Methods

In this section we describe the input data and the methodology used to map the Wealth Index in the Horn of Africa.

SOCIO-ECONOMIC SURVEYS

For many economists, household income would be the indicator of choice to determine economic status. It is, however, extremely difficult to measure income accurately for a number of reasons. People often try to hide their income from interviewers, for example by not providing accurate estimates; all elements of income may not be shared among household members; or income may vary considerably depending on the time of year. As well as difficulties in estimating income, it may not represent an equivalent estimate of welfare across different social contexts. For example, in pastoralist societies welfare tends to be more closely related to livestock assets held rather than to income generated. An alternative approach is to measure household consumption expenditure. Consumption expenditure estimates are generally easier to collect and more readily standardised across countries (World Bank, 2003). Moreover, consumption is thought to be a more stable measure of poverty over time than is income in agricultural economies (Deaton and Zaidi, 2002).

Income or expenditure data are collected either through household surveys that are specifically designed to collect such information (welfare monitoring surveys) or through more generic surveys, primarily designed to collect and update social and demographic indicators. Such surveys may also include socio-economic modules in their questionnaires.

In most IGAD member states, the World Bank, in collaboration with the national governments or other international agencies, has conducted a series of household surveys to collect socio-economic data at the household level. In some cases, such as Uganda, the central government conducts regular national household surveys, with very similar objectives. These surveys usually contain information on demographics, health, education, employment, income and expenditure, as well as household characteristics and agricultural and livestock assets. Whist these surveys provide good relative comparisons of welfare within a country at a particular time, the results are not comparable across countries (and not necessarily across different time periods in a particular country). This means that regional analyses and comparisons are not possible using such data.

Another type of household survey that may be able to overcome these problems of standardisation is the Demographic and Health Survey (DHS). The DHS program was established by the United States Agency for International Development (USAID) in 1984. It was designed as a follow-up to the World Fertility Survey and the Contraceptive Prevalence Survey projects. The DHS project was established at the Institute for Resource Development, Inc. (IRD), which was subsequently acquired in 1989 by Macro International Inc. (OCR Macro), the company that manages the collection, analysis and dissemination of data, and has been implemented in overlapping five-year phases. In 1993 DHS was folded into USAID's multi-project MEASURE program as MEASURE DHS+, which incorporated traditional DHS features, expanded the content on maternal and child health, and added biomarker testing to numerous surveys. The MEASURE DHS program is still funded principally by USAID with contributions from other donors.

The objectives of the DHS program are, among others, to provide decision-makers in participating countries with improved information and analyses in support of making informed policy choices; to improve coordination and partnerships in data collection at the international and country levels; and to develop in participating countries the skills and resources necessary to conduct high-quality demographic and health surveys. The basic approach of the DHS program is to collect data that are comparable across countries. To this end, standard model questionnaires have been developed, accompanied by user guides and manuals. Since 1984, more than 130 nationally representative household surveys in about 70 countries have been completed under the DHS project. Many of the countries have conducted multiple DHS surveys to establish trends, enabling them to gauge progress in their programs.

The DHS surveys are designed to collect household data on marriage, fertility, family planning, reproductive health, child health and HIV/AIDS (Rutstein and Rojas, 2003). They do not collect information on economic measures of poverty, such as income or expenditure, but data are collected about the dwelling itself, such as the source of water, type of toilet facilities, materials used to construct the house and ownership of various assets. These asset indices may be used as a proxy for the wealth status of the household (see the sub-section below on the DHS Wealth Index).

The most recent DHS survey data are accompanied by global positioning system (GPS) coordinates at the cluster level, where a cluster is usually a census enumeration area, sometimes a village in rural areas or a city block in urban areas. Collecting only one location point for a cluster greatly reduces the chance of compromising the confidentiality of respondents, but it is enough to allow the integration of multiple datasets for further analysis (Montana and Spencer, 2004).

DHS surveys have been carried out in all IGAD member states with the exception of Djibouti and Somalia. In Sudan the surveys are representative only of large administrative units so, for the current analysis, we used only the datasets for Eritrea, Ethiopia, Kenya and Uganda to develop welfare models, though predictions were made for all countries in the region. For each country, the most recent dataset available was used, specifically: Eritrea 2002, Ethiopia 2005, Kenya 2003 and Uganda 2001[2]. These four surveys included a total of 37 352 households, which were grouped into 1 519 geo-located clusters.

