate determinants. The socioeconomic factors in the basic framework include individual, household, and community-level variables, the latter including macro-environmental factors. Factors such as climate, rainfall, and soil are especially pertinent to children's survival in sub-Saharan Africa because widespread poverty leaves the population highly vulnerable to fluctuations in the availability of food and water, the transmission of infectious and vector-borne disease, and even the amount of time a mother spends working versus her time devoted to child care (Watson et al., 1997).

Factors such as climate, rainfall, and soil are especially pertinent to children's survival in sub-Saharan Africa.

This study examines the role of nonbiological variables in predicting childhood mortality, with an emphasis on environmental and spatially-determined variables. We control for a variety of proximate determinants including several maternal and demographic factors. Because our goal is to infer the causal role of socioeconomic and environmental characteristics, it is beyond the scope of this study to analyze the direct, biomedical causes of death such as complications of birth, malnutrition, and specific infections such as diarrhea, HIV, and acute respiratory infection. These biological causes of death are believed to be correlated with social factors (Cramer, 1987; Schultz, 1993).

1.2 Evidence from the Literature

The impact of proximate factors, and socioeconomic and environmental factors acting directly or indirectly through them, on childhood mortality has been studied for several decades. Some factors have been examined much more thoroughly than others, and the following review is intended to guide the choice of variables for our analysis rather than provide an exhaustive overview.

Proximate Determinants

The proximate determinants of child mortality include maternal and demographic factors, nutrition, illness, and injury (Mosley and Chen, 1984). Maternal risk factors are more closely related to neonatal or early infant deaths because they are associated with premature and low birth weight infants and delivery complications. One of the most important maternal factors found to be related to childhood mortality is the pace of childbearing (Hobcraft et al., 1985; Rutstein, 1984). In particular, short preceding birth intervals are believed to increase an infant's risk of mortality because the mother's nutritional reserves have not fully recovered from the previous birth. Short birth intervals may affect the older child as well by creating competition between young siblings for the mother's resources (Boerma and Bicego 1992).

Two other important maternal factors are the mother's age at birth and the birth order. Results from a proportional hazards model using data from the Malawi DHS survey show that both of these effects are important in determining risks primarily during infancy (Manda, 1999). In sub-Saharan Africa, where women marry at a young age, first births are associated with very young mothers. Theory suggests that these women's children carry a higher risk of death because young, first parity mothers may not have reached their full physical and reproductive maturity (Zenger, 1992). Findings regarding children of older mothers and mothers of high parity vary more, but because of the increased risk of delivering a genetically impaired infant later in life, these children are also likely to carry higher risk of death (Sullivan et al., 1994).

Demographic factors such as male sex (Sullivan et al., 1994), multiple births (Pison et al., 1989), and previous child deaths (Majumder et al., 1997; Mturi and Curtis, 1995) are associated with a high risk of infant death. Infant boys, especially during the neonatal period, have a higher risk of death than females. Early infant death is also significantly higher for multiple births, mainly because multiple births are most likely to be premature and/or low birth weight. If more than one birth survives delivery then there is competition for breast milk and the mother's resources.

Nutrition, illness, and injury are common proximate determinants of childhood death. Although these factors are not included in this analysis, they cannot be overlooked as key factors in predicting childhood mortality. Numerous studies examining mortality outcomes have researched mother and child nutritional status as direct or indirect causes of infant and child deaths through association with specific diseases (Onis, 2000; Rice et al., 2000; Rutstein, 2000).

HIV/AIDS, a major epidemic in sub-Saharan Africa, is not without repercussions on childhood mortality. Adetunji (2000) finds that improvements in under-five mortality are reversed in countries with high adult HIV prevalence (≥5 percent). At the end of 2001, several West African countries in this study had estimated adult prevalence between 5 percent and 10 percent (Burkina Faso, Côte d'Ivoire, and Togo), and Cameroon had a prevalence level of 11.8 percent. The remaining four countries had estimated prevalence under 5 percent (Benin, Ghana, Mali, and Senegal) (UNAIDS, 2002). About 25 to 35 percent of children born to HIV-positive mothers are also infected with the virus, and the median age at death for HIV-positive children in Africa is about two years (Boerma et al., 1998). Mortality rates for children of HIV-infected mothers are therefore much higher—two to five times higher—than those for children of HIV-negative mothers. Perhaps even more important are the indirect effects of adult HIV on child mortality. Elevated adult HIV prevalence rates also increase the risk of death for HIV-negative infants and children because a parent's death leaves them vulnerable. The death of an HIV-positive parent or guardian means a loss of income and an orphan's time and energy are likely diverted from school to helping maintain the household. Unfortunately, precise effects of the disease on childhood mortality levels are difficult to capture, not only because of these indirect effects, but also because detailed information on children of mothers who died of HIV (as well as other causes) have often been omitted from household surveys including the DHS surveys.

Maternal education has consistently been observed to have a strong impact on child survival.

Socioeconomic Determinants

Unlike the endogenous maternal and demographic factors that substantially increase an infant's risk of death, the effects of socioeconomic variables are enhanced as the child gets older (Manda, 1999). The reason usually cited for this is that a greater proportion of child deaths between age 1 and 4 years are due to exogenous factors over which parents potentially have control. Parents' education, access to health services, and the household environment represent a few of these factors.

Maternal education has consistently been observed to have a strong impact on child survival (Agha, 2000; McMurray, 1997; Rao et al., 1997; Root, 1997; Trussell and Hammerslough, 1983). Paternal education has also emerged as a significant factor (Majumder et al., 1997). In part, maternal education is positively correlated with using modern health services including prenatal care (Shakhatreh, 1996). More education is needed to counteract child mortality than infant mortality, presumably because older children are more reliant on health facilities, clean hygiene practices, and quantity and variety of solid food—factors that better educated parents are more likely to seek out and gain access to (Boerma, 1996).

The use of health services, especially prenatal and delivery care, which is often a function of other socioeconomic factors, also reduces infant mortality (Ahonsi, 1995; Forste, 1994; Gaminiratme, 1991). The use of preventive health services, such as immunization programs, has been determined to influence survival later in childhood (Ahonsi, 1995; Diamond, 2000).

The household environment, measured by factors such as source of drinking water and toilet facilities, provides important determinants covarying with older children's chances of survival (Esrey and Habicht, 1986; Merrick, 1985; Woldemicael, 2000). These factors are important not only for their direct effect on child survival, but because they may also indicate the overall resource level of a child's family. Poverty in and of itself is a key determinant of infant and child mortality (Gupta and Baghel, 1999; Hussain et al., 1999).

In addition to socioeconomic factors, cultural factors may influence mortality. Society's beliefs about disease, for example, may result in taboos or ritualistic treatments whose therapeutic effects are not supported by modern medicine (Fabrega, 1972). Cultural beliefs may lead to breastfeeding practices that are detrimental to the infant's growth (Lesthaeghe, 1989; van de Walle and van de Walle, 1991). Basu (1997) contends that behavioral underinvestment may underlie the biological determinants of mortality. Cultural factors such as these and others are important in understanding childhood mortality, but because they are difficult to quantify they are not explicitly considered in the present analysis.

Spatially-relevant Factors

Although demographic analyses are almost always place-based, much analysis is spatially general. Urban-rural distinctions are common but are nearly always expressed with a dichotomous variable. Descriptions of study sites may set the stage for an analysis and assist in the explanation of residual effects, but even basic factors, such as population density (which might affect disease transmission) or other environmental characteristics identified in Mosley and Chen's frequently tested framework (1984), are not often considered in the formal analysis of mortality.

Urban residence is one of the most commonly identified factors in mortality variation, and the main reasons given for its importance in contemporary developing countries are spatial. Urban residents (and, just as importantly for disease transmission, their neighbors) have greater access than their rural counterparts to resources such as health services, clean water, sanitation, and education. Entwisle et al. (1997) considered a spatially sophisticated measure of nearness to resources. Using a network analysis of data on roads, they found significant relationships between contraceptive choice and accessibility to towns and health centers. Specifically, travel time effects are important even at short distances, and road composition plays a part in method selection.

Urban areas also have higher population densities, making it easier to share information and resources. In a recent article, Woods (2003) argues that mortality varies along an urban-rural continuum, rather than between discrete urban and rural environments, and that at least in the past in Europe the rural end of this continuum favored survival. He suggests that future analyses of urban-rural differentials in mortality should focus on mortality in childhood, which "appears to be highly sensitive to differences in population density" (Woods, 2003: 43). Defo's (1994) study of child survival in Cameroon using longitudinal data found that overcrowding has deleterious effects on both infant and child survival. Nevertheless, Woods (2003) recommends distinguishing between infant and child deaths, in part because "an excess of the latter may be found especially in urban centers and at times before the medical control of childhood diseases became possible" (Woods, 2003: 43).

Using a fairly coarse but nonbinary measure of urbanness, Gupta and Baghel (1999) found that urban residence is an important factor in infant mortality. Mortality in the slums was found to be higher than in other parts of urban areas, but the rates in slums were more favorable than in rural parts of India. Further, mortality was found to be higher in the slums of major cities than in smaller metropolises.

Other recent work has shown the importance of spatial disaggregation. Root (1997) contends that population density is an important factor in spatial patterns of child mortality in Zimbabwe, although his test of this hypothesis is crude; he divides the country at a coarse level into high- and low-density regions. In his study of West Africa and East/Southern Africa, Root (1999) used DHS data to examine patterns at the subnational (regional) level. These data typically consisted of first-level administrative units or aggregations thereof. Root found important subnational patterns, and suggested that these patterns should be analyzed in connection with population density and vector habitat data, key factors in the transmission of infectious diseases.

Developments in small area estimation methods (Elbers et al., 2003) have enabled researchers in several countries to combine low spatial resolution household survey data with high-resolution census and physical data in order to estimate health and economic indicators at high resolution. Specifically, Fujii (2002) and Fujii et al. (2002) have combined Cambodian household survey and census data with spatial data including land use, agricultural production, climate, vulnerability to flooding, and distance to rivers, roads, towns, cities, and health facilities to generate estimates of poverty and malnutrition with acceptable standard errors for most communes. However, because the spatial data are used to estimate the demographic indicators, the two classes of data cannot be compared statistically.

Climate is of potential interest because it incorporates factors affecting agricultural production and disease transmission.

Lastly, demographic analysis has long been concerned with the relationship of population dynamics and agricultural production (Boserup, 1965; Malthus, 1798). Several recent studies have shown the importance of spatially-specific climatic factors on health and mortality outcomes (NRC, 2001). Climate is of potential interest because it incorporates factors affecting agricultural production and disease transmission (through vector, water, and airborne mechanisms). Curtis and Hossain (1998) examine the effects of aridity on child malnutrition and find it to be a significant predictor of wasting (see next section). Findley et al. (2002) find that the incidence of infectious diseases is closely linked with rainfall in Mali; malaria is most prevalent one to two months after peak rainfall, and acute respiratory infections peak in dry months. Quantitative research has received support from in-depth qualitative research. For example, Adams (1994) and colleague (Sauerborn and Adams, 1996) find complex connections between climate anomalies, household food security, and the health and nutrition of household members in rainfall-dependent agricultural communities in Mali. Pitt and Sigle (1997) find that seasonal variability in rain may cause problems in smoothing income and resource distribution across seasons, ultimately compromising the well-being of children in Senegal. This effect is magnified in rural areas, where households are often more vulnerable to environmental shock than urban households. Numerous studies have shown seasonality in the incidence of diarrhea (e.g., Armah et al., 1994; Muhuri, 1996). These climatic variables, while intrinsically spatial, are often specified only as time variances.

1.3 DHS Experience with Cluster-level Spatial Data

DHS data were first used in regional highly spatially disaggregated form in the West Africa Spatial Analysis Prototype Exploratory Analysis (WASAP) project. WASAP studies analyzed differences in demographic and health indicators across social and ethnic borders and aridity zones in 12 countries. Curtis and Hossain (1998) used WASAP data to consider the effects of aridity, population density, agricultural production, and market tension (a theoretical measure of the "pull" of local and international markets based on agroclimatic and infrastructure data) on child malnutrition. Controlling for correlates from the DHS data (maternal education, birth order, age, incidence of diarrhea), only aridity and nonfood crop production were significant predictors of wasting, and only market tension was a significant predictor of stunting. Saha (1998) linked increases in market tension and level of market tension and economic diversity with knowledge and use of modern methods of family planning.

Expected Gains From Current Approach

The current data mark an improvement over WASAP in several respects. First, the cluster locations have been geocoded more consistently, using handheld GPS units. Second, the component surveys were carried out over a shorter time interval (five years instead of ten years). Lastly, the increased availability of spatially explicit physical and population data allows for analysis with a wider range of variables at higher resolution. For example, WASAP took its population data from a set of agricultural censuses averaging less than 25 units per country. The current study uses population data from more than 1200 units in the ten survey countries.

2 Data and Study Design

The analysis is based on DHS data and linked information from a variety of spatial data sources. DHS data analyzed in this report are drawn from the ten most recent georeferenced surveys in West Africa: Benin, Burkina Faso, Cameroon, Côte d'Ivoire, Ghana, Guinea, Mali, Niger, Senegal, and Togo (see Map 3). Since data collection was carried out within a relatively short period of time, 1997 to 2001, period effects on mortality experience were minimized. Surveys were also conducted in this time period in neighboring Nigeria, Gabon, and Mauritania, but they did not include the georeferenced cluster data necessary for locating respondents accurately.

Data on 122,389 children from the selected surveys who were born during the ten years before the respective dates of interview were pooled into one data set. Since we are interested in exposure to death up to the fifth birthday, about half of the cases were right censored in the calculation of child mortality.[2] Surveyed births are located in 2,771 clusters across the ten countries. The locations of these clusters were recorded at the time of the survey using GPS devices (see Map 4).

2.1 Adjusted Weights

All of the DHS surveys used in this report are nationally representative.[3] The sample design is a probabilistic two-stage sample, where enumeration areas (EAs) are randomly selected with probability proportional to their size. The households within the selected EAs are randomly selected with equal probability, and sampling weights are assigned to individuals. A thorough review of sampling methodology is presented in the DHS Sampling Manual (Macro International, 1996).

For this analysis, information on the 122,389 children described above was pooled into one data set. Because of large differences across country populations and sample sizes, the sample weights in the pooled data set needed to be rescaled to represent the ten countries in proportion to their populations. For example, the births in the Côte d'Ivoire sample in the ten years preceding the survey represented only 0.06 percent of all the births in that country in the same time period. The births in the Togo sample in the ten years preceding the survey represented 0.71 percent of all the births in that country in the same time period. An expansion weight was calculated for each country and then multiplied by the original sample weight. The weights were then renormalized to average to one across the pooled sample. The new weights were applied in the analysis. Because of our primary interest in spatial clustering, we have not adjusted

Because of large differences across country populations and sample sizes, the sample weights in the pooled data set needed to be rescaled to represent the ten countries in proportion to their populations.

[2] Right censoring refers to those cases whose observed time is truncated before their fifth birthday. We have only partial information, that is, we know that they survived until at least the time of the interview.

[3] The Mali and Niger surveys exclude remote populations, totaling 2.6 and 4.7 percent of their populations, respectively. Details follow in the section on the aridity variable, which is most likely to be affected. Residents of refugee camps were not surveyed in Guinea.

for maternal clustering. Subsequent analyses could account for associations between siblings.

2.2 Data Quality

The main issue concerning data quality is that the age at death data, which is reported in months, shows considerable heaping at 12 months. Some of the deaths reported at 12 months may have actually occurred at 10, 11, 13, or 14 months. The interpretation of this response can be important for the estimation of infant mortality, in particular. To the extent that deaths at 10 or 11 months were misreported as 12 months, heaping will result in a an underestimate of infant mortality. Heaping of age at death at 12 months happens to some extent in all DHS surveys because of respondent error or interviewer error.

An index of heaping at 12 months may be calculated by dividing the number of deaths at 12 months by the mean number of deaths at months 10, 11, 12, 13, and 14 (Curtis, 1995). In this study, the mean heaping index was 2.8, ranging from 1.5 in Niger to 4.2 in Guinea. Overall, this implies that unadjusted infant mortality rates are underestimated by about 2 percent. Conversely, because there are fewer deaths in the 1 to 4 years age group, the corresponding unadjusted child mortality rates are overestimated by slightly more than 2 percent.

