TABLE T.3 DATA SOURCES FOR HAZARDS

Hazard type	Data source
Earthquakes	Council of the National Seismic System (as of 2002), *Earthquake Catalog,* http://quake.geo.berkeley.edu/cnss/
Cyclones	Carbon Dioxide Information Analysis Centre (1991), *A Global Geographic Information System Data Base of Storm Occurrences and Other Climatic Phenomena Affecting Coastal Zones,* http://cdiac.esd.ornl.gov/
Floods	U.S. Geological Survey (1997), *HYDRO1k Elevation Derivative Database,* http://edcdaac.usgs.gov/gtopo30/hydro/
Droughts (physical drought)	IRI/Columbia University, National Centres for Environmental Prediction Climate Prediction Centre (as of 2002), CPC *Merged Analysis of Precipitation (CMAP),* monthly gridded precipitation, http://iridl.ldeo.columbia.edu/

TABLE T.4 DATA SOURCES FOR VICTIMS, POPULATION AND VULNERABILITY VARIABLES

Theme	Data source
Victims (killed)	Université Catholique de Louvain (as of 2002), *EM-DAT: The OFDA/CRED International Disaster Database,* http://www.cred.be/ (for droughts, victims of famines were also included on a case by case basis by UNDP/BCPR)
Population (counts)	CIESIN, IFPRI, WRI (2000), *Gridded Population of the World (GPW), Version 2,* http://sedac.ciesin.org/plue/gpw/; UNEP, CGIAR, NCGIA (1996), *Human Population and Administrative Boundaries Database for Asia,* http://www.grid.unep.ch/data/grid/human.php
Vulnerability factors	
Human Development Index (HDI)	UNDP (2002), *Human Development Indicators,* http://www.undp.org/
Corruption Perceptions Index (CPI)	Transparency International (2001), *Global Corruption Report 2001,* http://www.transparency.org/
Soil degradation (% of area affected)	ISRIC, UNEP (1990), *Global Assessment of Human-Induced Soil Degradation (GLASOD),* http://www.grid.unep.ch/data/grid/gnv18.php
Other socio-economic variables	UNEP/GRID (as of 2002), *GEO-3 Data portal,* http://geodata.grid.unep.ch/ (data compiled from World Bank, World Resources Institute, FAO databases)

The list of factors to be considered for the analysis was set on the basis of the following criteria:

- *Relevance.* Select vulnerability factors (outputs orientated, resulting from the observed status of the population) not based on mitigation factors (inputs, action taken). For example, school enrollment rather than education budget.
- *Data quality and availability.* Data should cover the 1980-2000 period and most of the 249 countries and territories.

Examples of variables that were rejected for these two reasons were the percentage of persons affected by AIDS, the level of corruption and the number of hospital beds per inhabitant.

T.3.4 Data sources

Data sources ranged from universities and national scientific institutions to international data series collected by international organisations. Table T.3 presents the data sources used to obtain data on hazards.

Table T.4 presents the data sources used to obtain data on victims, population and vulnerability variables.

T.4 The Computation of Physical Exposure

T.4.1 General description

Two methods are available for calculating physical exposure. First, by multiplying hazard frequency by the population living in each exposed area. The frequencies of hazards were calculated for different strengths of event, and physical exposure was computed as in Equation 4.

EQUATION 4 COMPUTATION OF PHYSICAL EXPOSURE

EQ $$PhExp_{nat} = \sum F_i \cdot Pop_i$$

Where
$PhExp_{nat}$ is the physical exposure at national level
F_i is the annual frequency of a specific magnitude event in one spatial unit
Pop_i is the total population living in the spatial unit

A second method was used when data on the annual frequency of return of a specific magnitude event was not available. In this case (earthquake), physical exposure was computed by dividing the exposed population by the numbers of years when a particular event had taken place as shown in Equation 5.