The DHS Wealth Index

Whilst the DHS surveys do not collect information on income or expenditure, a proxy that can be used is the Wealth Index (WI), which is constructed from a number of indicators that are thought to be correlated with a household's economic status (Rutstein and Johnson, 2004). Component indicators include, for example, possession of assets such as a television, radio, telephone or refrigerator, and variables describing the dwelling, such as the type of flooring, water supply, sanitation facilities and number of people per sleeping room.

[2] More recent datasets were subsequently released for Kenya and Uganda (2008/2009 and 2006 respectively) but they were not available at the time this analysis was conducted.

The WI, as computed from individual national surveys, cannot be used for direct cross-country comparisons since the indicators included vary from country to country. The WI is in fact a relative measure of wealth within a given survey (Rutstein and Johnson, 2004). In an FAO study, the authors discuss the WI in relation to other welfare estimates within individual countries of the IGAD region, showing a good correlation between the different measures (FAO, 2008). The objective here is to explore its value as a regionally consistent measure of welfare that can be used to produce regional poverty maps.

The WI is constructed by way of a Principal Component Analysis (PCA) on the recorded set of assets and services (Filmer and Pritchett, 2001). DHS uses the SPSS factor analysis procedure (see for example Field, 2005). This procedure first standardises the indicator variables (by calculating z-scores); then the factor coefficient scores (factor loadings) are calculated; and finally, for each household, the indicator values are multiplied by the loadings and summed to produce final values on each PCA axis. Each resulting sum is a standardised score with a mean of zero and a standard deviation of one. In the present analysis, following convention, only the first of the factors produced is used to represent the WI.

Many of the indices of poverty cannot easily be put on a quantitative scale (Rutstein and Johnson, 2004), but they can usually be coded in some way, and hence included in quantitative analyses. PCA uses only quantitative data, but this can include binary or dummy-coded qualitative data such as the presence or absence of something. Care must be taken, however, not to introduce variables falsely giving the appearance of a quantitative scale. For example, one could assign scores of '1' to the possession of a bicycle, '3' for a motorcycle and '5' for a car, but such a weighting scheme would be arbitrary and would not provide an acceptable pre-treatment of data destined for PCA. Dummy coding these same variables (i.e. creating a separate variable for each mode of transport and assigning scores of 1 or 0 to indicate presence or absence, respectively) would, however, be acceptable.

The use of a single score (PCA axis 1) for any index of wealth assumes that the majority of the variation within the dataset can be captured within this one dimension alone. Whilst PCA axis 1 by definition captures the largest percentage of the variance within the dataset, in complex data sets (such as those contributing to the WI) this may in fact be a small proportion of the total variance. There are as many axes within a PCA as there are variables (n) in the original data set, and since each axis captures some of the variance (in decreasing amounts from PCA axis 1 to PCA axis n). The larger the number of variables, the less likely it is that PCA axis 1 captures an absolute majority of the total variance (in fact it will only do so if all the indicator variables are highly correlated with each other; in which case some of these variables are redundant, and could be excluded from the questionnaires, thus saving time and resources). Using PCA axis 1 scores alone to capture poverty must therefore be approached with caution.

As mentioned above, each of the country-specific datasets has in the past been subjected to a separate PCA. This has two consequences. First, each country's PCA will be derived from a different set of input variables, only some of which might be shared with other countries. Second, even if the same set of variables were used for each country (with PCA again carried out separately for each country), the WI cannot be directly compared between countries. This is because all PCA scores are specific to the datasets being analysed, and the PCA's outputs are mean-centred scores on each of the PCA axes. To illustrate this further, consider two countries, A and B, one on average much richer than the other: there is a range of variation of wealth around each country's average wealth. Let the richest people in the poorer country be poorer than the poorest people in the richer country (i.e. there is no overlap in the wealth of any of the citizens in the two countries). PCAs carried out on each country's data will put the relatively wealthier people of each country on the positive side of its PCA axis 1[3], and the relatively poorer people on the negative side. The WI values of certain individuals within both countries may therefore be the same, despite the fact that they are not equally wealthy in absolute terms. They are only equally wealthy in relative terms, and compared only with their fellow country-men and women, not with the foreigners from the other country. Individual-country PCAs thus hide the difference in mean wealth between the two countries.