Because there is no way to accurately redistribute individual deaths by changing the age at death, and because the adjustment would not significantly influence the relative rates of mortality of interest in this study, no adjustment was made for heaping in this report. Nevertheless, we have taken steps to avoid an ambiguity in the interpretation of month that is sometimes overlooked. Following the usual convention when age is reported in years, we assume that age at death in months means completed months of age. Therefore, to estimate exact age at death, 0.5 months was added to each age reported in months. By this reckoning, for example, 12 months becomes 12.5 months, which is clearly past the first birthday.

Following the usual convention when age is reported in years, we assume that age at death in months means completed months of age.

2.3 Measures of Infant and Child Mortality

DHS estimates of infant and child mortality use direct methods and are based on birth histories. They are period-specific rather than cohort-specific, meaning that children of a particular age were exposed to the risk of death during a five-year period prior to the survey date—but not necessarily the five years immediately preceding the survey. See Sullivan et al. (1994) for a detailed discussion of DHS childhood mortality estimates. Period-specific rates are synthetic cohort probabilities in which children of different birth cohorts contribute to the mortality experience of different subintervals of age. The advantage of calculating a synthetic rate is that in using partial survival time information at the date of interview, we have estimates for the most recent period, rather than only for children who have been observed for the full period of interest. Table 1 shows infant and child mortality rates for the early 1990s and the late 1990s for the countries selected for analysis in this report.

Two standard measures of child mortality are ${}_1q_0$, the probability of dying in the first year of life, and ${}_4q_1$, the probability of dying during ages 1–4, given that the child survived the first year. The infant mortality rate (IMR), when divided by 1000, is equivalent to ${}_1q_0$. The probability of surviving to age five can be expressed as $(1 - {}_1q_0) \times (1 - {}_4q_1)$. The analysis in this report is based on estimates of ${}_1q_0$ and ${}_4q_1$.

2.4 Geographic Data

Over the past decade, GIS software has become more accessible to social and health researchers, and geographic data have become increasingly available in formats that may be integrated with georeferenced survey data. Nevertheless, integration remains a nontrivial undertaking and therefore geographic variables must be selected with care. Using GIS software, geographic data were assigned to cluster locations, which were in turn appended to household, maternal, and child data from the DHS surveys. That is, the cluster locations were plotted on each geographic layer (e.g., rainfall, urban extents). Distances were then calculated to features including urban areas and the coast. The values of the geographic layers where the DHS clusters fell were then attributed back to the individual households in the clusters.

Using the newly updated Gridded Population of the World (GPW) (version 3, alpha), population densities (in the year 2000) were recorded for each cluster location, and calculated for the area within a 10 and 30-kilometer radius of each (CIESIN, 2003; see Map 4). The GPW database reallocates population estimates from the census units in which they were collected (roughly 1,200 for the ten countries in the study) to a 2.5-minute quadrilateral grid, a format easily overlaid with the DHS cluster points (see Table 2). Distances were calculated to the coast (using the Digital Chart of the World's coastal boundary data) and the nearest populated places of 20,000 and 50,000 residents, coded both as point locations and as urban extents of finite area (Balk et al., 2003; see Map 5). Similar to GPW, the database of populated settlements uses census data that is assigned to urban polygons as delineated by the Nighttime Lights dataset (Elvidge et al., 2001) and a few other sources, as the lights are of inferior quality in parts of Africa (Balk et al., 2003). (The Nighttime Lights data detect the amount of stable light at night, and are highly correlated with both urban areas and electrification.[4]) All of the above variables were calculated a second time ignoring any part of the above area that was on the other side of a national border.

Farming system (Dixon et al., 2001, see Map 6), arid zone (UNEP, 1997; WRI, 2002; see Map 7), average rainfall (CRU, no date, see New et al., 1999; see Map 8), growing season (Fischer et al., 2000, see Map 10) and an index of malaria risk (Kiszewski et al., forthcoming) were calculated at each cluster point. Each of these observational datasets was developed for application in agricultural, climate, or other research areas, but due to

Table 1
Basic data on survey countries

	5-year period preceding survey				10-year period preceding survey							
	Infant mortality ($_1q_0$)		Childhood mortality ($_4q_1$)		Infant mortality ($_1q_0$)	Childhood mortality ($_4q_1$)	Births	Female sample	Clusters	Year		
Country	Most recent survey	Prior survey	Most recent survey	Prior survey	Most recent survey	Most recent survey	Most recent survey	Most recent survey	Most recent survey	Most recent survey	Prior survey	Pop '95 (UN, '000)
Benin	89.1	93.9	77.8	80.0	94.8	75.0	10,395	6,219	247	2001	1996	5,336
Burkina Faso	105.3	93.7	127.1	103.1	108.6	129.5	11,734	6,445	210	1998-99	1992-93	10,415
Cameroon	77.0	64.3	79.9	65.2	79.8	72.3	7,859	5,501	203	1998	1991	13,182
Côte d'Ivoire	112.2	88.5	77.2	66.9	111.5	70.7	3,884	3,040	140	1998-99	1994	13,528
Ghana	56.7	66.4	53.9	56.8	61.2	52.4	6,555	4,843	400	1998	1993	17,649
Guinea		98.0		87.5	106.6	99.0	12,011	6,753	293	1999		7,153
Mali	113.4	122.5	130.5	131.2	126.2	128.3	25,984	12,817	403	2001	1995-96	9,944
Niger	123.1	123.1	171.8	222.6	135.8	193.0	15,333	7,577	268	1998	1992	9,150
Senegal	67.7	68.2	76.5	68.2	69.4	75.2	14,569	8,593	320	1997	1992-93	8,330
Togo	79.7	77.3	72.3	83.8	80.3	69.0	14,065	8,569	288	1998	1988	4,060

Source: DHS Statcompiler, United Nations

[4] Because the censor detects stable light sources, especially electricity, which is notably absent from many parts of Africa, including urban areas, supplemental sources and indirect estimation techniques were also applied to estimate the extent of urban settlements in Africa (see Balk et al., 2003 for additional details).

the common GIS format and scale, they are appropriate for integration with the DHS data. A full list of geographic variables appears in Table 2 and is described in detail in Chapter 3.

Table 2
Geographic variables

Variable	Source data	Source	Description	Variants	Resolution
		Demographic			
Population Density	Gridded Population of the World (GPW) v. 3	CIESIN		A1) at cluster, A2) within 10 km, and A3) within 30 km, B1) unconstrained and B2) constrained by national borders	2.5 minutes
Urban proximity	Urban-rural	CIESIN	Distance (euclidean) to nearest urban area	Urban areas > A1) 20,000 and A2) 50,000 people, coded as B1) points (presumed centroids) and B2) polygons, C1) unconstrained and C2) constrained by national borders	1 minute
Coastal proximity	Digital Chart of the World (DCW)-derived continent boundary	National Imagery and Mapping Agency (NIMA)	Distance (euclidean) to nearest point on the coastline		1:1,000,000
Distance to roads	VMAP roads data	National Imagery and Mapping Agency (NIMA)	Distance (euclidean) to nearest point on a road		1:1,000,000
		Ecological			
Farming system	Farming systems	FAO	Farming system based on the classification system by Dixon et al. (2001)		Unspecified
Arid zone	Arid zones	Millennium Ecosystem Assessment	Type of arid zone (non-arid zones are undifferentitated)		Unspecified
Stability of malaria transmission index		Kiszewski et al., forthcoming	Composite of environmental and epidemiological data		30 minutes
Rainfall	CRU05 0.5 Degree 1961–1990 Mean Monthly Climatology	Intergovernmental Panel on Climate Change/ International Research Institute for Climate Prediction (IPCC/ IRI)	Average monthly rainfall at cluster, in mm/day, 1961–1990	By month, yearly average, maximum month	30 minutes
Growing season		International Institute for Applied Systems Analysis (IIASA)	Length of growing season, in months		30 minutes

3 Methods

To describe differences in mortality by single variables we undertook a survival analysis using the Kaplan-Meier (product-limit) method. This method tracks the age pattern of mortality during the first year of life and during the next four years for selected covariates. For the multivariate analysis, we used a generalized linear model to fit $_1q_0$ and $_4q_1$. Each method is described in detail below, following a description of the covariates.

3.1 Variable Selection

Table 3 shows mean values or proportions of the variables included in the analysis for the sample as a whole and for each of the ten countries in the study. These variable sets form the basis of the models in the multivariate analysis.

3.1.1 Control Variables [Model 1]

Country and birth cohort are included as control variables. In the multivariate analysis below, Ghana, with the lowest mortality, is selected as the reference country. The ten surveys were conducted within a five-year period, and in each country we focus on the births during the ten years preceding the survey. The years of births in the pooled data file range from 1987 to 2001. As a partial control for period trends in fertility, we broke this range into three five-year intervals: 1987–1991, 1992–1996, and 1997–2001. We recommend caution in the interpretation of differentials and coefficients for the ten countries and the three cohorts. Country and cohort are likely to represent the net effect of many unmeasured influences that vary across countries and time. Moreover, they are somewhat confounded simply because the surveys were not conducted at the same time.

3.1.2 Proximate Determinants [Model 2]

The child's sex, birth order, and multiple birth status were included. Guinea had the highest proportion of male children, 51.4 percent, with a study average of 50.5 percent. Multiple births, accounting for 3.6 percent of all births, were much more common in Benin (5.8 percent) and Togo (4.9 percent). Mother's age at birth was also included: Ghana, Togo, and Senegal (with the lowest fertility among the ten countries) have relatively small shares of younger mothers and relatively large shares of older mothers.

Birth spacing was not explicitly included, although many studies have shown that when birth intervals are short, i.e., less than two years, both the child at the beginning and the child at the end of the interval are more likely to die. (This effect is due to the competition for maternal time and resources—similar to the competition between the children in a multiple birth, an included variable.) In our preliminary analysis we found that much of the effect of birth spacing is captured by birth order. A high proportion of the births are first births, and they have no preceding birth interval. We were also concerned by the censoring of the subsequent birth interval, a result of our focus on recent births. Lastly, there are endogenous or feedback effects because an early child death can

In our preliminary analysis we found that much of the effect of birth spacing is captured by birth order.

Table 3
Descriptive statistics of key covariates, by country

Variable	Benin	Burkina Faso	Came-roon	Côte d'Ivoire	Ghana	Guinea	Mali	Niger	Senegal	Togo	All	Type
Control variables												
Birth cohort												
Born 1987–1991	1.9	27.9	34.4	26.5	28.9	20.7	6.1	33.7	47.0	36.5	26.5	%
Born 1992–1996	48.9	51.2	51.9	50.8	49.7	53.6	50.7	51.0	50.7	49.9	50.9	%
Born 1997–2001	49.2	20.8	13.7	22.8	21.4	25.7	43.2	15.4	2.3	13.6	22.6	%
Proximate determinants												
Multiple birth	5.8	2.9	3.8	3.2	4.0	3.8	3.5	3.4	2.8	4.9	3.6	%
Male	50.4	50.9	49.5	50.0	50.6	51.4	50.9	50.9	50.8	50.3	50.5	%
Mother's age at birth												
Younger than 20	15.5	16.5	22.1	21.1	14.1	20.1	20.0	21.3	15.2	12.9	18.5	%
20–34	71.0	67.6	66.0	66.4	69.8	67.2	65.2	66.1	68.6	71.7	67.5	%
35 or older	13.5	15.9	11.9	12.5	16.1	12.7	14.8	12.6	16.2	15.5	14.0	%
Birth order												
First birth	20.2	17.6	21.1	22.8	22.9	18.1	16.7	16.3	17.7	19.0	19.6	%
Second birth	17.3	15.7	17.2	17.3	19.4	17.0	15.3	14.0	15.9	17.7	16.7	%
Third or higher birth	62.5	66.7	61.7	59.9	57.8	64.9	67.9	69.7	66.4	63.3	63.7	%
Socioeconomic factors												
Household water source												
Piped	37.4	8.5	31.3	46.7	30.9	17.8	26.7	16.0	44.1	30.1	28.6	%
Well	43.2	86.2	31.4	45.3	36.2	47.0	68.3	75.7	52.8	40.2	53.4	%
Surface	13.8	4.7	36.3	8.0	31.8	34.6	4.9	2.6	2.0	28.5	16.4	%
Other	5.6	0.6	1.0	0.0	1.1	0.5	0.1	5.6	1.1	1.2	1.5	%
Toilet facility												
Flush toilet	1.1	0.3	4.6	8.3	4.2	1.7	4.9	0.9	6.9	0.0	3.7	%
Pit latrine	23.7	18.2	84.3	51.2	68.8	60.9	73.1	16.9	59.0	27.8	50.7	%
Basic pit	74.5	81.4	11.0	40.4	26.9	37.4	22.0	82.1	33.8	70.0	45.4	%
Other toilet	0.7	0.1	0.1	0.0	0.1	0.0	0.0	0.1	0.3	2.2	0.2	%
Floor												
Natural floor	45.5	76.4	60.1	25.7	18.7	58.6	83.1	86.5	43.1	30.8	53.1	%
Rudimentary floor	0.5	0.0	0.2	0.4	0.1	1.3	0.0	0.0	0.0	0.2	0.2	%
Finished floor	53.9	23.3	39.7	73.9	81.3	40.0	16.8	13.1	56.7	68.7	46.5	%
Other floor	0.1	0.3	0.0	0.0	0.0	0.0	0.0	0.4	0.2	0.3	0.1	%
Household assets												
Electricity	15.5	3.8	36.4	45.2	32.0	13.7	9.6	5.8	28.0	10.9	22.2	%
Radio	76.8	61.9	55.6	66.5	48.7	58.2	73.9	37.4	70.3	54.5	59.3	%
Television	13.4	4.5	18.2	28.8	17.0	8.9	15.5	4.7	21.3	11.6	15.0	%
Refrigerator	3.9	1.8	9.6	14.0	10.9	6.2	4.6	2.2	10.8	3.2	7.3	%
Mother's education												
No education	75.4	92.3	37.7	66.7	40.3	87.0	85.5	89.3	78.1	62.6	69.7	%
Some primary	16.6	3.6	24.6	13.1	16.9	5.7	9.0	3.7	7.5	27.3	12.3	%
Completed primary	1.3	2.2	14.4	12.8	3.6	1.6	1.7	4.4	8.0	2.0	6.1	%
Some secondary or higher	6.6	2.0	23.3	7.3	39.2	5.8	3.8	2.6	6.4	8.1	12.0	%
Spatial variables												
Urban	30.5	9.9	27.2	31.9	25.2	24.9	22.2	16.2	33.5	23.5	23.5	%
Density within 30 km	262.9	72.8	105.4	343.1	301.4	91.3	56.3	59.8	947.6	196.0	225.4	mean
Distance to coast	190.2	758.6	417.0	196.7	169.6	203.2	792.3	877.2	76.6	199.0	417.1	mean
Distance to place of 50,000 population	37.5	73.9	34.5	40.7	28.8	44.5	76.4	63.2	29.9	29.5	47.3	mean
Arid zone												
Nonarid	24.9	0.0	71.2	68.7	59.8	93.4	0.9	0.0	30.0	75.0	41.7	%
Dry subhumid	71.0	12.6	6.6	31.3	23.0	6.6	25.6	0.0	7.7	13.2	18.9	%
Semiarid	4.1	87.4	21.2	0.0	17.2	0.0	64.8	92.9	50.8	11.9	36.8	%
Arid/hyperarid	0.0	0.0	1.0	0.0	0.0	0.0	8.7	7.1	11.4	0.0	2.6	%
Farming system												
Tree crop	0.0	0.0	32.4	57.5	58.6	1.6	0.0	0.0	0.0	18.8	22.0	%
Coastal artisanal fishing	36.6	0.0	3.2	21.3	26.0	18.8	0.0	0.0	0.5	20.1	11.9	%
All others	63.5	100.0	64.4	21.1	15.4	79.7	100.0	100.0	99.5	61.1	66.1	%
Average daily rainfall (mm)	3.1	2.0	4.4	3.9	3.3	5.2	2.0	1.3	1.8	3.1	3.0	mean
Malaria Stability Index	17.8	32.4	14.3	23.2	25.6	20.6	29.4	24.1	18.7	24.2	23.5	mean
Growing season (months)												
Less than 3	0.0	3.1	1.8	0.0	0.0	0.0	24.1	76.3	15.7	0.0	12.6	%
3–4	0.0	21.5	4.5	0.0	0.0	0.0	12.1	21.5	41.5	0.0	9.6	%
6–8	74.0	75.4	26.9	6.6	20.8	72.3	63.8	2.2	42.8	45.7	38.0	%
8–10	26.1	0.0	30.1	29.5	46.6	18.3	0.0	0.0	0.0	54.4	20.0	%
More than 10	0.0	0.0	36.8	63.9	32.7	9.4	0.0	0.0	0.0	0.0	19.8	%

tend to shorten the subsequent birth interval. Rather than divert attention from the primary concern of this study, we decided to avoid the measurement and modeling issues that would have been raised by the inclusion of the prior and/or subsequent birth intervals.