EQUATION 5 PHYSICAL EXPOSURE CALCULATION WITHOUT FREQUENCY

EQ5 $$PhExp = \sum \frac{Pop_i}{Y_n}$$

Where
Pop_i is the total population living in a particular buffer, the radius of which from the epicentre varies according to the magnitude
Y_n is the length of time in years
PhExp is the total physical exposure of a country, in other words the sum of all physical exposure in this country

EQUATION 6 COMPUTATION OF CURRENT PHYSICAL EXPOSURE

EQ6 $$PhExp_i = \sum \frac{Pop_i}{Pop_{1995}} \cdot PhExp_{1995}$$

Where
$PhExp_i$ is the physical exposure of the current year
Pop_i is the population of the country at the current year
Pop_{1995} is the population of the country in 1995
$PhExp_{1995}$ is the physical exposure computed with population as in 1995

Once the area exposed to a hazard was computed — using UNEP/GRID-Geneva methods for earthquakes, floods and cyclones and using a method for drought from the International Research Institute for Climate Prediction (IRI) — then the exposed population was calculated for each exposed area. This number was then aggregated at the national level to come to a value for the number of exposed people over the last 21 years for each hazard type.

Depending on the type of hazard and the quality of data, different methods were applied to estimate the size of populations exposed to individual hazards. Population data was taken from CIESIN, IFPRI and WRI Gridded Population of the World (GPW, Version 2) at a resolution of 2.5'[f] (equivalent to 5 x 5 km at the equator). This was supplemented by the Human Population and Administrative Boundaries Database for Asia (UNEP) for Taiwan and CIESIN Global Population of the World Version 2 (country level data) for ex-Yugoslavia. These datasets reflect the estimated population distribution for 1995. Since population growth is sometimes very high in the 1980-2000 period, a correction factor using country totals was applied in order to estimate current physical exposures for each year as follows (see Equation 6).

Due to the resolution of the dataset, the population could not be extracted for some small islands. This has meant some small islands had to be left out of parts of the analysis. This is a topic for further research (see recommendations in the Conclusions of the Technical Annex).

The main challenge lay in the evaluation of areas exposed to particular hazard frequency and intensity. At the global scale, data was not complete. Expert opinion was used to review the process of building datasets. Of the four hazards studied, only in the case of floods was it necessary to design a global dataset. This was constructed by linking CRED information with USGS watersheds. Drought maps were provided by IRI. For the other hazards, independent global datasets had already been updated, compiled or modelled by UNEP/GRID-Geneva and were used to extract population. The Mollweide equal-area projection was used when calculations of areas were needed.

T.4.2 The case of earthquake

A choice was made to produce seismic hazard zones using the seismic catalogue of the Council of the National Seismic System. The earthquakes records of the last 21 years (1980-2000) were grouped in five magnitude classes using a buffer with a radius from the epicentre that varied according to the magnitude class (see Table T.5).

The values in Table T.5 show estimated ground-motion duration for specific acceleration and frequency ranges, according to magnitude and distance from the epicentre.[8] Numbers in bold in Table T.5 show the duration for a particular acceleration and frequency range between the first and last acceleration excursions on the record greater than a given amplitude level (for example, 0.05 g).[9]

f. GPW2 was preferred to the ONRL Landscan population dataset despite its five times lower spatial resolution (2.5' against 30") because the original information on administrative boundaries and population counts is almost two times more precise (127,093 administrative units against 69,350 units). Furthermore, the Landscan dataset is the result of a complex model which is not explained thoroughly and which is based, among other variables, on environmental data (land-cover). That makes it difficult to use for further comparison with environmental factors (circularity).

TABLE T.5 LIMITS OF THE RADIUS FOR EARTHQUAKES HAZARD

Distance (km)	Magnitude						
	5.5	6.0	6.5	7.0	7.5	8.0	8.5
10	8	12	19	26	31	34	35
25	4	9	15	24	28	30	32
50	2	3	10	22	26	28	29
75	1	1	5	10	14	16	17
100	0	0	1	4	5	6	7
125	0	0	1	2	2	3	3
150	0	0	0	1	2	2	3
175	0	0	0	0	1	2	2
200	0	0	0	0	0	1	2

Source: [Bolt et al. 1975] Acceleration > 0.05 g = ~ 0,49 m/s2, frequency > 2 Hz

According to these figures, a specific buffer distance was defined for each class of magnitude to limit the area affected by ground motions: 75 km for Magnitude ≤ 6.2, 125 km for M = 6.3 – 6.7, 150 km for M = 6.8 – 7.2, 175 km for M = 7.3 – 7.7, 200 km for M ≥ 7.8. This approach did not take into account local conditions, for instance soil or geo-tectonic characteristics.

Assuming the limitations inherent in a mortality-based conceptual model, there were three key challenges to calculating the earthquake risk index.