Construction of a Regional Wealth Index

The solution to this dilemma, of course, is to carry out a single PCA for all countries together. To do this we must use a set of input socio-economic variables common to all countries. Table 1shows the DHS indicators used to construct the WI in individual countries in the Horn of Africa (Eritrea, Ethiopia, Kenya and Uganda), and whether they were used in the present analysis to build the Regional WI. Some of the indicators were excluded because data were not available in all 4 surveys.

It is also important to check that combining countries' data in this way does not distort the results from the individual countries' PCA (it should not do so because the data themselves are not transformed in any way within PCA: each point stays at the same distance from all other points in the dataset within the rotated axes as it was within the un-rotated axes). We want to achieve both the same relative ranking of individuals within each country as was obtained by the country-specific PCA and also a single-scale measure of WI applicable across all countries combined. In this way, the absolute richest individuals across all countries will end up with the same WI scores; similarly for people at all other absolute levels of wealth or poverty[4].

In order to check that a regional measure of WI was accurately reflecting the previous within-country estimates, we first calculated the correlations between the regional and country-specific WIs (shown in Tables A2 to A5 in Annex 1). The latter were calculated in two ways: first using all of the variables available within that country and secondly using only those variables common to all countries. It was expected, and generally found, that correlations of Regional WI with the first sort of within-country WI were less strong than they were with the second (respectively the red and blue highlighted figures in Tables A2 to A5), but the differences were small.

[3] PCA axis scores may have reversed signs, with the richest people ending up with the highest negative scores, and poorest people with the highest positive scores. It is the absolute difference between scores that is the real index of absolute differences in the WI.

[4] Combining data across countries in this way will give the same relative weight within the PCA – and therefore the calculation of WI - to the possession, for example, of a bicycle in all countries. This may not actually reflect reality on the ground. The song "Oh Lord, won't you buy me a Mercedes Benz?" might change in the poorest country of all to one that simply requests the Almighty's supply of a decent bicycle.

Table 1. Field survey indicator variables.

Indicator	Used to compute Regional WI
Has electricity	Y
Has radio	Y
Has television	Y
Has refrigerator	Y
Has bicycle	Y
Has motorcycle	Y
Has car	Y
Has telephone	Y
Drinking water is piped in residence	Y
Drinking water is piped in public tap	Y
Drinking water from well in residence	Y
Drinking water from public well	Y
Drinking water is from surface water	Y
Drinking water is rainwater	N
Other source of drinking water	Y
Has own flush toilet	Y
Uses shared flush toilet	Y
Has pit latrine	Y
Has ventilated pit latrine	Y
Uses bush as latrine	Y
Uses other type of latrine	Y
Has dirt, earth principal floor in dwelling	Y
Has wood planks principal floor in dwelling	Y
Has tile flooring	Y
Has cement flooring	Y
Has other type of flooring	Y
Has natural material roofing	N
Has corrugate iron roofing	N
Has roofing tiles	N
Has other roofing	N
Number of members per sleeping room	N
Has domestic servant	N
Household works own or family ag. land	N

Secondly, in order to ensure compatibility across datasets we also examined the weightings of individual indicators of wealth[5] across the within-country analyses. This addresses questions of the following sort (and raised in footnote 4) 'Does a bicycle in Eritrea have the same PCA loading as a bicycle in Uganda?' In general the correlations were very good (Figure1). The exceptions to this general rule were

[5] Weightings in PCA are the natural cosines of the angle between the original (raw variable) axis and each rotated (PCA) axis. The axis of a raw variable that is highly correlated with the WI will have a small angle to the rotated axis, and will therefore end up with a high weighting (since COS(0 degrees) = 1, and COS (90 degrees) = 0).

also of interest. For example, the variable *WCBush* (indicating use of surrounding vegetation for toilet purposes) had a weighting of -0.15 in Eritrea but only of -0.06 in Uganda. This indicates a greater (negative) correlation of this variable with the WI in Eritrea than in Uganda. Generally differences were smaller than this.