3.1.3 Spatial Variables [Model 3]

The spatial variables included here fall primarily into two types: those describing an urban-rural continuum, and those describing climatic parameters.

The classic urban-rural indicator is the usual dichotomous classification given during the enumeration phase of the survey implementation. The two additional measures considered are average population density within 30 kilometers and the distance to the nearest populated settlement of 50,000 persons or more (Balk et al., 2003), described above. We used a buffer of 30 kilometers to smooth differences in the density of population information, as well as in the distribution of surveyed households around a given cluster point. A DHS survey cluster is a representation of a group of households whose boundaries may or may not coincide with a census enumeration unit. The primary sampling unit for DHS surveys, the cluster, is usually a census enumeration area (EA). In rural or sparsely populated areas, density within these clusters may be highly varied. In many countries, rural clusters may contain more than one village, and they may be geographically large. In urban areas, EAs are often geographically small but more densely populated. Urban EAs are usually segmented during the listing process (see Macro International, 1996 for further details). Fifty thousand was chosen as the city population threshold because data on cities of that size are more consistently available than for smaller cities. All spatial indicators were chosen ignoring national boundaries in determining the 30-kilometer radius and the nearest city. While borders have obvious political and economic effects, they are less likely to impede disease vectors. The effect of specific borders is an open question beyond the scope of this report.

We explored five measures of climate—rainfall, aridity, farming systems, length of growing season, and the stability of malaria transmission.

Twenty-four percent of the births were in urban areas; but the average population density of urban births is 665 persons per square kilometer. One thousand persons per square kilometer is a conventional minimum for urban areas, although perhaps more applicable to North America and Europe than Africa (Rain, n.d.).

Additionally, a variable for the shortest distance to the coast is included. This variable has been shown to be an important correlate of economic development (Sachs et al., 2001) as a proxy variable for access to goods and services on the global market, trading potential, and so forth. We included it to determine whether there is evidence for a similar effect on mortality. As Table 3 indicates there is considerable variability in the national averages of this variable.

We have adopted several measures of climate in part because no single measure is expected to capture the inherent complexities. We explored five measures—rainfall, aridity, farming systems, length of growing season, and the stability of malaria transmission. Theory suggests that excess dryness or wetness will increase the risk of mortality. In dryer areas, increases in rain will be expected to improve child survival by providing sources of water, inputs to agricultural production, and improved sanitation. In wetter areas, excess rain may reduce crop yield (due to pests present only in very wet areas) and provide a more fertile vector habitat.

Rainfall, as shown in Maps 8 and 9, clearly has wide temporal and spatial variation in West African countries. Two regimes are most prominent. In most of the survey region August is the wettest month. In southern and central Côte d'Ivoire, Ghana, Togo, Benin, and Cameroon, there are two rainfall peaks—one in May/June and the second in August/September. January/December is the driest period in nearly all areas, although similarly low rainfall extends from November to March in Burkina Faso, Mali, Niger, and Senegal.

Aridity, another measure of dryness, combines precipitation and evapotranspiration rates into distinct classes. However, aridity zones proved to be too problematic to use. Nomadic populations in Mali and entire *arrondissements* (districts) and the rural population of other districts in Niger were excluded from the sampling frame. This is noteworthy because these areas contribute disproportionately to the population of arid and hyperarid zones. Only four countries have clusters in the arid and hyperarid zones, and a disproportionate number of those are in Mali. Mali is the only country contributing to all four zones, and four countries contribute to only two zones (Burkina Faso, Côte d'Ivoire, Guinea, and Niger). Additionally, aridity is highly correlated with rainfall (–0.79), such that inclusion into the model with rainfall would overspecify it. Although we did some preliminary analysis of this variable, we omit it from further treatment here due to sampling concerns and the availability of substitute measures.

Farming systems, delineated in Dixon et al. (2001), provide an indication of the likely potential of the agro-climatic zone. Delineations are coarse, however, and cannot be considered accurate indicators of food supply or employment type for surveyed households or for surrounding communities. Preliminary analysis with a limited model including all ten farming systems present in the study region indicated that two systems, coastal artisanal fishing and tree crops, showed the strongest relationship with mortality.

The length of the growing season has long been associated with agricultural productivity (FAO, 1978). Seasons of fewer than 70 days are considered too short for sustainable agriculture, and long seasons, of greater than 300 days, are considered not optimal because the excess rain fosters pests that damage crops. The range of 120–240 days is considered good, with 240–300 days being considered optimal. The variable is so highly correlated with rainfall, at 0.84, that we could not add it to the multivariate model while also controlling for rain.

One further variable, closely related to climate, was an index of the stability of malaria transmission (Kiszewski et al., forthcoming). This data set is constructed by "incorporating published estimates on the proportion of blood meals taken from human hosts, daily survival of the vector, duration of the transmission season, and extrinsic incubation" (Kiszewski et al., forthcoming, 1). However, because the majority of inputs to the index are at a relatively coarse resolution, it interacted too strongly with the country variable and was therefore omitted from the main models.

Variables associated with land cover and land use were also omitted. These variables might include land cover classification or land use and vegetation indices (e.g., Normalized Difference of Vegetation Index [NDVI]). Such variables might serve as proxies for vector habitats and ecological factors influencing agro-pastoral economic life. Several possible datasets were considered for use, but all were too complex to be introduced in a systematic and rigorous way in the short term. Thus, rain and selected farming systems were the only variables ultimately selected for inclusion in the multivariate model.

3.1.4 Socioeconomic Variables [Model 4]

Socioeconomic variables included in the analysis reflect the household environment and the household assets. The distinction between household environment and assets is somewhat arbitrary because environmental characteristics may be heavily influenced by assets—that is, a household's ability to afford high-quality water, sanitation, and flooring, or the community's capacity to fund public infrastructure (e.g., to provide safe water and sewer services). The distinction is maintained nevertheless because environmental characteristics are often somewhat exogenous to the household (e.g., community-level services) and because they may directly mediate vectors of disease transmission or otherwise influence the level of contamination in the child's home environment. Household environment variables included in the analysis were the source of drinking water, type of toilet, and type of flooring. Piped water, modern toilet facilities, and finished flooring are believed to improve chances of survival by minimizing contamination. Their effects are expected to be significant, especially for older infants and children who are more exposed to them through drinking the water, crawling or playing on the floor, and using the toilet. The effects for the latter may also affect young infants by indirect exposure to contamination via the mother using unsanitary toilet conditions.

Household assets in the model include electricity, radio, television, and refrigerator. These are indicators of the socioeconomic status of members of a household. Households with higher socioeconomic status (more assets) are believed to have a positive impact on infant and child survival. We experimented with combining them into a single assets index but found it more informative to retain separate variables. As expected, households with these assets were more likely to be households in which women had a higher than average level of education. The exception was in the case of radios; the majority of households possessed a radio regardless of the mother's level of education.

Data on the mother's current partner's education and occupation, although important socioeconomic indicators, were omitted from the analysis because the information was either not available or not comparable for all countries included. Similarly, while mother's marital status (including informal union) is an important predictor of mortality, it was not included in this analysis because, since it shows so little variation, it is of limited use as a measure of current status.

Piped water, modern toilet facilities, and finished flooring are believed to improve chances of survival by minimizing contamination. Their effects are expected to be significant.

3.1.5 Omitted Variables

Other potentially important types of variables were omitted from this study, notably variables relating to nutritional and health status. While information on breastfeeding and young infant feeding are collected in DHS surveys, it is only for a subsample of children born in the three years preceding the survey. Likewise, for anthropometric information, only children born in the five years preceding the survey had their weight and height measured. Similarly, basic health information about recent episodes of diarrhea, cough, and fever were available for children under age five. In Cameroon, Niger, and Togo, these anthropometric and recent disease episode data were collected for children age 0–3 years.[5] As stated earlier, this study includes a much larger sample of children born in the ten years preceding the survey. Furthermore, the nutrition and health data were collected only for children who were living at the time of the survey, thereby excluding the data for children who died during the same period.

Analyzing the risk of malaria transmission to child survival is limited to a bivariate examination. It could not be considered in the full multivariate model because of issues of multicollinearity and specification. No other disease transmission factors are considered because of data constraints.

3.2 Survival Models

An analysis of selected survival functions served to model the distribution of deaths over time stratified by selected covariates. The nonparametric Kaplan-Meier (product-limit) method was used to generate maximum likelihood estimates of $S(t)$, the probability that death occurs at an age greater than t.

Survival distributions were generated using the SAS 8.2 Lifetest procedure. By incorporating information on age at death, the distribution curves demonstrate the differential pace and level of mortality for infants and children. DHS data provides age at death in months for children under age 24 months and age at death in years for older

[5] Senegal is the other exception: No anthropometry measures were collected there and only recent diarrhea (not cough and fever) is included for children age 0–5 years.

children. A quantitative evaluation of the stratified survival curves at age 12 months (for infants) and 59 months (for children) highlights the cumulative impact of these factors on the two age groups.

We stratified survival distributions by selected factors hypothesized to influence childhood mortality. The survival distribution estimates can be compared visually or by log-rank statistics that adjust for stratum scores and test for homogeneity of strata. The survival curves reveal initial confirmation of expected findings from both individual-level factors (such as maternal education) and environmental factors (such as population density). It is important, however, to recall that in this part of the analysis no other factors have been controlled. An analysis of the independent effects of these factors on infant and child mortality (i.e., controlling for an ensemble of other determinants) is presented in Part 4.

3.3 Generalized Linear Model

For the multivariate analysis, we used a generalized linear model (GLM). This technique (cf. McCullagh and Nelder, 1989) is similar to a hazard model or a survival analysis (cf. Namboodiri and Suchindran, 1987) but produces coefficients that are more analogous to the usual ${}_1q_0$ and ${}_4q_1$. The computer analyses were done with the GLM procedure in Stata, versions 7 and 8. A brief description of the modeling strategy follows, as implemented for infant mortality; similar logic applies to deaths among children age 1–4.

At the level of the individual child, we define a binary outcome, *died0*, coded 1 if the child died before reaching exactly 12 months (one year) of age and 0 if it survived. We also code a measure of exposure to the risk of dying, called time0, which can be between 0 and 1. If the child was observed to die any time in the first year of life, or was observed to survive the full first year, *time0* is coded 1. However, if the case was censored (i.e., the child was born during the year before the survey, and was still alive at the time of the survey), then *time0* is the fraction of the year for which the child was observed. Then, for a given sample of children, the standard estimate of ${}_1q_0$ will be equivalent to the sum of *died0* for those children, divided by the sum of *time0*.

An individual-level statistical model that gives this same estimate will be a generalized linear model with outcome *died0*, a binomial error distribution with binomial denominator *time0*, and a log link function. When this model is run with no covariates, the output will produce a constant which, if exponentiated, will be the estimate of ${}_1q_0$. When covariates are included, the exponential of the constant term will be a fitted ${}_1q_0$ for the reference combination of the covariates. The exponential of a coefficient for a covariate will be the relative risk for that covariate. For example, Table 4 gives the coefficients (before and after exponentiation) for the covariate "country," a categorical covariate; the reference country is Ghana.

A quantitative evaluation of the stratified survival curves at age 12 months (for infants) and 59 months (for children) highlights the cumulative impact of these factors on the two age groups.

In Table 4, all numbers except those in the last column come directly from the (Stata-generated) computer output. The last column is obtained by exponentiating the first column. The exponentiated constant term, 0.0605, is the estimate of ${}_1q_0$ for Ghana, the reference (or "omitted" country). It is equivalent (when multiplied by 1000) to an infant mortality rate (IMR) of 60.5 deaths per 1000 births. The report on the 1998 Ghana survey (GSS, 1999: 83) gives an IMR of 56.7 for 0–4 years before the survey and 65.8 for 5–9 years before the survey. Our estimate of 60.5 for 0–9 years before the survey is consistent with those values. The exponentiated coefficient for Burkina Faso, for example, is 1.7664, meaning that its ${}_1q_0$ is $0.0605 \times 1.7664 = 0.1069$. This ${}_1q_0$ is about 77 percent $[(1.7664 - 1) \times 100 = 76.64)]$ higher than the ${}_1q_0$ for Ghana.

In the tables, ** after an exponentiated coefficient indicates that it is significantly different from 0 in a two-tailed 0.01 test or one-tailed 0.005 test; * indicates significance at the two-tailed 0.05 or one-tailed 0.025 level, and # indicates significance in a two-tailed 0.10 or a one-tailed 0.05 level. We use the # symbol and refer to one-tailed

Table 4
Results of GLM model for infant mortality by country

Country	Coefficient	Robust standard error	z	Exponentiated coefficient
Burkina Faso	0.568964	0.072360	7.86	1.7664**
Benin	0.427995	0.072815	5.88	1.5342**
Côte d'Ivoire	0.600754	0.102399	5.87	1.8235**
Cameroon	0.262742	0.084478	3.11	1.3005**
Guinea	0.596614	0.071421	8.35	1.8160**
Mali	0.741359	0.067393	11.00	2.0988**
Niger	0.743393	0.070315	10.57	2.1031**
Senegal	0.152848	0.075390	2.03	1.1651*
Togo	0.263117	0.073111	3.60	1.3010**
_cons	-2.805231	0.061622	-45.52	0.0605**

Note: The estimates in Table 4, as in all other tables in this report, are weighted (see section 2.1 for more detail on the weights), with robust estimates of the standard errors that take into account the cluster design of the data. Clustering at the household level is not taken into account.

tests because many potential hypotheses about mortality differentials are indeed one-tailed rather than two-tailed. Significance levels are determined from the z column of the computer output (the ratio of the coefficient to its standard error) and describe the significance of the difference from the reference category.

We have used a log probability model because of the familiarity of ${}_1q_0$ and ${}_4q_1$ to all demographers, but some analyses of infant mortality use logit regression, another generalized linear model. In logit regression, it is the logit of the probability of a death, rather than the log of the probability, that is linear in the predictors. In logit regression, exponentiated coefficients are interpreted as relative *odds*, rather than as relative risks. Hazard or survival models are linear in the log and are also similar, but in those models the probability of death refers to an instantaneous rate of change in the survivorship function, rather than the change from exact age 0 to 1 and from exact age 1 to 5.

For all of these models, the estimated probability of dying must be less than one for every case. Logit and hazard models are constructed in such a way—through the logit link and the instantaneous rate of change, respectively—that this condition is always satisfied. In our data, the maximum predicted probability of dying is always less than one (the maximum is about 0.80), but this is an empirical result and for other data sets or age intervals the log link function might not be usable.

4 Results

In general, we anticipate that the risks an infant faces during birth and the first month of life are very different from those faced after this period. Infant deaths are more closely linked to endogenous factors that are difficult to prevent (e.g., congenital malformations, hereditary diseases, and low birth weight). Older children are more likely to die from preventable diseases including infectious diseases and malnutrition. This pattern occurs because older children are more mobile and, in interacting with their environment, are more exposed to contamination in the air, water, and food. For these reasons, we anticipate that proximate factors act more strongly on infant mortality and socioeconomic and spatial factors act more strongly on child mortality.

The survival curves, which show a clear bivariate picture of mortality differentials, confirm these associations (see Figures 1–10 and Table 6). The multivariate analysis that follows shows a somewhat more complex picture. Overall, there is confirmation of conventional factors and support for inclusion of many of the spatial factors.

4.1 Survival Analysis

Table 5 provides a summary of the Kaplan-Meier estimates for strata in each covariate. With the exception of sex of children age 1–4 years, strata for all variables shown here have significantly different survival functions. Figures 1–10 show the survival density functions for infants and children by covariates.

Among the most important proximate determinants that influence child survival are the mother's age at birth and the birth order of the child (Sullivan et al., 1994). These maternal factors have differential impacts on infants and children: infant deaths among mothers under age 20 typically occur in early infancy; young motherhood has less impact for children age 1–4 years (Figure 1). Birth order is closely related to mother's age at birth. In sub-Saharan Africa this is due in part to early age at marriage and the resulting early age at initiation of childbearing. Similar to infants of young mothers, first births are less likely to survive infancy than higher order births. Likewise, the impact of birth order on survival is greatly reduced for children age 1–4 years (Figure 2). Multiple births face a much higher risk of death, especially during infancy (Figure 3).