The first and most difficult challenge was the necessity to use a restricted time-frame for analysis of risk (1980-2001). Twenty years is a short time-span to analyze the occurrence of geological phenomena such as earthquakes, which are low frequency/high impact events. For this reason, risks are overestimated by the model for some countries and underestimated for others. Armenia provides an example of a high-impact single earthquake in a small-sized country (29,000 square kilometres), with a high population density (117 per square kilometre). The earthquake that affected this former Soviet Republic in 1998 killed 25,000 people, left 514,000 people homeless and prompted the evacuation of almost 200,000 people. The high losses recorded in this event appear to exaggerate Armenia's long-term calculated risk value, in comparison with countries known to be at risk but where no event took place during the time period used to calculate the risk model. An example of this is the Algerian earthquake in 2003, which is later than the period used in the DRI exercise. In order to partly overcome such limitation, frequency was derived using data from 1964-2000 in order to take advantage of the time-span available globally.

Secondly, in the delimitation of areas at risk from individual earthquake zones, it was not possible to consider intervening factors (such as soil types and geology) in the transmission of earthquake energy. In explaining the ground motions of earthquakes and therefore the severity of impact, soil conditions play a major role. Inclusion of this data would have allowed for a more accurate delimitation of areas and thus populations exposed to earthquake risks of various magnitudes and intensities. While values for peak ground acceleration were available from the Global Seismic Hazard Assessment Programme, they did not allow for the calculation of frequencies. Consequently, the analysis was based solely on magnitude values that were taken from the Council of the National Seismic System (CNSS).

FIGURE T.3 POPULATION, INTENSITY AND PHYSICAL EXPOSURE FOR EARTHQUAKES

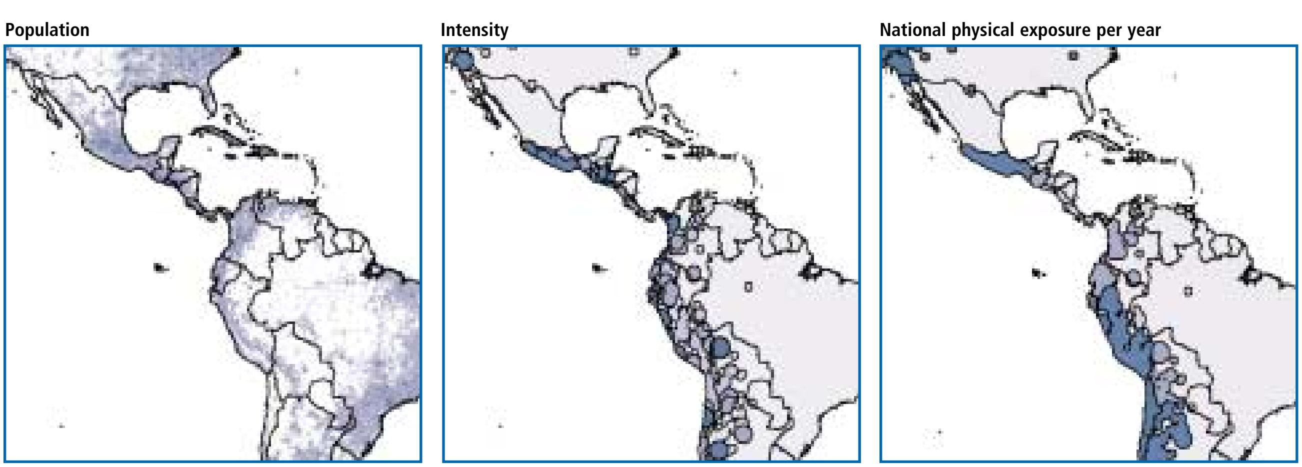


TABLE T.6 WIND SPEEDS AND APPELLATIONS

Wind speeds	Name of the phenomenon
≥ 17 m/s	Tropical storms
≥ 33 m/s	Hurricanes, typhoons, tropical cyclones, severe cyclonic storms (depending on location)
≥ 65 m/s	Super-typhoons

A third and more generic challenge for the risk model was the lack of casualty and death data and a lack of underlying socio-economic and environmental data for some countries. This is particularly problematic for mapping global earthquake risk because some gaps in national level data led to the exclusion of some countries — known to be at particularly high risk from earthquakes — from the calculation of the vulnerability indicators. This was the case for Afghanistan, Sudan, Tajikistan and Guinea. Future improvements in statistical records will enhance the scope of future assessments.