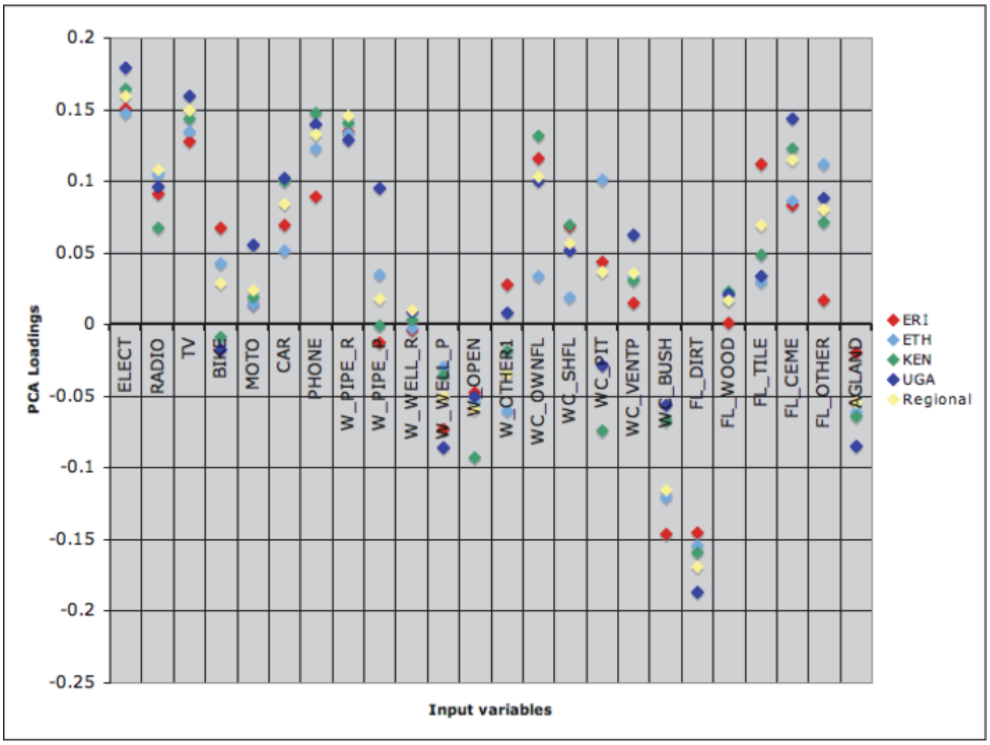
After the above comparisons were examined, it was decided that a single Regional WI provides an acceptable measure of region-wide, rather than country-specific, welfare, and all the results here are based on the Regional WI values, using the common variables listed in Table 1. Figure 2 shows the resulting Regional WI for the geo-referenced clusters.

PCA may be carried out on the raw data or on the standardised data (in the latter case the mean value for that variable is subtracted from the data value and the result is divided by the standard deviation of the variable concerned; standardised variables tend to be in the range -3.0 to +3.0). The original WI analyses first standardised the input socio-economic data (Rutstein and Johnson, 2004), and that was also the practice adopted here.

The WI values were either based on the data aggregated to the cluster level (i.e. the values of each socio-economic variable were the average for all households within each cluster), or on the individual household survey data. There were 37 352 households in the entire dataset, grouped into 1 519 geo-located clusters. Both sets of data were eventually modelled.