Maternal education has been observed to have a strong impact on child survival. Unlike maternal factors that have a differential impact on infant and child survival, education is a socioeconomic characteristic that influences both age groups. Infants and children of mothers with no education both have only an 89 percent chance of survival at 12 months and at 59 months (Figure 4). Infants and children of mothers with secondary or higher education have greatly improved chances of surviving, 95 percent and 97 percent, respectively.

Infants and children residing in urban areas have, on average, better survival chances than those in residing in rural areas. This advantage is usually assumed to be related to better infrastructure and access to services. When the survival curves of residence are overlaid with population density classes, the subtleties often disguised by the dichotomous urban-rural variable are exposed (Figure 5). While it is still clear that infant

Table 5
Summary table of S(t) for selected covariates

Characteristic	S(t) at 12 months	S(t) at 59 months
Mother's age at birth		
Younger than 20	86.9	89.3
20–34	91.0	91.2
35 or older	90.6	91.3
Birth order		
First birth	88.4	91.3
Second birth	91.0	91.2
Third or higher birth	90.5	90.7
Birth cohort		
Born 1987–1991	90.0	90.3
Born 1992–1996	90.1	91.1
Born 1997–2001	91.0	92.1
Mother's education		
No education	88.9	89.2
Some primary	91.5	93.1
Completed primary	92.3	94.1
Some secondary or higher	94.9	96.6
Multiple birth		
Single birth	90.9	91.2
Multiple birth	71.0	86.5
Sex of child		
Female	90.9	90.9
Male	89.4	90.9
Residence and density		
Rural	89.3	89.8
Urban	93.0	94.4
<25 per sq km	87.6	89.1
25–100 per sq km	89.4	89.3
100–500 per sq km	92.3	93.6
500–1000 per sq km	94.7	96.0
>1000 per sq km	93.9	96.5
Distance to city of 50,000 population		
In urban area	92.5	93.9
1–25 km	91.8	92.3
25–100 km	90.0	90.1
100–150 km	88.3	89.3
>150 km	87.6	87.6
Rainfall		
<2 ml per day	88.6	86.5
2–4 ml per day	90.6	92.2
>4 ml per day	91.2	93.4
Farming		
Tree crops	92.4	94.8
Root crops	87.9	90.9
Cereal/root crops	89.4	89.5
Agro-pastoral	89.3	87.7
Fishing	92.9	95.2
Growing season (months)		
Less than 3	87.0	84.1
3–4	89.8	87.8
4–8	89.7	90.4
8–10	91.3	93.9
More than 10	92.1	94.5
Malaria		
Low index	91.2	92.3
Medium index	90.2	91.4
High index	88.6	87.2

Note: The Chi-square for the log-rank statistics were strongly rejected at $p < 0.0001$ for all variables except sex of children age 1–4 years ($p < 0.0537$).

mortality is higher in rural areas than in urban areas, in rural areas there is a density continuum indicating that infants living in the most sparsely populated areas (less than 25 inhabitants per square kilometer) suffer the lowest probability of survival. These very sparse areas may have the least adequate infrastructure to support prenatal and delivery services. Similarly, although infants generally enjoy greater chances of survival in urban areas, for infants who live in the most densely populated areas (more than 1000 persons per square kilometer) survival chances appear to be compromised. This is likely a reflection of overcrowded or slum conditions where, similar to remote rural areas, maternal services are inadequate (Defo, 1994; Gupta, 1999; and Woods, 2003).

Compared with infants, the survival pattern of children reveals a continuum of population density that is more closely clustered around rural residence (Figure 5). This

Figure 1
Mother's age at birth

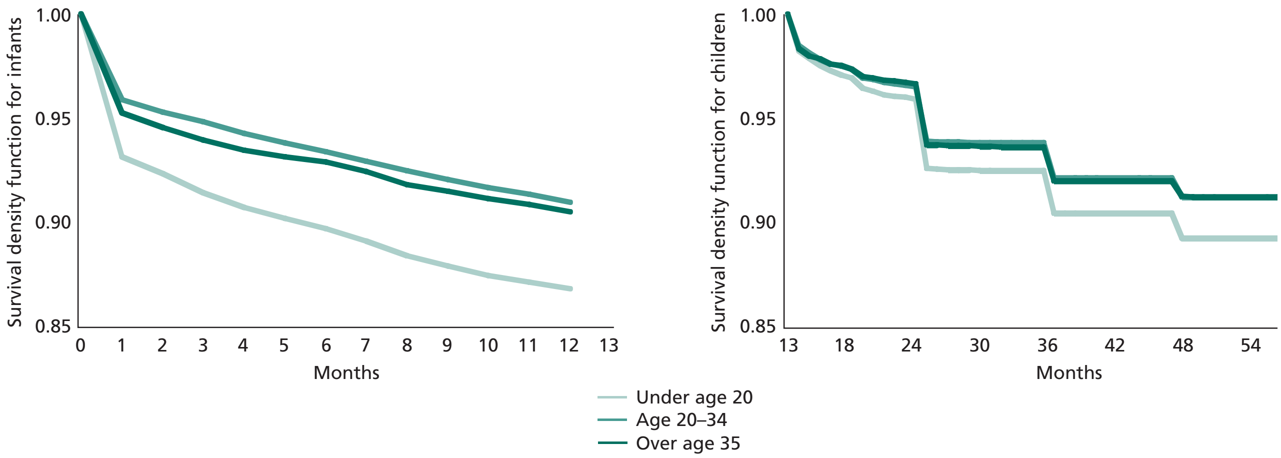


Figure 2
Birth order

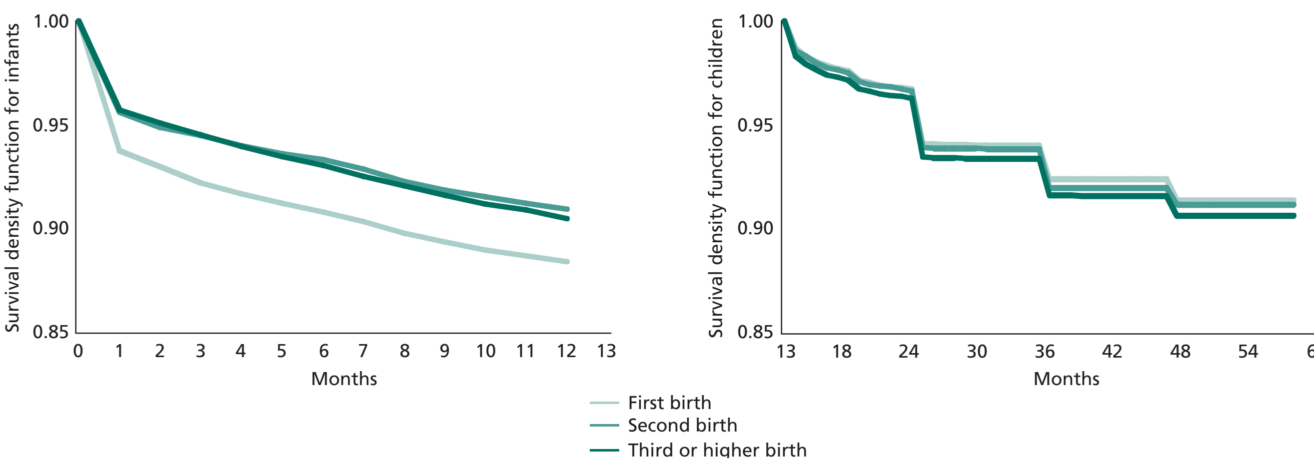


Figure 3
Multiple births

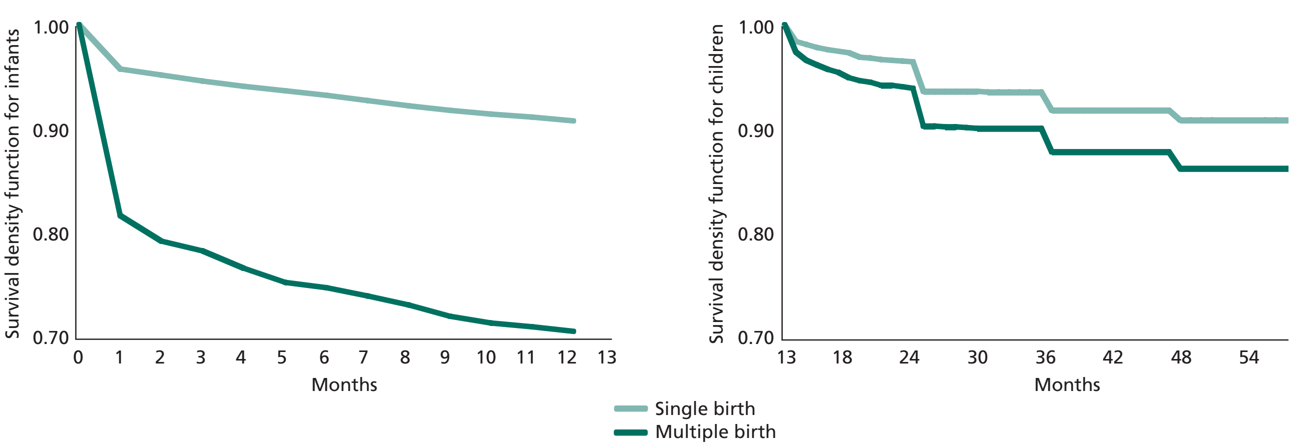


Figure 4
Mother's education

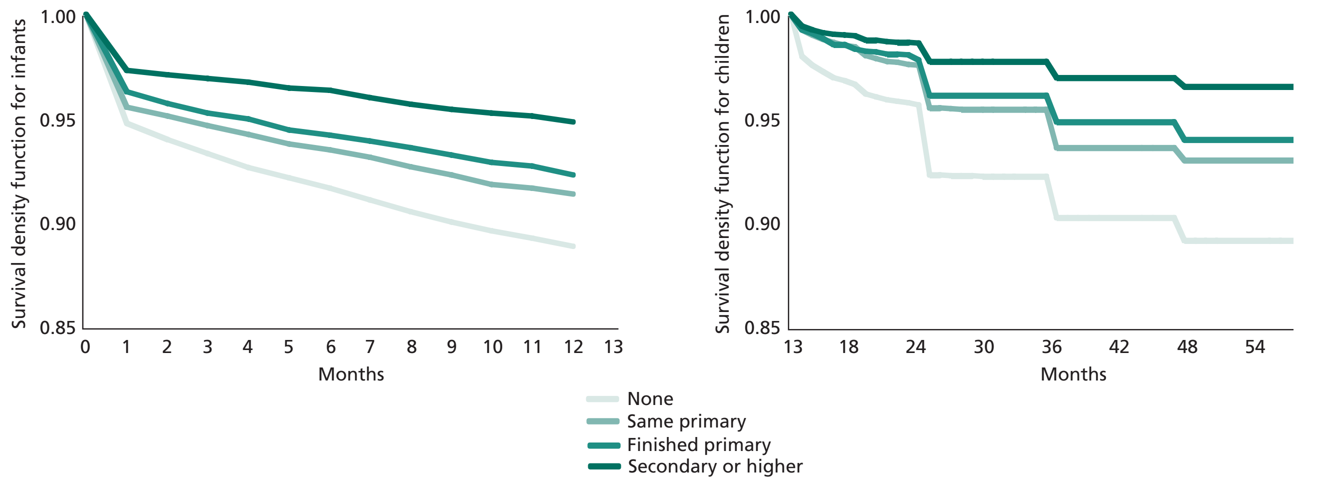
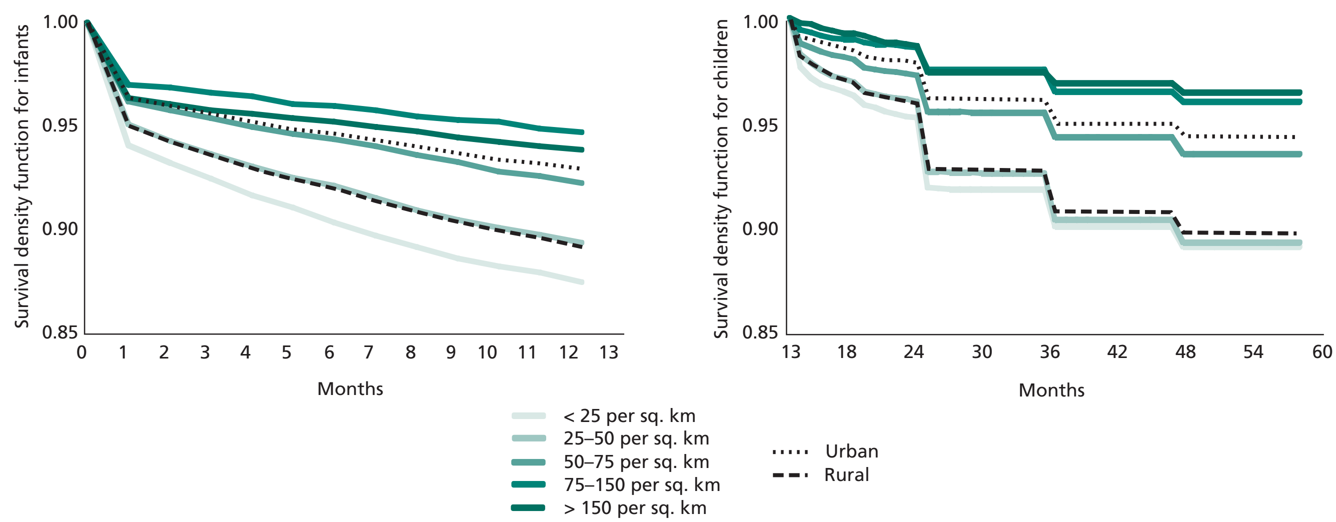


Figure 5
Density per square kilometer and residence



suggests that if children survived infancy in the most sparsely populated areas then, despite the measure of sparseness, they are equally likely to survive to their fifth birthday. Children in the sparsest settings, although having greater risk of death than children in urban, more densely populated areas, still have better chances of survival than infants in the sparsest settings. For both infants and children, mortality increases monotonically the further one resides from an urban area (Figure 6). The negative effect of the highest population densities on infants does not have a parallel in the distance measure, perhaps because the highest density areas cannot be distinguished from slightly lower density areas in urban areas.

Variation in average daily rainfall has a greater impact on children age 1–4 than on infants. One explanation for this is that their dietary needs are varied and dependent on agricultural production, whereas an infant's dietary needs are met by breastfeeding. Figure 7 shows that for children living in areas with an average of less than 2 ml of rainfall daily, the probability of survival after 59 months is 86.5 percent. In comparison, children living in areas with higher average daily rainfall have a 92–93 percent chance of surviving after 12 months. Figure 8 shows similar patterns for the length of the growing season on infant and child survivorship, with the lowest survival rates for children

Children in the sparsest settings, although having greater risk of death than children in urban, more densely populated areas, still have better chances of survival than infants in the sparsest settings.

living in arid and semiarid zones, that is, those with the shortest (less than 3 months, and 3–4 months) growing seasons. Even children in the main agricultural band of 4–8 months have lower chances of survival than children in the most sustainable regime (8–10 months). The effect on infants is weaker, although the least advantaged growing season (i.e., the arid zone) stands apart from the others. Figure 9 shows that farming systems differ considerably in both infant and child mortality.

Malaria transmission factors are important in child survival. We stratified the malaria stability index into three categories corresponding to the 20 percent highest, 20 percent lowest, and middle 60 percent of transmission likelihood (Figure 10). The impact is in the expected direction for both age groups, that is, the stratum with a high transmission index has a faster pace of mortality than the low and medium transmission groups. However, the impact of a high transmission index appears to be more intense for children age 1–4 years than for infants, perhaps because older children are more likely to be exposed to repeated malarial infections that contribute to the development of other diseases that increase the risk of death, such as severe anemia (Menendez et al., 2000; Slutsker et al., 1994). Further analysis is needed to determine if this trend persists when other factors are controlled. For reasons detailed below, this analysis cannot be undertaken here.