T.4.3 The case of tropical cyclone

The data used to map tropical cyclone hazard areas were produced by the Carbon Dioxide Information Analysis Centre.[10] The spatial unit is a 5 x 5 decimal degrees cell. Return probabilities were based on tropical cyclone activity over a specific record period. Exceptions were made for several estimated values attributed to areas that may present occasional activity, but where no tropical cyclones were observed during the record period.

The Saffir-Simpson tropical cyclones classification is based on the maximum sustained surface wind. Systems with winds of less than 17 m/s are called Tropical Depressions. If the wind reaches speeds of at least 17 m/s, the system is called a Tropical Storm. If the wind speed is equal to or greater than 33 m/s, the system is named, depending on its location:[g] Hurricane, Typhoon, Severe Tropical Cyclone, Severe Cyclonic Storm or Tropical Cyclone. Systems with winds reaching speeds of 65 m/s or more are called Super-typhoons.[11]

The CDIAC provided the probability of occurrence for these three types of events. The average frequency (per year) was computed using Equation 7.

To obtain physical exposure, a frequency per year was derived for each cell. Cells were divided to follow country borders, then population was extracted and multiplied by frequency in order to obtain the average yearly physical exposure for each cell. This physical exposure was then summed by country for the three types of cyclones.

Physical exposure to tropical cyclones of each magnitude was calculated for each country using Equation 5.

There is room for improving the human exposure calculation by more accurate delimitation of exposed population zones for tropical cyclone tracks. Even though accurate zoning was possible for many tropical cyclone-prone countries, data on tracks, central pressure and sustained winds were not available for some heavily populated and high-risk countries, such as India, Bangladesh and Pakistan. While these data exist they were not accessible.

T.4.4 The case of flood

The only global database on floods that was identified was the Dartmouth Flood Observatory, but this database did not cover the time period under study. Due to the lack of information on the duration and severity of floods, only one class of intensity was made. Using the EM-DAT database, a geo-reference of each recorded flood was produced and the watershed related to each flood event was identified. Watersheds affected were mapped for the period 1980-2000. A frequency was derived for each watershed by dividing the total number of events by 21 years. The watersheds were then split to follow country borders. Next, population was extracted and multiplied by the event frequency. The average yearly physical exposure was then summed at a country level using Equation 3.