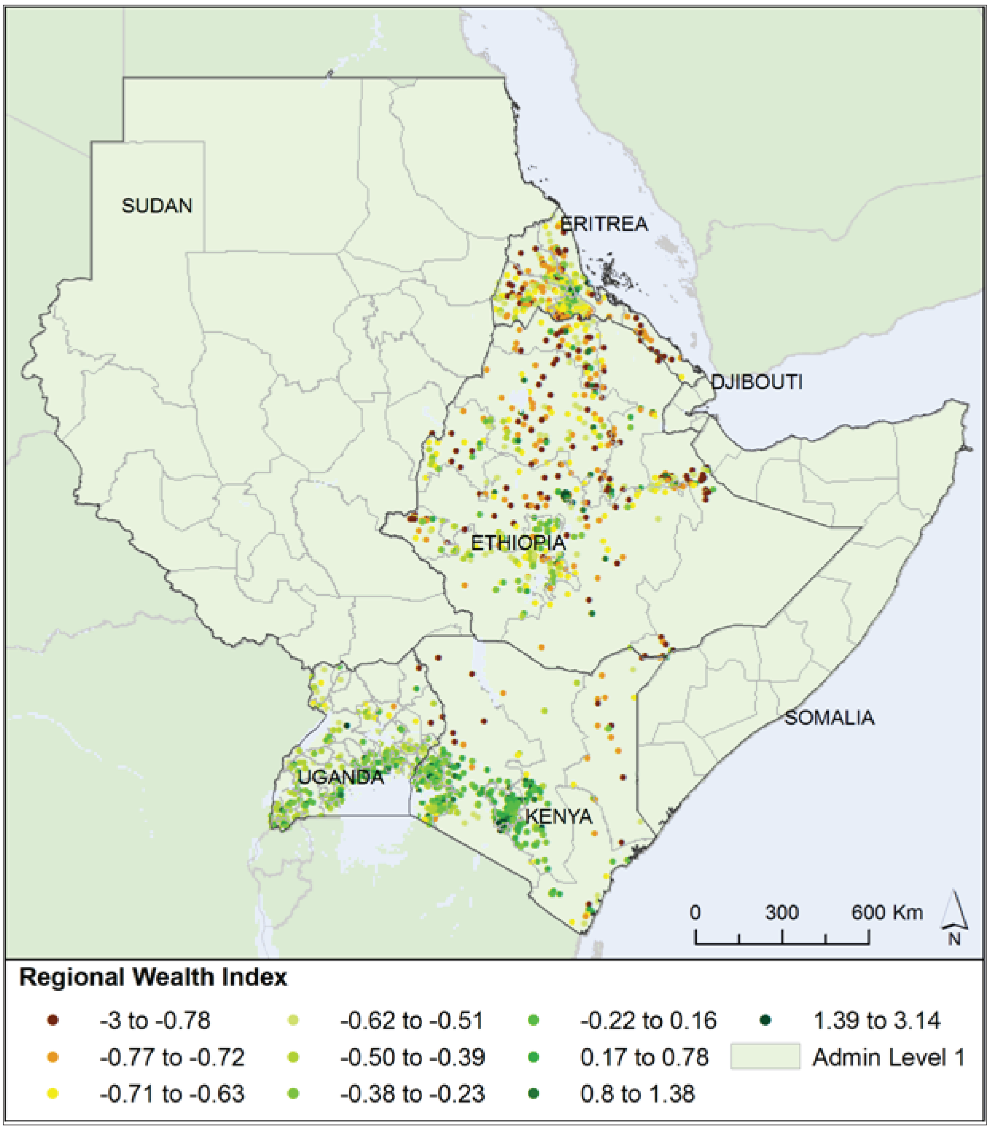
A first test was carried out to investigate whether the authors could repeat the results of the original analyses for each country. Care was taken to establish that we knew exactly how the data had been processed within the PCAs carried out previously. We were able to obtain the same values for the WI as the original survey analyses reported.

Figure 1. PCA Loadings.



Note: The figure shows PCA loadings (y-axis) of the input variables (x-axis) for each country separately, and regional loadings for the same variables. The fact that the individual country values tend to share similar PCA loadings indicates a similar contribution of each variable to the country-specific WI calculations. The fact that the regional loadings are within the range of values of the individual countries indicates that calculating a regional rather than country-specific WI does not change the relative contribution of each variable to the single, regional index.

Figure 2. Regional Wealth Index, for the 1 519 geo-referenced DHS clusters.



REMOTELY SENSED ENVIRONMENTAL DATA

Since the original environmental analysis of poverty in Uganda was conducted (FAO, 2006; Robinson *et al.*, 2007), the team in Oxford has processed the 2001 to 2005 series of satellite data from the MODerate-resolution Imaging Spectroradiometer (MODIS) sensor on board the newer Terra and Aqua satellites. These data are spectrally similar to (though not identical with) the AVHRR channels, used in the Uganda study, but offer much better geo-registration and spectral stability. In short, they are a better measure of environmental conditions for the period in question. The MODIS datasets used in this analysis include daytime and night-time land surface temperature (LST), the Middle-infrared (MIR) reflectance and the vegetation indices: Normalised Difference Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI)[6], and an Evapotranspiration layer, at a coarser resolution.

[6] NDVI is calculated as [Near infrared (NIR) – RED] / [NIR + RED], where NIR is MODIS band 2 and RED is MODIS band 1. EVI is calculated as (2.5 * [NIR – RED] / [NIR + 6*RED – 7.5*BLUE + 1], where BLUE is MODIS band 3.