Figure 6
Distance to city

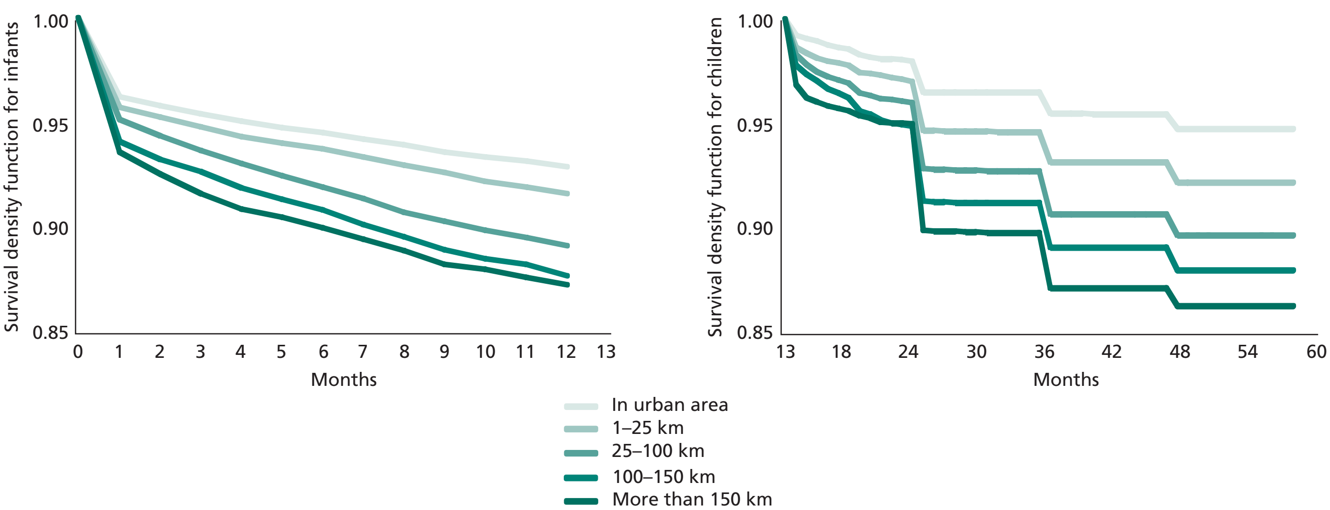


Figure 7
Average daily rainfall

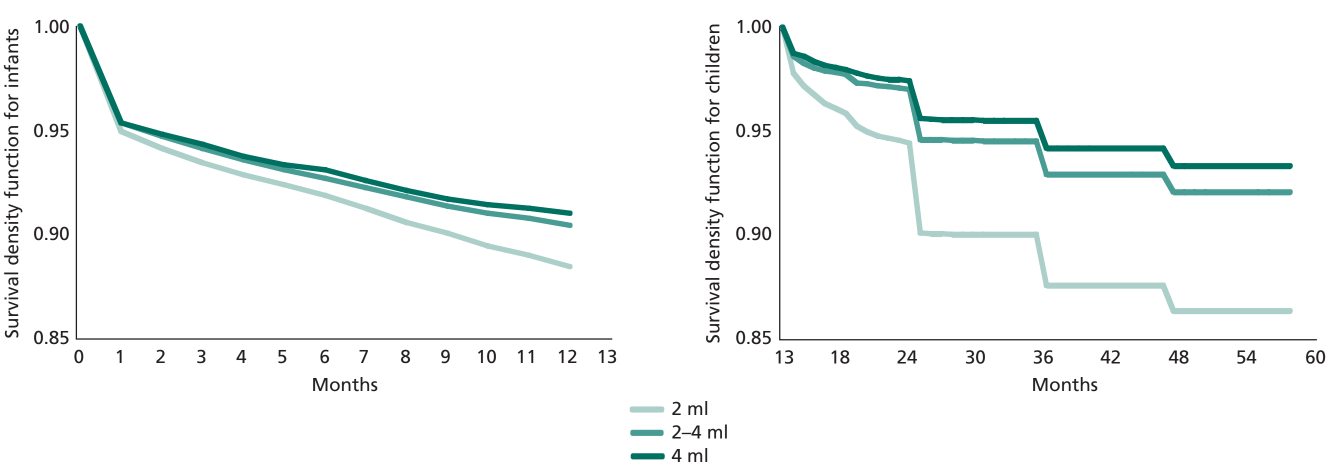


Figure 8
Growing season

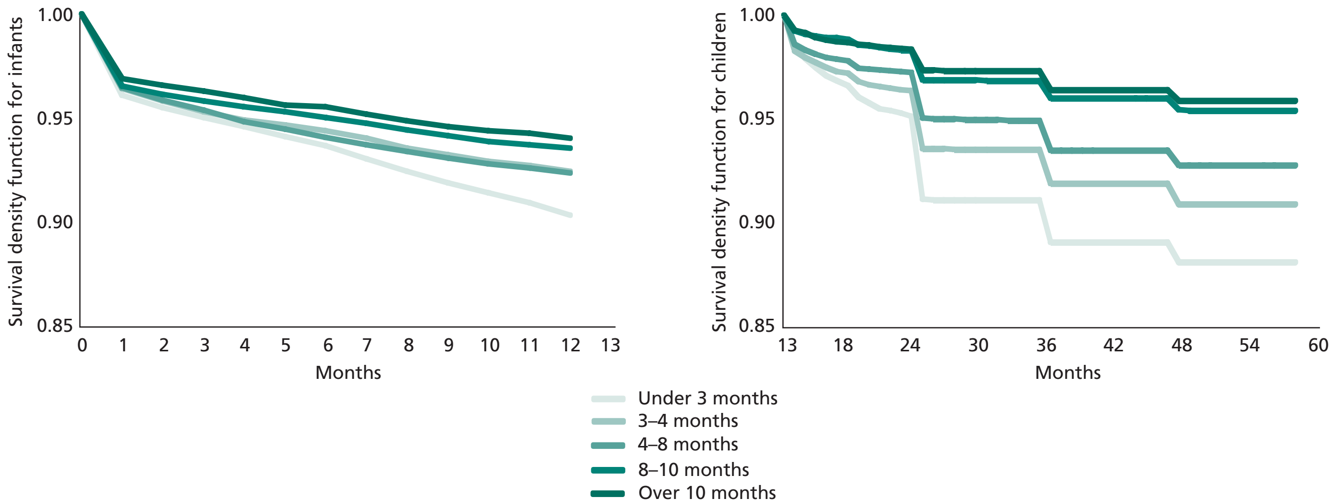


Figure 9
Farming system

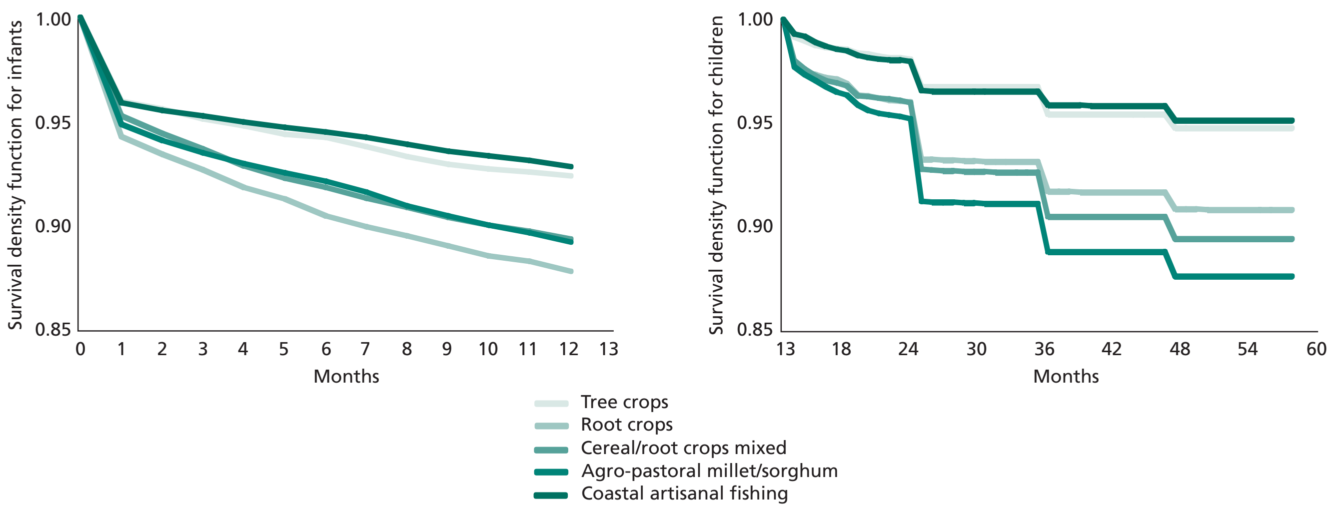
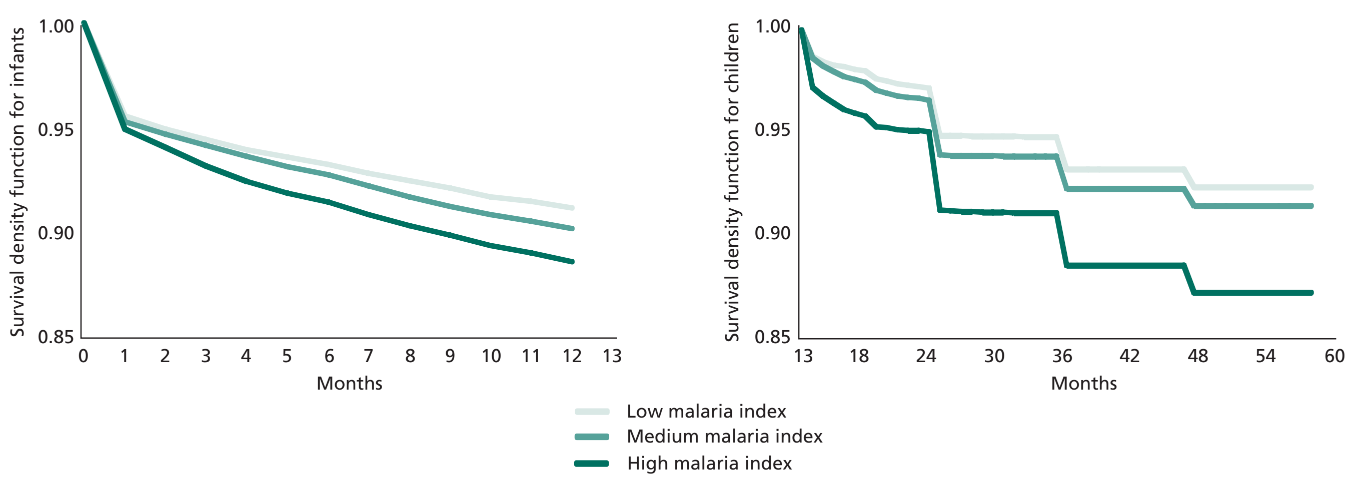


Figure 10
Malaria index



4.2 Generalized Linear Model

Tables 6 and 7 present the results from a series of five GLM or log probability models applied to ages 0 and 1–4, respectively. The variables in the five models may be summarized as follows.

Model 1: Country and time period. This is a baseline model; country and time period are largely interpreted as control variables and are included in all subsequent models. There is wide variation in the coefficients in this model, and one goal of the subsequent models is to explain or reduce this variation.

Model 2: Model 1 plus four demographic characteristics of the child and mother: sex, multiple birth, birth order, and age of mother. These four variables are included in all subsequent models, and as expected their coefficients are quite consistent across models.

Model 3: Model 2 plus household characteristics. These include source of water; type of toilet; type of floor; whether the household has electricity, radio, television, refrigerator; and mother's education. These would be the standard kinds of variables in a model for infant or child mortality. Note that this model does not include the urban-rural classification, which is available in the DHS data but which we regard as a spatial variable.

Model 4: Model 2 plus spatial characteristics. These include the urban-rural classification, population density (taken as the natural log thereof), rainfall (both linear and quadratic terms), distance to coast, and a three-category version of farming system.

Model 5: Model 3 plus spatial characteristics. A comparison of this model with model 3 provides our best evidence of the additional explanatory value of spatial variables, above and beyond the standard model.

We now turn to a systematic discussion of the results in Tables 6 and 7.

4.3 Discussion

In Table 6, which gives the models for age 0, the country effects (expressed as ratios to Ghana's infant mortality) largely become insignificant after the household and spatial characteristics have been included. In model 1, eight countries have significantly (at the 0.01 level) higher infant mortality than Ghana, but by model 5 only three countries meet this criterion: Côte d'Ivoire, Mali, and Niger. Most of the change can be attributed to the addition of the spatial variables, as can be seen by comparing model 4 with model 2 or comparing model 5 with model 3. The countries that change the most with the addition of these variables—that is, the countries whose higher mortality can be most strongly attributed to unfavorable spatial characteristics—are Burkina Faso, Cameroon, Guinea, Mali, and Niger. The effects for time periods are small in all models, although the second time period achieves a low level of significance on some models.

In all models, males have about 17 percent higher mortality than females.

The proximate determinants behave in ways consistent with well-established effects in the literature. In all models, males have about 17 percent higher mortality than females, children in multiple births are about 3.5 times as likely to die as singletons, second and later births have a risk that is 10 to 20 percent less than first births, and children born when the mother is age 20 or above have a risk that is 25 to 30 percent less than when the mother is less than age 20. These differences are all highly significant, but are regarded mainly as controls here.

The eight household-level variables appear in models 3 and 5, and their coefficients

Table 6
Log probability models for $_1q_0$

	Model 1	Model 2	Model 3	Model 4	Model 5
Country					
Benin	1.5614**	1.4959**	1.2323**	1.3174**	1.1831*
Burkina Faso	1.7896**	1.7972**	1.3070**	1.1922#	1.1168
Cameroon	1.3156**	1.2707**	1.1900*	1.0187	1.0085
Cote d'Ivoire	1.8328**	1.7998**	1.6770**	1.7223**	1.6284**
(R)Ghana					
Guinea	1.7887**	1.7391**	1.3924**	1.3771**	1.2067#
Mali	2.1042**	2.0787**	1.6528**	1.3896**	1.3473**
Niger	2.1857**	2.1481**	1.5781**	1.4665**	1.3814**
Senegal	1.1434#	1.1667*	1.0112	1.1852	1.0477
Togo	1.3159**	1.2947**	1.0796	1.1466#	1.0421
Birth cohort					
(R)1987–1991					
1992–1996	0.9503	0.9489#	0.9369*	0.9502	0.9407*
1997–2001	0.9619	0.9547	0.9473	0.9535	0.9503
Sex of child					
(R)Female					
Male		1.1695**	1.1731**	1.1702**	1.1726**
Multiple birth					
(R)Single birth					
Multiple birth		3.4267**	3.5405**	3.4353**	3.5254**
Birth order					
(R)First birth					
Second birth		0.8458**	0.8055**	0.8262**	0.8048**
Third or higher birth		0.9036*	0.8041**	0.8507**	0.8016**
Mother's age at birth					
(R)Younger than 20					
20–34		0.6931**	0.7417**	0.7242**	0.7452**
35 or older		0.7242**	0.7523**	0.7388**	0.7533**
Household water source					
(R)Piped					
Well			1.0772#		1.0418
Surface			1.081		1.0566
Other			0.9431		0.9309
Toilet facility					
(R)Flush toilet					
Pit toilet			1.3355**		1.3080**
No toilet			1.4462**		1.3461**
Other toilet			1.1869		1.197
Floor					
(R)Natural floor					
Finished floor			0.8621**		0.9011**
Other floor			0.7555		0.7498
Household assets					
Electricity			0.9289		0.9744
Radio			0.9902		0.9916
Television			0.9311		0.952
Refrigerator			0.8557#		0.8492#
Mother's education					
No education					
Some primary			0.916		0.9469
Completed primary			0.7756*		0.8205*
Some secondary or higher			0.6940**		0.7364**
Rainfall					
Average daily rainfall				1.1292#	1.0996
Rainfall squared				0.9887	0.9924
Residence					
(R)Rural					
Urban				0.7218**	0.8722**
Log density				0.9576**	0.9716#
Distance to coast				1.0005**	1.0003*
Farming system					
Tree crop				0.7835**	0.8127*
Fishing				0.976	0.9817
(R)Other					
Constant	0.0643**	0.0803**	0.0792**	0.0867**	0.0777**
Number of cases	122389	122389	118701	122389	118701
df	11	17	32	24	39
Log lik.	-37485.04	-36551.4	-34776.94	-36286.14	-34722.08
Pseudo R^2	0.01042	0.03507	0.04432	0.04207	0.04583

Note: The exponentiated coefficients are the relative risk of dying at age 0.
significant at one-tailed .05 level
* significant at two-tailed .05 level
** significant at two-tailed .01 level
(R) reference category