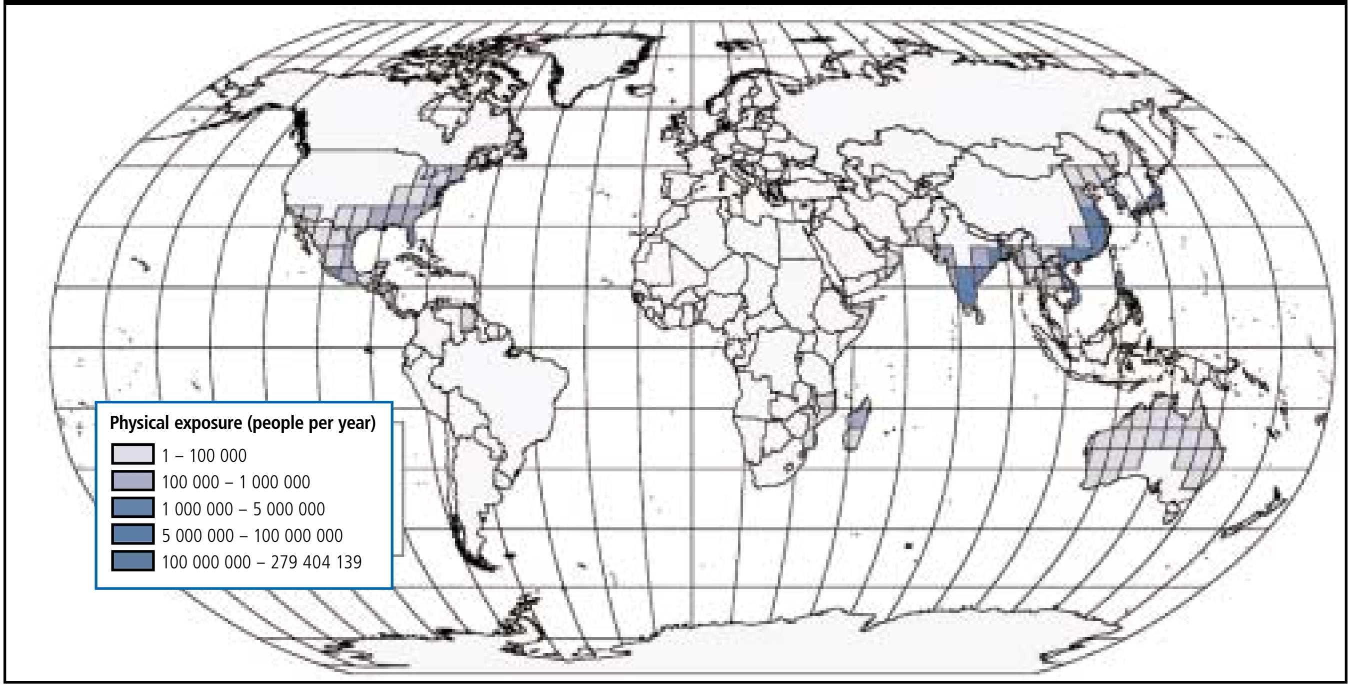
EQUATION 7 FROM PROBABILITY TO ANNUAL FREQUENCY FOR CYCLONES

EQ7 $$E(x) = \lambda = -ln(1 - P(x \geq 1))$$

Where
E(x) is the statistical expectation, i.e. the average number of events per year = λ
P(x) is the probability of occurrence

g. Hurricane: North Atlantic Ocean, Northeast Pacific Ocean east of the dateline, or the South Pacific Ocean east of 160E; Typhoon: Northwest Pacific Ocean west of the dateline; Severe tropical cyclone: Southwest Pacific Ocean west of 160E and Southeast Indian Ocean east of 90E; Severe cyclonic storm: North Indian Ocean; Tropical cyclone: Southwest Indian Ocean; Source: NOAA/AOML, FAQ: *Hurricanes, Typhoons and Tropical Cyclones,* http://www.aoml.noaa.gov/hrd/tcfaq/tcfaqA.html#A1

FIGURE T.4 AN EXAMPLE OF PHYSICAL EXPOSURE FOR TROPICAL CYCLONES



Source: Carbon Dioxide Information Analysis Centre: A Global Geographic Information System Database of Storm Occurrences and Other Climactic Phenomena Affecting Coastal Zones; CIESIN, IFPRI, WRI: Gridded Population of the World (GPW), Version 2 (population); Compilation and computation by UNEP/GRID-Geneva

Assuming the limitations inherent in a mortality-based conceptual model there were two key challenges to measuring flood risk.

First, there remains a need for refining the calculation of human exposure and vulnerability to floods in the formulation of the DRI. The use of watersheds affected by floods to delimit hazard exaggerates the extent of flood-prone areas, subsequently exaggerating human exposure and diminishing proxies of vulnerability.

Second, in the absence of historical flood event data, annual probabilities of floods should be based on hydrological models rather than being inferred from the flood entries in the EM-DAT database.

T.4.5 The case of drought

Identification of drought

The data used in this analysis consisted of gridded monthly precipitation data for the globe for the period 1979-2001. This dataset was based on a blend of surface