In addition to the MODIS data, the team obtained actual and potential evapotranspiration and precipitation data derived from METEOSAT, provided by EARS-NL, a high-tech remote sensing company, based in the Netherlands.

Both the MODIS and the METEOSAT data (listed in Table 2), were temporal Fourier processed to extract the seasonal fingerprint of each pixel in each channel (Scharlemann *et al.*, 2008). Temporal Fourier analysis transforms time-series satellite observations into a set of (uncorrelated) sine curves, or harmonics, of different frequencies, amplitude and phases that often have a clear biological interpretation (Rogers and Williams, 1994). For each variable the Fourier process outputs the mean, amplitude and phases of annual, bi-annual and tri-annual cycles, and, in addition, the minimum, maximum and variance of the smoothed data values.

ADDITIONAL DESCRIPTOR DATA

Additional data made available to the models are also listed in Table 2. These included distance to markets, population and livestock densities and the probability of occurrence of different tsetse species. Some of these variables (e.g. livestock densities) were themselves derived by modelling and the process by which they were derived is explained in Annex 2.

The population density layers used in this analysis were those developed by CIESIN, and in particular the Gridded Population of the World (GPW) version 3, (CIESIN and CIAT, 2005) and the Global Rural and Urban Mapping Project (GRUMP) (CIESIN *et al.*, 2004). Both GPW and GRUMP gridded data are derived from a simple proportional allocation gridding algorithm of national and sub-national level population data. GPW data are available at a resolution of 2.5 arc-minutes (Balk and Yetman, 2004). GRUMP distinguishes urban and rural population from around the year 2000 and is available at the finer spatial resolution of 30 arc-seconds (Balk *et al.*, 2004). GRUMP also supplies a database of human settlements, which comprises some 55 000 cities and towns with populations of 1 000 or more, and a map of urban extents, which was derived largely from the night-time lights (Elvidge *et al.*, 1997).

In order to determine the contribution of human population, not only in terms of population density, but also in relation to its impact on the environment, it was decided to include also the Human Footprint layer, from the Last of the Wild project (WCS and CIESIN, 2002; Sanderson *et al.*, 2002). The Human Footprint (HF) layer is produced through an overlay of a number of global data layers that represent the location of various factors presumed to exert an influence on ecosystems: human population distribution, urban areas, roads, navigable rivers, and various agricultural land uses. The combined influence of these factors yields the Human Influence Index. The Human Influence Index (HII), in turn, is normalised by global biomes to create the HF data set, according to the methodology developed by Sanderson *et al.* (2002). HF values range from 1 to 100. A score of 1 in moist tropical forests indicates that that grid cell is part of the 1 percent least influenced or 'wildest' area in its biome, the same as a score of 1 in temperate broadleaf forests (although the absolute amount of influence in those two places may be quite different). The areas that have the least influence (HF grid values less than or equal to 10) are included in The Last of the Wild data set (WCS and CIESIN, 2002). For this analysis, version 1 was used.

Arguably some of these additional variables are effectively the same, but it was decided to include them all and see which ones the models selected. Variables which are perfectly correlated with those already selected will not themselves be selected in the step-wise approach adopted here, since the inclusion of effectively the 'same' data for a second time cannot possibly improve the fit of any model. The same argument applies to closely correlated variables; these too are unlikely to be selected together within the final predictor variable set unless some important differences between them allow an improvement in the overall model fit.

Table 2. Predictor variables used in the WI analysis. The 'total number of files' column indicates the number of files contained in each set of variables, which, in the case of the satellite data, results from the temporal Fourier processing. The last column indicates the number of files actually used in the model.