Table 7
Log probability models for $_4q_1$

	Model 1	Model 2	Model 3	Model 4	Model 5
Country					
Benin	1.3177**	1.2924**	1.0261	1.0739	0.9462
Burkina Faso	2.4616**	2.4319**	1.6759**	1.2546*	1.2008#
Cameroon	1.5199**	1.4802**	1.4491**	1.0359	1.0711
Côte d'Ivoire	1.4193*	1.3939*	1.3995*	1.3874**	1.4467**
(R)Ghana					
Guinea	1.8177**	1.7725**	1.3410**	1.2846#	1.0544
Mali	2.2881**	2.2293**	1.7616**	1.1622	1.1923
Niger	3.5445**	3.4371**	2.3779**	1.6650**	1.6014**
Senegal	1.4892**	1.4771**	1.2689**	1.3018#	1.1544
Togo	1.3124**	1.3014**	1.0322	1.0874	0.9566
Birth cohort					
(R)1987–1991					
1992–1996	1.0469	1.048	1.0341	1.0347	1.0331
1997–2001	1.4311**	1.4368**	1.3815**	1.4140**	1.3811**
Sex of child					
(R)Female					
Male		0.9938	0.9958	0.995	0.9953
Multiple birth					
(R)Single birth					
Multiple birth		1.5850**	1.6567**	1.6097**	1.6493**
Birth order					
(R)First birth					
Second birth		1.0998#	1.0449	1.0704	1.0469
Third or higher birth		1.2035**	1.07	1.1295*	1.0742
Mother's age at birth					
(R)Younger than 20					
20–34		0.7473**	0.7929**	0.7806**	0.7916**
35 or older		0.7352**	0.7628**	0.7534**	0.7619**
Household water source					
(R)Piped					
Well			1.0742		1.0614
Surface			1.1258#		1.1516#
Other			0.7866*		0.7708*
Toilet facility					
(R)Flush toilet					
Pit toilet			1.3241#		1.3397*
No toilet			1.4539*		1.3721*
Other toilet			1.322		1.4098
Floor					
(R)Natural floor					
Finished floor			0.8963*		0.9308
Other floor			1.6601#		1.6629#
Household assets					
Electricity			0.7536**		0.7738**
Radio			0.9185*		0.9233*
Television			0.8943		0.8985
Refrigerator			0.8103		0.7882#
Mother's education					
No education					
Some primary			0.8305*		0.8918
Completed primary			0.7485**		0.8122*
Some secondary or higher			0.5803**		0.6446**
Rainfall					
Average daily rainfall				0.9478	0.9141
Rainfall squared				1.0111	1.0174*
Residence					
(R)Rural					
Urban				0.6323**	0.8808*
Log density				0.9844	1.0105
Distance to coast				1.0006**	1.0004*
Farming system					
Tree crop				0.6810**	0.6976**
Fishing				0.9234	0.9389
(R)Other					
Constant	0.0530**	0.0588**	0.0647**	0.0835**	0.0732**
Number of cases	98249	98249	95381	98249	95381
df	11	17	32	24	39
Log lik.	-24513.11	-24443.8	-23332.5	-24170.9	-23272.65
Pseudo R^2	0.0269	0.02972	0.045	0.04055	0.04745

Note: The exponentiated coefficients are the relative risk of dying at ages 1 to 4.
significant at one-tailed .05 level
* significant at two-tailed .05 level
** significant at two-tailed .01 level
(R) reference category

are almost identical in those two models, although they are closer to unity in model 5 because of some association with the spatial variables. Source of water is not significant; type of toilet is highly significant, with a flush toilet being by far the most beneficial category;[6] type of floor is highly significant, with a natural floor being the least desirable type. We would hypothesize lower mortality for households with electricity, radio, television, or a refrigerator, but only the last of these four achieves significance with a one-tailed 0.05 test. Finally, maternal education has a monotonic protective effect. Children whose mother had at least some secondary schooling have nearly a 30 percent advantage.

We can summarize the potential protective effect of the household-level variables by multiplying together the eight lowest values of these eight variables. For a child in the optimal category of every variable, the relative risk (compared with a child in the reference category of every variable) would be 0.9309 × 1 × 0.7498 × 0.9744 × 0.9916 × 0.9520 × 0.8492 × 0.7364 = 0.4014. In contrast, a hypothetical child in the worst category of every variable would have a relative risk of 1.0418 × 1.3461 × 1 × 1 × 1 × 1 x 1 × 1 = 1.4024. Thus, a convenient measure of the combined effect of these eight variables—irrespective of the choice of reference categories—is 1.4024 / 0.4014 = 3.4938. That is, a hypothetical child in the *worst* category of all eight household-level variables would have a fitted value of $_1q_0$ that is 3.49 times greater than for a hypothetical child in the *optimal* category of the eight variables, holding everything else constant.

The five spatial variables appear in models 4 and 5. Their coefficients are similar in those two models but are closer to unity in model 5 than in model 4 because of the association with household-level variables. Rainfall is weakly significant for age 0 in the absence of household-level variables; urban residence is highly beneficial, as is higher density; infant mortality tends to increase with distance from the coast of Africa; and tree crops are the most advantageous type of farming system.

Density and distance to the coast are all interval-level variables. To give a better sense of their importance, we can calculate their effect on infant mortality at specific values. For example, the 10th percentile of the density measure is 16.7496 and the 90th percentile is 374.9888. When converted to (natural) logarithms, giving a much better fit, the 10th and 90th percentiles are 2.8184 and 5.9269, respectively. In model 5, the exponentiated coefficient for the log of density is 0.9716. Therefore, the relative risk for density is $0.9716^{2.8284} = 0.9220$ at the 10th percentile and $0.9716^{5.9269} = 0.8430$ at the 90th percentile. Both of these are the risk relative to ln(distance) = 0, i.e. distance = 1. As in the above scenario comparing the best and worst scenarios of household-level factors, it may be more useful to compare the two ends of the density distribution. Thus, 0.8430/0.9220 = 0.9143 is the risk of an infant death at the 90th percentile of the density distribution, relative to the risk at the 10th percentile. This is about a 9 percent reduction in the risk of an infant death. Given that these effects are over and above those of urban residence, which itself lowers the risk of death by 13 percent, we consider this effect to be substantial.

The 10th and 90th percentiles of distance to the coast are 12.52742 and 892.0779, respectively. Going through the same steps as above, the risk of an infant death is about 30 percent greater at the 90th percentile of the distance distribution than at the 10th percentile.

Now consider child mortality for age 1–4 as described in Table 7. There are many similarities to the results for age 0, but some differences as well. Most of the country effects become insignificant by model 5, with the notable exceptions of Côte d'Ivoire and Niger, two of the three countries that were significantly higher than Ghana in terms of infant deaths. The covariates introduced in models 2–5 have virtually no effect for

[6] Since children in this age group do not themselves use toilets, we interpret this variable as a proxy of the general hygiene and sanitation infrastructure of the household.

Côte d'Ivoire; indeed, there is even a slight increase in its coefficient as other variables are added. The third time period appears to have significantly higher mortality, but we must interpret this coefficient with care. The coefficient is affected by the timing of the specific surveys and the increased censoring of the most recent time period as well as by possibly genuine trends in mortality, perhaps due to HIV infection.

The multiple birth and birth order effects are smaller for age 1–4 than for age 0, although the birth order effect remains substantial and highly significant. Multiple births have a relative risk about 60 to 65 percent higher than singletons, even after the effects of low birth weight and competition for the mother's milk are largely past. Higher age of mother (at time of birth) continues to have a beneficial effect, reducing the risk by 20 to 25 percent. The child's gender has no effect on survival beyond infancy.

Source of water becomes more important as the child is weaned. Surface water is clearly inferior to piped water. The class "other water," however, is optimal, lowering the risk of death by 13 percent. Unfortunately, this classification (1.5 percent of the sample) was used primarily in the Benin and Niger surveys and no additional information was provided to aid interpretation (or to allow us to group these cases with other known water types). Anything other than a flush toilet increases the risk of death by about 32 to 45 percent (model 3). A finished floor is optimal and "other floor" (again, no interpretation or additional aggregation possible) is worst. Electricity is highly protective for age 1–4, although it was not for age 0; households with electricity have child mortality probabilities about 23 percent below households without it. Radio and refrigerator also have a protective effect.[7] Mother's education has an even more beneficial monotonic effect for age 1–4 than for age 0. The fitted probability of dying is about 36 to 42 percent less for women with some secondary education.

Source of water becomes more important as the child is weaned. Surface water is clearly inferior to piped water. The class "other water," however, is optimal, lowering the risk of death by 13 percent.

In the final model, a hypothetical child in the optimal combination of the eight household predictors would have a risk of 0.7708 × 1 × 0.9308 × 0.7738 × 0.9233 × 0.8985 × 0.7882 × 0.6446 = 0.2340, relative to a child in the reference combination. Another hypothetical child, in the worst combination, would have a risk of 1.0614 × 1.3721 × 1.6629 × 1 × 1 × 1 × 1 × 1 = 2.4218. The ratio of the highest risk combination to the lowest risk combination is 2.4218 / 0.2340 = 10.3494. That is, the fitted risk is more than ten times as great in the worst combination, compared with the best one. This is a much greater degree of variation than was found for age 0.

Of course, the household variables are to a large degree proxies for a whole package of characteristics representing standard of living, hygienic practices, and so on. Individual effects should not be taken completely at face value. For example, separate tabulations show that more education and having a refrigerator are highly correlated, and the parents in households with refrigerators tend to have even more than "some secondary" education. Being able to preserve food safely is undoubtedly important for child survival, but households with refrigerators usually have many additional advantages.

The spatial effects for mortality during age 1–4 are somewhat different than for age 0. Urban residence is still protective, but higher density is not. Distance from the coast is highly significant. The effects of climate and related agricultural production are more important determinants of death among children age 1–4 than among infants: Tree crops are the optimal farming system, with about 30 percent lower risk of death than the "other" category. There is a significant but nonlinear effect for rainfall. The coefficient for the quadratic term for rainfall is greater than one, which means effect of rainfall is curvilinear (concave).[8]

[7] The correlation between ownership of a television and electricity is 0.627—the highest of pairwise correlation among these four variables—suggests that it may be overspecified to include both in the model.

[8] The survival analysis revealed significant differentials by the malaria transmission index, which ranges from 0 to 38 in the study region. As noted above, the index is primarily a nation-

Children in the two shortest-season areas, that is, in the arid and semiarid range, had 15 percent and 12 percent higher risk of death than children in the optimal range.

Recall that because rainfall and growing season were so highly collinear we could not include both terms in the model. Instead, we ran the models replacing rainfall with growing season—a variable that might not pick up the effect of disease vectors, for example. The results (not shown in the tables) suggest that growing seasons under 120 days have significant negative effects on child mortality but not infant mortality. Children in the two shortest-season areas, that is in the arid and semiarid range, had 15 percent and 12 percent higher risk of death than children in the optimal range.[9] The risk of death was not higher for children in the wettest range (more than 10 growing months) although agricultural research indicates that growing seasons of this length are not optimal (FAO, 1978). When this variable is introduced in model 5 it also reduces the impact of the distance to the coast. While still weakly significant, coastal zones in this region are wetter than interior areas, and this is accounted for more directly with the growing season data. Nevertheless, residual effects associated with coastal proximity remain.

4.4 Overall Effects and Interpretations

The Pseudo R^2 in our tables is calculated in the standard way as $R^2 = 1 - (LLm / LL0)$, where LLm is the log of the likelihood function for the specific model and LL0 is the log of the likelihood function for the null model, which has no covariates and is restricted to exactly the same cases that appear in the specific model. It can be interpreted as the proportion of the total deviance that is explained by the covariates in the model.

The overall effect of the household-level variables can be measured by the increase in Pseudo R^2 when model 3 is compared with model 2, or when model 5 is compared with model 4. Similarly, the overall effect of the spatial variables is shown by the increase when model 4 is compared with model 2, or when model 5 is compared with model 3. We will not list these differences numerically, but it can easily be seen that the overall effect of both sets of variables is generally small for both age 1–4 and age 0. As a set, the spatial variables appear most important when they are added to model 2 for age 1–4; the Pseudo R^2 for model 4 is increased by $0.04055 - 0.02972 = 0.0108$, about 1 percent of the total deviance. The overall effect of the household variables is greater in every such comparison, which is consistent with the discussion of the levels and significance of coefficients.

The urban-rural distinction, as noted before, has been included as a spatial characteristic for conceptual reasons but is actually available in the DHS surveys and would often be grouped with what we have called household characteristics. Much of the importance of the spatial variables can be attributed to this inclusion. Further, model 5 does not account for distance to nearest populated settlement or interaction terms between urban residence and density; for example, to consider the possibility that urban proximity is not a uniform effect (e.g., interurban high-density residence may increase the risk of death). Some of these possibilities were entertained separately and are shown in Table 8. When the interaction of density and urban residence is considered, urban residence loses its significance and urban density lowers infant mortality (but not child mortality). The further an infant lives from a city of moderate size, the greater the risk

al-level composite (Kiszewski et al., forthcoming), thus it was removed from the multivariate model. Nevertheless, had it been included in model 1 (not shown), mortality would be shown to raise the risk of an infant death (1.005) and child death (1.007), respectively. These effects were not sustained, as additional variables are entered, and produced some confounding effects at the country level, indicating that more evaluation of the variable or its specification is needed and that the coefficients should be interpreted with caution.

[9] The optimal cutoff of 70 days for arid was not possible given the original classification of the data, so the data were classified as 0–90 for arid and 91–120 for semiarid. The reference category was a growing season of 120–240 days.

Table 8
The relative risk of dying at $_1q_0$ and $_4q_1$: variations on urban-type variables in model 5

Variable	A	B	C	D	E
			$_1q_0$		
Urban	0.857**	-	0.818	-	0.875**
Density (ln)	-	0.964*	0.966+	-	0.977
Density (ln) × urban	-	-	1.015	-	-
Distance to populated place	-	-	-	1.001*	1.000
Density (ln) × distance	-	-	-	-	-
			$_4q_1$		
Urban	0.886**	-	1.085	-	0.876**
Density (ln)	-	1.004	1.027	-	1.002
Density (ln) × urban	-	-	0.953	-	-
Distance to populated place	-	-	-	1.000	1.000
Density (ln) × distance	-	-	-	-	-

Note: All other variables in model 5 are controlled for here, but in none of these models is model 5 exactly replicated.
+ $p < 0.10$ ** $p < 0.01$
* $p < 0.05$ *** $p < 0.001$

The further an infant lives from a city of moderate size, the greater the risk of death; however, this effect is not observed for children age 1–4.

of death; however, this effect is not observed for children age 1–4. The effect is eliminated if it is entered along with the dichotomous urban-rural variable and population density. While this approach is far from satisfying in terms of explaining the continuum of urban-rural phenomena, alternatives are not intuitive. Not shown is the substitution of the urban-rural and density variables in model 5 with a series of rank-order variables of urban-density and rural-density classes. The risk of infant death is greatest in the sparsest and most dense rural areas and higher in all rural areas than in urban areas. In urban areas, the risk of infant death is lowest in the densest areas. The effects on children are not noteworthy. While this part of Africa is not known for having high density urban areas, it is somewhat surprising to find that density differences have no effect on children's mortality in urban and rural areas.

The coastal effect is another one of the robust spatial variables predicting infant and child deaths. The effect was also found to be important in economic development (Sachs et al., 2001) because, it is argued, coastal zones tend to be advantaged in their ability to transport goods, services, and ideas. The coastal countries in this study tend to have a higher gross domestic product (GDP) per capita—regardless of whether GDP is measured by purchasing power parity (PPP) or otherwise—than the landlocked countries, Burkina Faso, Mali, and Niger (see Table 9). Because country is also controlled for here[10] and because the distance to coast measure is continuous, the impact of coastal proximity may be interpreted as an inter- and intranational access measure, above and beyond country-level economic development. That is to say, interior dwellers in coastal countries are at greater risk of death than their coastal counterparts. No subnational level income or GDP measures are available, but it may be that in coastal countries the coastal zone is disproportionately well off.

4.5 Lessons from Extreme Cases

As a way of highlighting the extremes in the probabilistic distribution of death, we identified from model 5 above the 1000 cases with the highest and lowest probability of death for infants and children. Table 10 shows the percentage of cases by selected variables. We found extremely high Pearson's chi-squared values for each of these cross-tabulations (all with $p < 0.000$) indicating that these extreme cases differ significantly from each other and from the remaining cases in the dataset.