FIGURE T.5 POPULATION, FREQUENCY AND PHYSICAL EXPOSURE FOR FLOODS

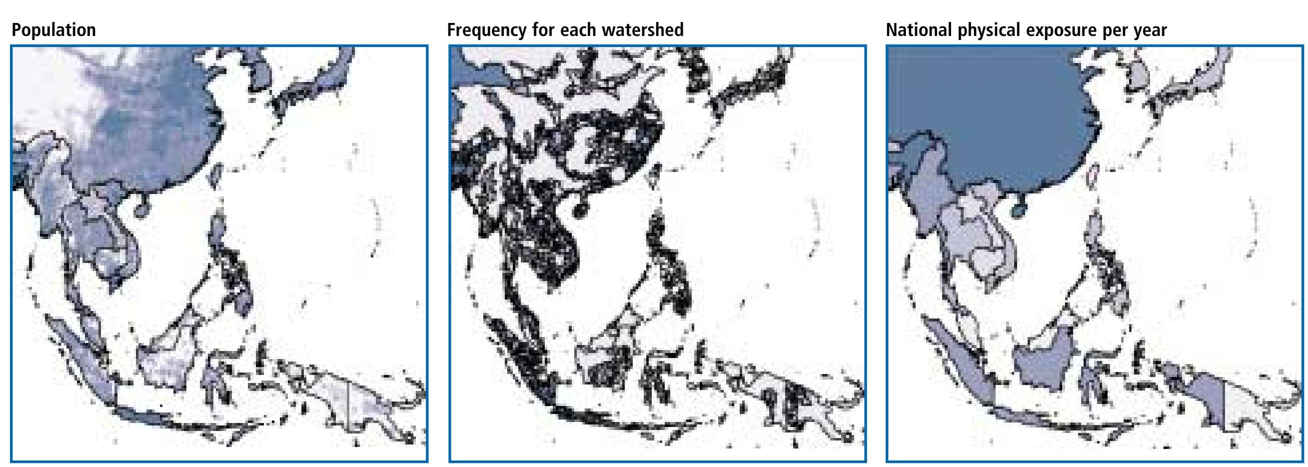


TABLE T.7 DEFINITION OF DROUGHT

Duration	Severity
3 months	90% of median precipitation 1979-2001 (-10%)
3 months	75% of median precipitation 1979-2001 (-25%)
3 months	50% of median precipitation 1979-2001 (-50%)
6 months	90% of median precipitation 1979-2001 (-10%)
6 months	75% of median precipitation 1979-2001 (-25%)
6 months	50% of median precipitation 1979-2001 (-50%)

station observations and precipitation estimates drawn from satellite observations. The first step in assessing exposure to meteorological drought was to compute, for each calendar month, the median precipitation for all grid points between the latitudes of 60S and 70N over the base period 1979-2001 (the 23-year period for which the data was available). Next, for each grid-point, the percent of the long-term median precipitation was computed for every month during the period January 1980 to December 2000. For a given month, grid-points with a long-term median precipitation of less than 0.25 mm/day were excluded from the analysis. Such low median precipitation amounts can occur either during the 'dry season' at a given location or in desert regions. In both cases our definition of drought does not apply.

A meteorological drought event was defined as having occurred when the percent of median precipitation was at or below a given threshold for at least three consecutive months. The different thresholds considered were 50 percent, 75 percent and 90 percent of the long-term median precipitation, with the lowest percentage indicative of the most severe drought according to this method. The total number of events during the period 1980-2000 were thus determined for each grid-point and the results plotted on global maps.

Computation of physical exposure

Using the IRI/Columbia University dataset, physical exposure was estimated by multiplying the frequency of hazard by the population living in an exposed area. The events were identified using different measurements, based on severity and duration as described in Table T.7. For each of the following six definitions, the frequency was then obtained by dividing the number of events by 21 years, thus providing an average frequency of events-per-year.

Physical exposure was computed, as in Equation 5, for each drought definition. The statistical analysis selected the best fit. This was achieved with droughts of three months duration and 50 percent decrease in precipitation.

EQUATION 8 ESTIMATE OF KILLED

EQ8 $$K = C \cdot (PhExp)^{\alpha} \cdot V_1^{\alpha_1} \cdot V_2^{\alpha_2} \dots \cdot V_p^{\alpha_p}$$

Where
- K is the number of persons killed by a certain type of hazard
- C is the multiplicative constant.
- PhExp is the physical exposure: population living in exposed areas multiplied by the frequency of occurrence of the hazard
- V_i are the socio-economic parameters
- α_i is the exponent of Vi , which can be negative (for ratio)

T.5 Statistical analysis: Methods and results

T.5.1 Defining a multiplicative model

The statistical analysis is based on two major hypotheses. First, that risk can be understood in terms of the number of victims of past hazardous events. Secondly, that the equation of risk follows a multiplicative model as in Equation 8.

Using logarithmic properties, the equation was re-formulated as in Equation 9. This equation creates a linear relationship between logarithmic sets of values. This allows significant socio-economic parameters V_i (with transformations when appropriate) and exponents α_i to be determined using linear regressions.

T.5.2 Detailed process

Data on victims

Numbers of killed were derived from the EM-DAT database and computed as the average number killed per year during the 1980-2000 period.

EQUATION 9 LOGARITHM PROPERTIES

EQ9 $$ln(K) = ln(C) + \alpha(PhExp) + \alpha_1 ln(V_1) + \alpha_2 ln(V_2) + \dots \alpha_p ln(V_p)$$