Electronic Supplementary Material

Modelling the influences of climate change-associated sea-level rise and socioeconomic development on future storm surge mortality

Journal: Climatic Change

Simon J Lloyd, R Sari Kovats, Zaid Chalabi

Department of Social and Environmental Health Research, London School of Hygiene and Tropical Medicine, 15-17 Tavistock Place, London. WC1H 9SH. United Kingdom.

Sally Brown, Robert J Nicholls

Faculty of Engineering and the Environment and Tyndall Centre for Climate Change Research, University of Southampton, University Road, Highfield, Southampton. SO17 1BJ., United Kingdom

Corresponding author: Sari Kovats

Phone: +44 (0) 20 7927 2962

sari.kovats@lsthm.ac.uk

Abbreviations: DIVA: Dynamic Interactive Vulnerability Assessment; EM-DAT: Emergency Events Database; GDP: Gross Domestic Product; HDI: Human Development Index; LHS: Left-Hand Side; RHS: Right-Hand Side; WBDI: World Bank Development Indicators.

Table S1: Countries by region, based on regions used in the Global Burden of Disease Study (Institute for Health Metrics Evaluation, 2010). All listed countries were used to calibrate the model and when making projections, except: countries not included when calibrating the model are marked with a hash (#), and, countries not used in projections are marked with an asterisk (*) [Countries were excluded if data were not available]. Note that, when reporting mortality results at regional level, we have aggregated Asia, Central into Europe, East; and, Latin America, Andean, - Southern, & -Tropical into a single region named Latin America, South.

Asia Pacific, High Income	**Europe, Central**	**Latin America, Andean**	**North Africa/Middle East**
Brunei Darussalam	Albania	Ecuador	**cont'd**
Japan	Bosnia and Herzegovina	Peru	Occupied Palestinian
Republic of Korea	Bulgaria	**Latin America, Central**	Territory
Singapore	Croatia	Colombia	Oman
Asia, Central	Poland	Costa Rica	Qatar
Georgia	Romania	El Salvador	Saudi Arabia
Asia, East	Serbia*	Guatemala	Syrian Arab Republic
China	Slovenia	Honduras	Tunisia
Dem P's Rep of Korea	**Europe, East**	Mexico	Turkey
Hong Kong*	Estonia	Nicaragua	United Arab Emirates
Macao*	Latvia	Panama	Western Sahara
Taiwan*	Lithuania	Venezuela	Yemen
Asia, South	Russian Federation	**Latin America, South**	**Oceania**
Bangladesh	Ukraine	Argentina	Fiji
India	**Europe, West**	Chile	French Polynesia
Pakistan	Belgium	Uruguay	Kiribati
Asia, South East	Cyprus	**Latin America, Tropical**	Marshall Islands
Cambodia	Denmark*,[2]	Brazil	Micronesia (Federated
Indonesia	Finland	**North America, High**	States of)
Malaysia	France	**Income**	Nauru
Maldives	Germany	Canada	New Caledonia
Myanmar	Greece	United States	Palau
Philippines	Iceland	**North Africa/Middle East**	Papua New Guinea
Sri Lanka	Ireland	Algeria	Samoa
Thailand	Israel	Bahrain	Solomon Islands
Timor-Leste	Italy	Egypt	Tonga
Viet Nam	Malta	Iran (Islamic Republic of)	Tuvalu
Australasia	Monaco	Iraq	Vanuatu
Australia	Netherlands	Jordan	
New Zealand	Norway	Kuwait	**(continued over page)**
Caribbean	Portugal	Lebanon	
Caribbean cluster[1]	Spain	Libyan Arab Jamahiriya	
	Sweden	Morocco	
	United Kingdom		

Table S1, continued.

Sub-Saharan Africa, Central	**Sub-Saharan Africa, East cont'd**	**Sub-Saharan Africa, South**	**Sub-Saharan Africa, West cont'd**
Angola	Kenya	Namibia	Ghana
Congo	Madagascar	South Africa	Guinea
Dem Rep of the Congo	Mauritius#	**Sub-Saharan Africa, West**	Guinea-Bissau
Equatorial Guinea	Mozambique	Benin	Liberia
Gabon	Seychelles	Cameroon	Mauritania
Sub-Saharan Africa, East	Somalia	Cape Verde	Nigeria
Comoros	Sudan	Cote d'Ivoire	Sao Tome and Principe
Djibouti	United Republic of Tanzania	Gambia	Senegal
Eritrea			Sierra Leone
			Togo