Table 9
National-level indicators for the study region

	Population indicators				Poverty indicators					DHS health indicators				
					GDP per capita[3]									
	Population, 2003 estimate (in millions)[1]	Population growth rate, 2003 estimate[2]	Life expectancy, 2003 estimate[2]	Fertility, 2003 estimate[2]	2001 estimate	2001 est. expressed as PPP	GDP growth (Average Annual Growth 1990-2000)[4]	Poverty (percentage of population below poverty line, 2001)[2]	Poverty (Gini index: Distribution of family income [year])[2]	HIV prevalence, adults 15-49 years (end of 2001)[5]	Malnutrition (weight-for-height < –3 SD)[6]	Vaccination coverage[6, 7]	Use of health services (assisted delivery[6, 8])	Year of DHS survey
Benin	7.0	2.95	51.0	6.30	368	980	4.7	37	u	3.6	1.7	59.0	72.9	2001
Burkina Faso	13.2	2.60	44.5	6.34	215	1,120	4.9	45	48.2 (1994)	6.5	4.3	29.3	31.0	1998–99
Cameroon	15.7	2.02	48.0	4.63	559	1,680	1.7	48	47.7 (1996)	11.8	0.8	29.4	58.3	1998
Côte d'Ivoire	17.0	2.15	42.7	5.51	634	1,490	3.5	37 (1995)	36.7 (1995)	9.7	1.0	50.7	47.1	1998–99
Ghana	20.5	1.45	56.5	3.32	269	2,250	4.3	31 (1994)	40.7 (1999)	3.0	1.7	62.0	44.3	1998
Guinea	9.0	2.37	49.5	5.90	394	1,960	4.3	40 (1994)	40.3 (1994)	u	2.9	32.2	34.8	1999
Mali	11.6	2.82	45.4	6.66	239	810	3.8	64	50.5 (1994)	1.7	1.9	28.7	39.0	2001
Niger	11.0	2.71	42.2	6.91	175	890	2.4	63 (1993)	50.5 (1995)	u	3.7	18.4	17.6	1998
Senegal	10.6	2.56	56.4	4.93	476	1,500	3.6	54	41.3 (1995)	0.5	u	u	46.6	1997
Togo	5.4	2.37	53.4	4.97	270	1,650	2.3	32 (1989)	u	6.0	2.1	30.8	50.5	1998

[1] IPC
[2] CIA Factbook, available at http://www.odci.gov/cia/publications/factbook/geos/bn.html
[3] UN Human Development Report Office, available at http://www.undp.org/hdr2003/indicator/indic_111_1_1.html
[4] 2002 World Development Indicators, available at http://www.worldbank.org/data/wdi2002/pdfs/table%204-1.pdf
[5] UNAIDS/WHO
[6] ORC Macro, MEASURE *DHS*+ STATcompiler
[7] Children who are fully vaccinated are those who have received BCG, measles, and three doses of DPT and polio (excluding polio 0).
[8] Doctor or trained midwife/health professional
PPP = Purchasing power parity
u = Unknown (not available)

Infants with a high predicted probability of dying were disproportionately located in Mali and Niger (77 percent of high-risk cases, compared with 33.7 percent of the full sample). All were multiple births, nearly 80 percent were third or higher birth order, and 70 percent were males. Almost all (93.8 percent) were born to mothers with no education. Few infants (under 3 percent) were born into households with amenities such as electricity or television. Factors related to differences in environment and proximity to urban or coastal areas are substantially different in infants with a high probability of dying. In particular, these infants tend to live in dry zones far from both coastal and urban areas. Nearly half were born into areas with low annual rainfall, virtually all were born into areas more than 200 kilometers from a coastline (99.4 percent), and two-thirds were born more than 50 kilometers from a populated place. Moreover, nearly 80 percent of these infants live in sparsely populated areas (fewer than 50 persons per square kilometer). Of all infants with a high predicted probability of death, 44.8 percent in fact died.

Infants with a high probability of survival, conversely, were not as geographically concentrated, with no country having a disproportionate number of these cases. Moreover, these cases are more dispersed among clusters; whereas five clusters in Mali and two clusters in Niger had eight or more infants with high probability of dying, no cluster had more than five infants with high probability of survival. Very few cases with high probability of survival were multiple births, and only 41.1 percent were male. Infants with a higher probability of survival tend to be more coastal and more urban than the full sample; however, these differences are not as marked. Frequencies of cases with household amenities are also slightly higher than the full sample. None of these 1000 cases died.

Infants with a high predicted probability of dying were disproportionately located in Mali and Niger. All were multiple births.

Like infants, children 1–4 who are at high risk of death are geographically concentrated, with 78.1 percent of all cases found in Niger, and none found in Benin, Ghana, Senegal, and Togo. Moreover, these cases are concentrated in clusters, with 10 percent of all cases found in seven clusters in Niger. Environmental and spatial factors also appear to affect infants and children in the same way. Nearly all live more than 200 kilometers from a coast, and virtually none (0.2 percent) live in areas of moderate or higher population densities (150 or more persons per square kilometer). (However, chances of survival are not similarly linear: for children 1–4, survival is more likely at either the highest or lowest densities.) Low maternal education is also prominent among children with high predicted probability of death, with nearly all (97 percent) of these children being born to mothers with no education and none being born to mothers with secondary or higher education. Differences in household amenities are especially pronounced, with virtually none of these children living in households with a television, refrigerator, or electricity. Of these 1000 children, 39.1 percent died.

An examination of these extreme cases serves primarily to confirm the findings of the more thorough analysis above. Differences in proximate determinants such as sex and birth order are more pronounced in the infant analysis, while differences in household characteristics are more pronounced in the child analysis. Differences in maternal education are prominent throughout. The spatial characteristics of these extreme cases, however, provide some unique insights into the role of spatial factors in infant and child mortality. While cases with extremely high probability of dying are contained in a relatively compact area of southern Niger and Mali, cases with low probability of dying are less concentrated and more coastal. Cases (both infants and children) with particularly high probability of dying are notably not urban.

[10] When country is omitted from model 5 (not shown) the effects of coastal proximity are raised to risk ratios of 1.0004 and 1.0006 for infants and children, respectively.

Table 10
Profiles of extreme cases

The 1,000 cases with the lowest and highest predicted probability of dying

	Infants ($_1q_0$)		Children ($_4q_1$)		
Variable	Most likely to survive	Most likely to die	Most likely to survive	Most likely to die	Total
Died	0.0	44.8	0.0	39.1	8.4
Country					
Burkina Faso	6.8	4.6	9.8	3.9	9.6
Benin	8.4	3.3	8.3	0.0	8.5
Côte d'Ivoire	2.4	3.0	4.1	0.3	3.2
Cameroon	8.4	0.5	7.0	0.6	6.4
Ghana	7.4	0.0	8.7	0.0	5.4
Guinea	10.0	10.0	10.0	1.3	9.8
Mali	17.6	47.5	20.8	15.8	21.2
Niger	8.9	29.5	2.8	78.1	12.5
Senegal	17.8	0.8	15.9	0.0	11.9
Togo	12.3	0.8	12.6	0.0	11.5
Birth cohort					
1987–1991	0.0	25.5	0.0	14.5	24.9
1992–1996	2.7	49.4	16.2	18.1	50.8
1997–2001	97.3	25.1	83.8	67.4	24.3
Proximate determinants					
Male child	41.1	70.8	48.5	48.3	50.6
Multiple birth	0.8	100.0	1.7	50.1	3.7
Birth order					
First birth	13.7	9.3	19.5	20.3	18.5
Second birth	17.4	10.9	17.6	13.6	16.3
Third or higher birth	68.9	79.8	62.9	66.1	65.3
Mother's age at birth					
Younger than 20	10.0	22.2	13.3	36.3	17.9
20–24	69.9	58.3	70.4	53.4	67.9
35 or older	20.1	19.5	16.3	10.3	14.3
Household water source					
Piped	35.0	10.6	33.8	5.2	27.2
Well	50.5	76.3	50.6	85.3	56.7
Surface	12.0	12.2	13.8	9.1	14.2
Other	2.5	0.9	1.8	0.4	1.9
Household assets					
Electricity	25.2	2.5	24.6	0.1	16.3
Radio	66.0	52.1	66.7	30.1	61.7
Television	20.0	2.8	21.6	0.2	13.2
Refrigerator	11.3	0.5	10.9	0.0	5.9
Mother's education					
No education	66.1	93.8	66.3	97.0	76.5
Some primary	12.3	5.6	13.7	2.4	11.3
Completed primary	5.5	0.4	5.1	0.6	4.2
Some secondary or higher	16.1	0.2	14.9	0.0	8.0
Rainfall (quintiles)					
First quintile (dry)	27.2	49.8	19.8	88.6	27.4
Second quintile	19.1	19.8	24.0	8.6	22.8
Third quintile	22.8	16.6	26.9	1.2	25.3
Fourth quinitile	14.3	6.4	13.5	0.5	10.8
Fifth quintile (wet)	16.6	7.4	15.8	1.1	13.8
Urban	32.2	6.3	33.0	4.1	25.1
Population density (ln) (quintiles)					
First quintile (dispersed)	19.7	41.7	21.2	32.4	23.8
Second quintile	18.2	36.4	22.3	27.4	23.2
Third quintile	15.6	13.5	15.7	27.6	17.3
Fourth quinitile	19.5	5.9	15.4	12.4	17.2
Fifth quintile (concentrated)	27.0	2.5	25.4	0.2	18.5
Distance to coast (quintiles)					
First quintile (coastal)	27.0	1.6	25.3	0.3	18.8
Second quintile	18.6	2.7	15.9	0.5	14.8
Third quintile	16.8	13.6	20.4	1.0	18.4
Fourth quinitile	20.3	30.3	26.0	14.3	25.1
Fifth quintile (inland)	17.3	51.8	12.4	83.9	23.0
Farming system					
Tree crop	12.7	0.6	11.4	0.0	8.2
Coastal fishing	10.6	1.1	11.5	0.3	9.2

Note: Chi-squared values for each of these cross-tabulations are highly significant ($p < 0.000$).

5 Conclusions

Although country-level mortality ranges from 61 to 135 and from 52 to 130 per 1000 live births for infants and children respectively (in the 10-year period surveyed in each country), the impacts of most country-level differences become insignificant when household and spatial characteristics are included. The notable exceptions are Côte d'Ivoire and Niger, for which the risk of infant and child deaths remains significantly higher, as well as Mali for infant deaths. Spatial factors (e.g., proximity to urban areas, population density, farming systems), which appear to have an overall modest effect on both infant and child mortality, especially when the usual demographic and household characteristics are included, explain away a good deal of the country-specific variation in mortality. The improved definition of spatial factors may alert policymakers to address geographic parameters such as providing services to areas further from the coast.

The effects of proximate determinants (e.g., multiple births, birth order, sex, and mother's age) are consistent with findings often cited in the literature, specifically that they act more strongly on the risk of infant death than on the risk of child death. Similarly, household-level variables, including maternal education and housing quality, play an important role in the determination of both infant and child survival. Our estimates suggest that infants and children living in the most disadvantaged conditions (e.g., those whose mothers have no schooling, whose households do not have flush toilets or electricity, and who obtain drinking water from surface sources) are at risk of death 3.5 and 10.4 times that of infants and children, respectively, living in the most optimal conditions (e.g., those with mothers having at least some secondary schooling, piped water for drinking, and flush toilets), holding all else constant. The spatial variables are associated with household characteristics and may have an indirect effect mediated through these characteristics. A meaningful future analysis would be to explore the degree to which household characteristics are themselves determined by the physical environment. In the meantime, results from the present analysis suggest that policy efforts to reduce infant and child mortality should incorporate programs to increase mothers' education and improve household sanitation.

The impacts of most country-level differences become insignificant when household and spatial characteristics are included.

Population density and distance from the coast are significant determinants of infant and child mortality even when urban residence, which lowers the risk of death by 12–13 percent, is considered. The risk of infant death is about 30 percent greater at the 90th percentile of the coastal distance distribution than at the 10th percentile, and the relative risk is somewhat greater for children far from the coast. Tree crops were found to be the optimal farming system, with about 30 percent lower risk of child death than other systems; the impact was smaller—only 20 percent lower risk—on infants, again suggesting the greater importance of environmental factors for children than for infants, who are protected from some environmental effects, for example through breastfeeding. A direct and consistent effect of rainfall was not found, but it was determined that children living in areas with the shortest growing seasons, classified as arid and semiarid, had 15 percent and 12 percent higher risk of death, respectively, than children in the optimal range. No effect was found on infants.

Finally, a set of extreme cases—the 1000 cases with highest and lowest probabilities of death—were examined to flesh out some of the complexities in the full model, especially with regard to spatial characteristics. Cases with extremely high probability of dying were found to be contained in a relatively compact area of southern Niger and Mali, whereas cases with low probability of dying were less concentrated and more likely to be coastal. Notably, infants and children with particularly high probabilities of death were also found to be located in sparsely populated and nonurban areas. Survival chances for children, however, were greatest in the most dense and least dense areas, which suggests that the influence of population density on child death and survival is complex and bears additional consideration.

Population density and distance from the coast are significant determinants of infant and child mortality even when urban residence, which lowers the risk of death by 12–13 percent, is considered.

Suggestions for Further Research

In addition to determining how spatial variables impact household characteristics, as mentioned above, future studies should strive to optimize spatial information. For example, urban and climate variables can be classified into subcomponents that provide more insight. A dichotomous urban-rural variable is complemented by specific information on population density and distance to urban center; climate variables reflecting rainfall, growing seasons, and farming systems are important in the assessment of the effects of disease transmission and food production on mortality. But how these variables should be optimized—that is, combined, interacted, and transformed—depends on theoretical and statistical concerns beyond those that were considered here.

There are several reasons to both widen and narrow the geographic scope of this study.[11] Including countries with a wider range of physiographic features would facilitate comparisons that we could not undertake here. For example, elevation is thought to be an important component of malaria transmission in Africa, although in this study region there would have been little variability to evaluate. Including one additional country, Nigeria, would also provide a more complete regional picture. With more than 100 million inhabitants, Nigeria's population is comparable in size to that of the entire region studied, and it is nearly surrounded by these countries. Results for the 2003 Nigeria DHS survey were unavailable at the time of this study, but any future studies should include it.

The geographical scope may be narrowed to more precisely detect variations within a country. Because there are large differences (e.g., in population density or rainfall) over a large area, some of the effects within a country may be overpowered by the intercountry emphasis here. Comparative country-level studies would facilitate a more systematic assessment of hypothesized interactions between spatial and household characteristics as well as among spatial characteristics. Furthermore, questionnaire design for a single country may include country-specific covariates that are not necessarily comparable across multiple countries. Therefore, important covariates, such as partner's occupation, that could not be used in the present study could be included in single country studies.

Future survey implementation may also benefit by incorporating additional geographic concerns in the sampling frame. Currently, surveys are representative within political regions, but not other geographic regions.[12] To the extent that particular geographic parameters are believed to be important, oversampling in some places, such as hyperarid zones, could be a valuable undertaking. At the least, future studies incorpo-

[11] It is important to recognize and attempt to reconcile differences in variable coding so as to lose as few covariates as possible. While in-country survey implementation teams may have an interest in making their surveys as country specific as possible, it may be possible post hoc to determine complementarities across surveys for the purpose of relative ranking (e.g., best to worst condition).

[12] Some older DHS surveys such as Burundi 1987 and Côte d'Ivoire 1994 used environmental regions aggregated from small administrative units.

rating distinct geographic zones should take care to ensure that the sample sizes in the various classes of zones are sufficient to generate robust results.

Future work should attempt a more systematic examination of the spatial patterns of mortality and its determinants. This analysis has confirmed spatial and nonspatial risk factors, but it came short of examining cluster-level or fine-scale spatial patterns. This was not attempted in the current analysis because of concerns over using the cluster as a unit of analysis. Additional statistical work and consideration of sampling issues would clarify the feasibility of this approach. Spatial statistical programs are becoming increasingly sophisticated, perhaps accelerating this line of inquiry.

Another matter of spatial concern is access to resources. To assess a country's quality and coverage of health care services, DHS has begun collecting data, including geographic location, on health facilities in its service provision assessment (SPA) surveys. The SPA includes a nationally representative sample of health facilities, including national- and provincial-level hospitals, health centers, and dispensaries managed by the government or by NGOs. (Although the samples have until now excluded privately run, for-profit pharmacies and clinics, it has been proposed to include them in future SPAs.) Information is collected from service providers and clients at these facilities on facility infrastructure, specific child health, family planning and maternal health services, and services for sexually transmitted diseases and HIV/AIDS. This is a potentially rich source of data on provision of national health care that could be linked to household survey data where both types of data are georeferenced. In some countries, health facilities data may be available from other sources. Senegal, for example, through its Averted Maternal Deaths and Disability (AMDD) program (Moreira et al., n.d.) collects spatial information on all hospitals (government and private) and their service provision levels; and some countries have georeferenced some or all of their health facilities (e.g., Namibia, Malawi), which may provide indicators of subnational levels of service availability. Using such data in connection with DHS surveys might be a valuable exercise.[13]

Similarly, other physical datasets may be of interest for future work in addition to some of those mentioned above (e.g., land use or land cover). The Total Ozone Mapping Spectrometer aerosol index, which measures dust, could provide other correlates of mortality, especially in arid regions where dust is especially problematic. Model data sets for some disease vectors other than malaria may also be considered.