Data	Total no. of files	No. of files used in model	Resolution of original data	Source of original data
Vegetation Indices (NDVI, EVI)	102	60	1 km	NASA, MODIS version 4
Daytime and Night-time Land Surface Temperature (LST)				
Middle Infra-Red (MIR)				
Evapotranspiration			5 km	
Potential Evapotranspiration	10	10	3 km	EARS-NL
Actual Evapotranspiration				
Precipitation				
Global Land Cover	1	1	1 km	JRC, Global Land Cover 2000 (GLC2k)
Length of Growing period *	1	1	1 km	FAO/ILRI global livestock production systems
Digital Elevation Model (DEM)	1	1	1 km	NOAA, Global Land One-kilometer Base Elevation (GLOBE)
Slope	1	1		
Distance to Rivers	1	1	Calculated on 1 km grid	Local data where available, otherwise Africover (Eritrea and Kenya) and VMap0 (Djibouti)
Distance from Wetland	1	1	Calculated on 1 km grid	WWF, Global Land and Wetlands Database (GLWD)
Distance to Major roads	1	1	Calculated on 1 km grid	NIMA Digital Chart of the World (DCW) roads data layer, with the exception of Somalia, where the roads layer was provided by FAO-FSNAU
Distance to All Roads	1	1	Calculated on 1 km grid	Individual countries' road layers
Distance to Populated places (Gazetteer)	1	1	Calculated on 1 km grid	NIMA and GeoNames
Distance to Populated Places (Vmap0)	1	1	Calculated on 1 km grid	VMap Level 0

(cont.)

Table 2. *(cont.)*

Data	Total no. of files	No. of files used in model	Resolution of original data	Source of original data
Access to Markets *	1	1	Calculated on 1 km grid	CIESIN, Human settlements database from GRUMP, with the exception of Somalia, where market locations were provided by FAO-FSNAU
Population Density - GPW	1	1	5 km	CIESIN, Gridded Population of the World (GPWv3)
Population Density - GRUMP	1	1	1 km	CIESIN, Global Rural and Urban Mapping Project (GRUMP)
Urban Extents	1	1	1 km	CIESIN, Global Rural and Urban Mapping Project (GRUMP)
Human Footprint	1	1	1 km	WCS/CIESIN, Last of the Wild Project, v1
Night-time lights – City Lights	1	1	1 km	DMSP night-time lights
Night-time lights – Average Radiance	1	1	1 km	DMSP night-time lights
Cattle Density *	1	1	5 km	FAO, Gridded Livestock of the World (GLW)
Camel Density *	1	1		
Sheep Density *	1	1		
Goat Density *	1	1		
Pig Density *	1	1		
Chicken Density *	1	1		
Cropping *	1	1	1 km	
Tsetse *	3	3	1 km	FAO PAAT information system
IGAD Mask	1	1	1 km	Land/Water recode on NDVI image (MODIS)
Country Layer			Gridded at 1 km	FAO Global Administrative Unit Layers

Note: * indicates variables that were derived from interpolated or modelled data.

MODELLING APPROACH

As in the Uganda case study, the modelling approach was based on non-linear discriminant analysis (FAO, 2006), which allows the prediction not only of binary (presence/absence) data, but also of continuous (i.e. socio-economic) and multiple category data. In the present discriminant analyses the WI data (our ‘poverty' measure), divided into ten approximately equal-sized categories, were the dependent variable and the environmental and temporal Fourier data layers were the independent or predictor variables.

The algorithm examined the predictor variables one at a time to discover which one maximised the discriminant criterion selected by the user (in our case kappa, the index of agreement between model-predicted and observed data). This variable became the first selected variable of the eventual predicted WI map. The algorithm then went through the remaining variables, again one at a time, to select which one, in association with the first one selected, maximised the same discriminant criterion. The algorithm continued in this stepwise fashion until a pre-set number of variables

(10 in the present case) was selected. The set of selected variables was then used to make a map of the model-predicted poverty categories. More details on the discriminant analytical methods and on various metrics of model accuracy are provided by in Annexes C and D of FAO (2006).

Two types of model were run. In the first, the clustered data were used, in the second the individual household data. By definition, the predictor variable values for all households within the same cluster must be the same (because the households are given the same geo-location), and each cluster of households formed a single point (i.e. a single mean WI) which could go into only one of the ten WI categories in the model based on the clustered data. For the model based on individual household data, however, each household in a cluster might be assigned to a different WI category, depending on its individual WI. Thus one might expect a different set of predictor variables for each WI category of the clustered or individual household data.

A first model was run using the clustered data (Model 1). Then a second model was run using the individual household data (Model 2). The mapped results of these two models differed in appearance, and it was thought that this might arise because the category boundaries in the two models differed (although, it seemed, only marginally). To test whether the differences were due just to category boundaries, a third model was run using the clustered data sorted into ten categories and using the same category boundaries as were used in the household level model (Model 3).