[1]Caribbean countries were grouped into a single cluster for model calibration due to data availability. Countries in the cluster are: Anguilla[+], Antigua and Barbuda, Aruba, Bahamas, Barbados, Belize, Bermuda, Cayman Islands[+], British Virgin Islands, Cuba, Dominica, Dominican Republic, French Guiana[+], Grenada, Guadeloupe[+], Guyana, Haiti, Jamaica, Montserrat[+], Martinique[+], Netherlands Antilles, Puerto Rico, Saint Kitts and Nevis, St. Lucia, St Vincent and the Grenadines, Suriname, Trinidad and Tobago, Turks and Caicos Islands[+], US Virgin Islands. Countries were included individually when making projections, except those marked with a plus sign (+), for which projections were not made.

[2] Denmark was not included in projections due to inconsistencies in territorial definitions across scenarios.

INSTITUTE FOR HEALTH METRICS EVALUATION. 2010. *Global Burden of Disease Study: Operations Manual 2009* [Online]. Available: http://www.globalburden.org/GBD_Study_Operations_Manual_Jan_20_2009.pdf [Accessed December 19th 2012].

Table S2: Baseline (*j*=2000) average annual population at risk of storm surge exposure as modelled by DIVA, average annual surge-specific mortality derived from EM-DAT (standardized to population in the year 2000), and mortality risk per million at risk of exposure, at regional level.

Region	Baseline average annual risk of surge exposure	Baseline average annual killed by surge, standardized to population in 2000	Baseline surge mortality risk per million at risk of exposure
Asia Pacific, High Income	2,896	52.1	17,998.4
Asia, Central	540	0.0	0.0
Asia, East	659,115	271.0	411.1
Asia, South	825,023	18,731.8	22,704.6
Asia, Southeast	1,464,261	4,052.1	2,767.4
Australasia	2,922	2.4	835.5
Caribbean	3,814	196.8	51,606.1
Europe, Central	1,136	0.7	592.7
Europe, Eastern	46,032	2.3	50.1
Europe, Western	13,941	23.4	1,679.4
Latin America, Andean	2,666	10.1	3,777.5
Latin America, Central	6,443	375.2	58,229.2
Latin America, Southern	4,705	0.0	0.0
Latin America, Tropical	10,572	0.3	23.9
North Africa / Middle East	65,438	3.7	55.9
North America, High Income	26,008	50.2	1,931.4
Oceania	845	16.7	19,738.2
Sub-Saharan Africa, Central	861	0.7	776.2
Sub-Saharan Africa, East	274,157	64.4	234.9
Sub-Saharan Africa, Southern	153	0.0	0.0
Sub-Saharan Africa, West	29,745	0.0	0.0
Global	**3,441,274**	**23,853.8**	**6,931.7**

Table S3: The 7 GCMs used in this paper.

MRI-CGCM2.3.2
IPSL-CM4
CSIRO-MK3
ECHAM5-MPI-OM
UKMO-HadCM3
UKMO-HadGEM
NCAR CCSM3

For details on individual models, see:

Randall, D.A., R.A. Wood, S. Bony, R. Colman, T. Fichefet, J. Fyfe, V. Kattsov, A. Pitman, J. Shukla, J. Srinivasan, R.J. Stouffer, A. Sumi and K.E. Taylor, 2007: Cilmate Models and Their Evaluation. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M.Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

For details of the associated sea-level rise scenarios, see:

Brown, S., Nicholls, R. J., Lowe, J. A. & Hinkel, J. 2013. Spatial variations of sea-level rise and impacts: An application of DIVA. *Climatic Change,* (forthcoming). DOI: 10.1007/s10584-013-0925-y

Table S4: Regional level projections of population (thousands) and Human Development Index (HDI) analogue (as population weighted average) for baseline, 2050 and 2080. (Note that data includes only countries for which projections were made).