Finally, the geographic variables in this study were static, relating to one time period, while the surveys themselves were carried out over a five-year period and provide data about births from a fifteen-year period (1987–2001). Data such as rainfall and the Normalized Difference Vegetation Index (NDVI) are reported on a monthly basis, so that a location- and period-specific average rainfall could be calculated for each child in the survey. Or, more modestly, country-specific rainfall datasets could be calculated averaging over the ten-year period preceding each survey. Other data such as population density could be projected backwards in time to account for national decadal growth rates of 25–40 percent in the region in the 1990s to generate a more accurate estimate of population density at the time of the survey. Using time-varying data would require slightly different multivariate models than the ones pursued here, but it is a direction worth pursuing.

[13] In connection with use of health service data, it would be informative to know more about the nutritional status and health care seeking behavior of children who died. However, the logistics of collecting this data and the potentially poor quality of the retrospective information may outweigh the benefits.

References

Adams, A. 1994. Seasonal variations in nutritional risk among children in central Mali. *Ecology of Food and Nutrition* 33: 93–106.

Adetunji, J. 2000. Trends in under-5 mortality rates and the HIV/AIDS epidemic. *Bulletin of the World Health Organization* 78(10): 1200–1206.

Agha, S. 2000. The determinants of infant mortality in Pakistan. *Social Science and Medicine* 51(2): 199–208.

Ahonsi, B. A. 1995. Age variations in the proximate determinants of child mortality in south-west Nigeria. *Journal of Biosocial Science* 27(1): 19–30.

Armah, G. E., J. A. Mingle, A. K. Dodoo, A. Anyanful, R. Antwi, J. Commey, and F. K. Nkrumah. 1994. Seasonality of rotavirus infection in Ghana. *Annals of Tropical Paediatrics* 14(3): 223–229.

Balk, D., F. Pozzi, G. Yetman, A. Nelson, and U. Deichmann. 2003. The spatial distribution of global population in urban and rural areas: African examples. Paper presented at the Annual Meeting of the Association of American Geographers, New Orleans, Louisiana, March 2003.

Basu, A. M. 1997. Underinvestment in children: A reorganization of the evidence on the determinants of child mortality. In *The continuing demographic transition*, edited by G. W. Jones et al. Oxford: Oxford University Press, 1997. 307–31.

Boerma, J. T. 1996. *Child survival in developing countries: Can demographic and health surveys help to understand the determinants?* Amsterdam, Netherlands: Royal Tropical Institute.

Boerma, J. T. and G. T. Bicego. 1992. Preceding birth intervals and child survival: Searching for pathways of influence. *Studies in Family Planning* 23(4): 243–56.

Boerma, J. T., A. J. Nunn, and J. A. G. Whitworth. 1998. Mortality impact of the AIDS epidemic: evidence from community studies in less developed countries. *AIDS* 12 (Suppl. 1): S3–S14.

Boserup, E. 1965. *The conditions of agricultural growth: The economics of agrarian change under population pressure.* Chicago: Aldine Press.

Center for International Earth Science Information Network (CIESIN), Columbia University. 2003. *Gridded population of the world, version 3 (beta).* Available at <http://beta.sedac.ciesin.columbia.edu/gpw/>.

Climate Research Unit (CRU), University of East Anglia, n.d. 0.5 degree 1961–1990 mean monthly climatology, downloaded August 2003 from <http://iri.ldeo.columbia.edu/> (also available from the Intergovernmental Panel on Climate Change (IPCC) at <http://ipcc-ddc.cru.uea.ac.uk>).

Cramer, J. C. 1987. Social factors and infant mortality: Identifying high-risk groups and proximate causes. *Demography* 24(3): 299–322.

Curtis, S. L. 1995. *Assessment of the quality of data used for direct estimation of infant and child mortality in DHS-II surveys.* Occasional Papers No. 3. Calverton, Maryland: Macro International Inc.

Curtis, S. L., and M. Hossein. 1998. *The effect of aridity zone on child nutritional status.* West Africa Spatial Analysis Prototype Exploratory Analysis. Calverton, Maryland: Macro International Inc.

Defo, B. K. 1994. Determinants of infant and early childhood mortality in Cameroon: The role of socioeconomic factors, housing characteristics, and immunization status. *Social Biology* 41(3–4): 181–211.

Deichmann, U. and L. Eklundh. 1991. *Global digital data sets for land degradation studies: A GIS approach.* GRID Case Study Series No. 4. Nairobi: United Nations Environment Programme (UNEP) and Global Resource Information Database (GRID).

Diamond, I. 2000. Child mortality—the challenge now. *Bulletin of the World Health Organization* 78(10): 1174.

Dixon, J. and A. Gulliver with D. Gibbon. 2001. *Farming systems and poverty: Improving farmers' livelihoods in a changing world*. United Nations Food and Agriculture Organization (FAO) and World Bank, Rome and Washington D. C. Available at <http://www.fao.org/farmingsystems/>.

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

Elvidge, C., M. Imhoff, K. Baugh, V. Hobson, I. Nelson, J. Safron, et al. 2001. Nighttime lights of the world: 1994–95. *ISPRS Journal of Photogrammetry and Remote Sensing* 56: 81–99.

Entwisle, B., R. R. Rindfuss, S. J. Walsh, T. P. Evans, and S. R. Curran. 1997. Geographic information systems, spatial network analysis, and contraceptive choice. *Demography* 34(2): 171–187.

Esrey, S. A. and J. P. Habicht. 1986. Epidemiological evidence for health benefits from improved water and sanitation in developed countries. *Epidemiology Review* 8: 117–128.

Fabrega, H. 1972. Medical anthropology. In *Biennial review of anthropology*, ed. B. Siegel. Stanford: Stanford University Press. 167–22.

Findley, S., D. Balk, M. Barlow, and N. Sogoba. 2002. Putting climate in the service of public health in Mali. Paper presented at the Annual Meeting of the Population Association of America, Atlanta, Georgia, April 2002.

Fischer, G., et al. 2000. *Global agro-ecological zones*. United Nations Food and Agriculture Organization (FAO) and International Institute for Applied Systems Analysis (IIASA). Available at <http://www.iiasa.ac.at/Research/LUC/GAEZ/index.htm>.

Food and Agriculture Organization of the United Nations (FAO). 1978. *Report on the agro-ecological zones project.* Volume 1. *Methodology and results for Africa.* World Soil Resources Report 48. Rome: FAO.

Forste, R. 1994. The effects of breastfeeding and birth spacing on infant and child mortality in Bolivia. *Population Studies* 48: 497–511.

Fujii, T. 2002. *Estimation of poverty rates at commune-level in Cambodia: Using the small-area estimation technique to obtain reliable estimates.* Phnom Penh: Ministry of Planning and United Nations World Food Programme.

Fujii, T., S. Alayon, and L. Montana. 2002. *Micro-level estimation of the prevalence of stunting and underweight among children in Cambodia.* Phnom Penh: Ministry of Planning and United Nations World Food Programme.

Gaminiratme, K. H. 1991. Socioeconomic and behavioral determinants of diarrhoeal morbidity among children in Sri Lanka. In *DHS World Conference proceedings,* Vol. 1, 757-784. Calverton, Maryland: Macro International Inc.

Ghana Statistical Service (GSS) and Macro International Inc. (MI). 1999. *Ghana Demographic and Health Survey 1998*. Calverton, Maryland: GSS and MI.

Gupta, H. S. and A. Baghel. 1999. Infant mortality in the Indian slums: Case studies of Calcutta metropolis and Raipur city. *International Journal of Population Geography* 5(5): 353–66.

Hobcraft, J. N, J. W. McDonald, and S. O. Rutstein. 1985. Demographic determinants of infant and early child mortality: A comparative analysis. *Population Studies* 39(3): 363–85.

Hussain, A., S. M. Keramat Ali, and G. Kvåle. 1999. Determinants of mortality among children in the urban slums of Dhaka city, Bangladesh. *Tropical Medicine and International Health* 11: 758–764.

Kiszewski, A. et al. Forthcoming. A global index representing the stability of malaria transmission. *American Journal of Tropical Medicine and Hygiene*.

Korenrump, E. L., F. Arnold, B. G. Williams, B. L. Nahlen and R. W. Snow. Forthcoming. *Monitoring trends in under-five mortality through national birth history surveys.*

Lesthaeghe, R. J., ed. 1989. Reproduction and social organization in sub-Saharan Africa. *Studies in Demography* No. 4. Berkeley, California: University of California Press.

Macro International Inc. 1996. *Sampling manual*. DHS-III Basic Documentation Number 6. Calverton, Maryland: Macro International Inc.

Malthus, T. R. 1798. *An essay on the principle of population, as it affects the future improvement of society, with some remarks on the speculations of Mr. Godwin, M. Condorset, and other writers.*

Manda, S. O. M. 1999. Birth intervals, breastfeeding and determinants of childhood mortality in Malawi. *Social Science and Medicine* 48(3): 301–12.

Majumder, A. K., M. May, and P. D. Pant. 1997. Infant and child mortality determinants in Bangladesh: Are they changing? *Journal of Biosocial Science* 29(4): 385–99.

McCullagh, P. and J.A. Nelder. 1989. *Generalized linear models.* Second edition. London: Chapman and Hall.

McMurray, C. 1997. Measuring excess risk of child mortality: An exploration of DHS I for Burundi, Uganda and Zimbabwe. *Journal of Biosocial Science* 29(1): 73–91.

Menendex, C., A. F. Fleming, and P. L. Alonso. 2000. Malaria-related anaemia. *Parasitol Today* 16(11): 469–476.

Merrick, T. W. 1985. The effect of piped water on early childhood mortality in urban Brazil, 1970-1976. *Demography* 22: 1–23.

Moreira, I., E. H. Ousseeynou, F. Sarr, S. D. Djiba, N. M. N. Gako, N. Fall, et al. *Evaluation de la disponibilité, de l'utilisation et de la qualité des soins obstétricaux d'urgence au Sénégal*. Ministère de la Santé, Républic du Sénégal and UNFPA, and Columbia University, Rapport Final (Draft). Forthcoming.

Mosley, H. W. and L. C. Chen. 1984. An analytical framework for the study of child survival in developing countries. In *Child survival: Strategies for research,* ed. W. G. Mosley and L.C. Chen. New York: Population Council. 25–44.

Mturi, A. J. and S. L. Curtis. 1995. The determinants of infant and child mortality in Tanzania. *Health Policy and Planning* 10: 384–94.

Muhuri, P. K. 1996. Estimating seasonality effects on child mortality in Matlab, Bangladesh. *Demography* 33(1): 98–110.

Namboodiri, K. and C. M. Suchindran. 1987. *Life table techniques and their applications*. Orlando, Florida: Academic Press.

National Research Council. 2001. *Under the weather: Climate, ecosystems, and infectious disease*. Washington, D. C.: National Academy Press.

New, M., M. Hulme, and P. Jones. 1999. Representing twentieth-century space-time climate variability. Part I: Development of a 1961–90 mean monthly terrestrial climatology. *Journal of Climate* 12: 829–856.

de Onis, M. 2000. Measuring nutritional status in relation to mortality. *Bulletin of the World Health Organization* 78(10): 1271–1274.

Patz, J., D. Engelberg, and J. Last. 2000. The effects of changing weather on public health. *Annual Review of Public Health* 21(1): 271–307.

Pison, G. 1989. Les jumeaux: fréquence, statut social et mortalité. In *Mortalité et société en Afrique au sud du Sahara*, ed. G. Pison, E. van de Walle, and M. Sala-Diakanda. Coll. Travaux et Documents, Cahier No. 124. Institut National d'Études Demographiques (INED) and Presses Universitaires de France (PUF): Paris, France.

Pitt, M. M. and W. Sigle. 1997. Seasonality, weather shocks and the timing of births and child mortality in Senegal. Paper presented at the 10th Anniversary Conference of the Centre for the Study of African Economies, Oxford University, April 1997.

Rain, D. U.S. Census Bureau. Personal communication.

Rao, S. R., A. Pandey, and K. I. Shajy. 1997. Child mortality in Goa: A cross-sectional analysis. *Social Biology* 44(1–2): 101–10.

Rice, A. L., L. Sacco, A. Hyder, and R. Black. 2000. Malnutrition as an underlying cause of childhood death associated with infectious diseases in developing countries. *Bulletin of the World Health Organization* 78(10): 1207–1221.

Ronsmans, C. 1995. Patterns of clustering of child mortality in a rural area of Senegal. *Population Studies* 49(3): 443–61.

Root, G. 1997. Population density and spatial differentials in child mortality in Zimbabwe. *Social Science and Medicine* 44(3): 413–421.

Root, G. 1999. Disease environments and subnational patterns of under-five mortality in sub-Saharan Africa. *International Journal of Population Geography* 5: 117–132.

Rutstein, S. 2000. Factors associated with trends in infant and child mortality in developing countries during the 1990s. *Bulletin of the World Health Organization* 78(10): 1256–1270.

Rutstein, S. O. 1984. *Infant and child mortality: Levels, trends, and demographic differentials.* WFS Comparative Studies No. 43. Voorburg, Netherlands: International Statistical Institute.

Sachs, J. D., A. D. Mellinger, and J. L. Gallup. 2001. The geography of poverty and wealth. *Scientific American* 284(3): 70–75.

Saha, T. 1998. *The effect of settlement zones and economic diversity on reproductive behavior in West Africa*. West Africa Spatial Analysis Prototype Exploratory Analysis. Calverton, Maryland: Macro International Inc.

Sauerborn, R. and A. Adams. 1996. Household coping with the economic costs of illness. *Social Science and Medicine* 43(3): 291–301.

Schultz, T. P. 1993. Mortality decline in the low-income world: Causes and consequences. *American Economics Review* 83(2): 337–42.

Shakhatreh, F. M. N., A. A. Abbas, and A. A. Issa. 1996. Determinants of infant mortality and the use of maternity services in Jordan. *Dirasat, Medical and Biological Sciences* 23(2): 59–69.

Slutsker, L, T. E. Taylor, J. J. Wirima, and R. W. Stiketii. 1994. In-hospital morbidity and mortality due to malaria-associated severe anaemia in two areas of Malawi with different patterns of malaria infection. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 88(5): 548–51.

Sullivan, J. M., S. O. Rutstein, and G. T. Bicego. 1994. *Infant and child mortality.* DHS Comparative Studies No. 15. Calverton, Maryland: Macro International Inc.

Trussell, J. and C. Hammerslough. 1983. A hazards-model analysis of the covariates of infant and child mortality in Sri Lanka. *Demography* 20: 1–26.

UNAIDS. 2002. *Report on the global HIV/AIDS epidemic 2002*. Geneva: UNAIDS.

UNICEF. 2000. *UNICEF end decade databases, 2000.* Available at <http://www.childinfo.org/cmr2.html>.

United Nations Development Program (UNDP). 2001. *The Gambia human development report 2000: Promoting good governance for human development and poverty eradica-*

tion. Banjul: UNDP.

United Nations Environment Programme (UNEP). 1997. *World atlas of desertification.* Second edition. London: UNEP.

van de Walle, E. and F. van de Walle. 1991. Breastfeeding and popular aetiology in the Sahel. *Health Transition Review* 1(1): 69–81.

Watson, R. T., M. C. Zinyowera, and R. H. Moss, eds. 1997. *The regional impacts of climate change: An assessment of vulnerability.* A special report of the Intergovernmental Panel on Climate Change, Working Group II. Cambridge: Cambridge UP.

Whitworth, A. and R. Stephenson. 2002. Birth spacing, sibling rivalry, and child mortality in India. *Social Science and Medicine* 55: 2107–2119.

Woldemicael, G. 2000. The effects of water supply and sanitation on childhood mortality in urban Eritrea. *Journal of Biosocial Science* 32: 207–227.

Woods, Robert. 2003. Urban-rural mortality differentials: An unresolved debate. *Population and Development Review* 29(1): 29–46.

World Resources Institute (WRI). 2002. *Drylands, people, and ecosystem goods and services: A web-based geospatial analysis.* Available at <http://www.wri.org>.

Zenger, E. 1992. Siblings' neonatal mortality risks and birth spacing in Bangladesh. *Demography* 30: 477–488.