Region	Population (000s)			HDI		
	Baseline	2050	2080	Baseline	2050	2080
Asia Pacific, High Income	177,782,421	185,109,174	175,139,847	0.88	1.00	1.00
Asia, Central	5,289,937	4,038,361	3,463,931	0.49	0.79	0.91
Asia, East	1,303,883,210	1,321,223,430	1,085,734,830	0.54	0.87	0.98
Asia, South	1,300,287,100	1,859,036,800	1,693,470,500	0.43	0.71	0.88
Asia, Southeast	530,834,036	723,985,173	653,067,977	0.52	0.84	0.95
Australasia	22,512,966	33,701,855	31,560,441	0.90	1.00	1.00
Caribbean	38,174,654	44,915,535	42,723,398	0.57	0.79	0.90
Europe, Central	82,918,029	75,669,709	64,794,812	0.70	0.96	1.00
Europe, Eastern	203,621,260	170,566,453	153,170,216	0.68	0.94	0.98
Europe, Western	374,630,959	410,883,024	407,621,269	0.83	1.00	1.00
Latin America, Andean	37,833,520	57,604,010	59,056,170	0.63	0.91	0.99
Latin America, Central	200,113,548	310,279,301	316,064,008	0.61	0.87	0.97
Latin America, Southern	54,858,783	75,511,878	76,122,635	0.74	0.98	1.00
Latin America, Tropical	169,385,000	223,467,300	210,984,300	0.57	0.90	1.00
North Africa / Middle East	314,777,540	441,733,270	454,303,440	0.93	1.00	1.00
North America, High Income	383,026,306	659,840,237	691,381,898	0.56	0.82	0.95
Oceania	7,643,246	15,543,614	16,190,872	0.45	0.70	0.84
Sub-Saharan Africa, Central	69,005,974	166,715,574	193,315,935	0.27	0.53	0.73
Sub-Saharan Africa, East	145,668,350	232,539,576	237,246,005	0.29	0.58	0.77
Sub-Saharan Africa, Southern	46,016,828	30,536,239	26,414,898	0.58	0.84	0.94
Sub-Saharan Africa, West	217,830,590	352,625,078	347,990,731	0.32	0.58	0.76
Global total	5,686,094,257	7,395,525,591	6,939,818,113	-	-	-

Appendix S1: The DIVA flood model

Abbreviations:
DIVA: Dynamic Interactive Vulnerability Model
GDPC: Gross domestic product per capita
GLOBE: Global Land One km Base Elevation project
HDI: Human development index

Symbols:
$X_{i,j}$: Average annual people exposed to surges

A description of the DIVA model which estimates the average annual number of people potentially exposed to surges ($X_{i,j}$), (i.e. expected number of people flooded per year) is detailed in the supplementary material of Brown et al. (2013). This is reiterated, with additional context for the health model.

To determine the average annual number of people exposed to surges due to rising sea-levels and socio-economic change, a global vulnerability assessment – the **D**ynamic **I**nteractive **V**ulnerability **A**ssessment (DIVA) model was used. DIVA is an integrated model of coastal systems that assesses biophysical and socio-economic impacts of sea-level rise and socio-economic development (Vafeidis et al. 2008; Hinkel and Klein 2009; Hinkel 2005). The model breaks the world's coast (excluding Antarctica) into 12,148 linear segments, each associated with over 100 pieces of physical, ecological and socio-economic coastal characteristics (Vafeidis et al., 2008; Hinkel and Klein 2009). Land use is categorised by 16 different types, based on the dominant land use within a 0.5º x 0.5º global grid cell. The socio-economic scenarios are formed from land-use class, population growth (assumed at national level growth, but with the base year representing coastal population density (from CIESIN et al. 2000), thus taking into account the higher levels of population density often found in coastal zones). The climate scenarios represent changes in mean surface temperatures and sea levels, calculated on a 5º x 5º grid. For the climate scenarios, data is input every five years, from 1995 to 2100.

Flooding and submergence of coastal zones is caused by mean sea-level rise and extreme events, and also vertical land movement. Extreme events are produced by a combination of storm surges and astronomical tides. For each segment, relative sea-level rise is calculated by combining sea-level change with vertical land movement. Vertical land movement is a combination of glacial-isostatic adjustment according to the geophysical model of Peltier (2000a, b) - assumed uniform natural subsidence in deltas of 2 mm/yr – and that of eustatic sea-level rise. DIVA downscales the global sea-level scenarios to segment level. As sea levels rise, the return period of an extreme event is reduced, so that extreme water levels happen more often. This can be measured per segment of coast on an exceedance curve. It is assumed that there is no increase in the frequency or intensity of coastal storms, so present storm surge characteristics are simply displaced upwards with the rising sea level on the exceedance curve. Surges represent return periods of the 1-in-1 to the 1-in-1000 year floods. DIVA also considers flooding in the coastal part of rivers, known as the backwater effect. The total relative sea-level rise per segment under extreme conditions was mapped against the low-lying

coastal zone (Vafeidis et al. 2008) based on the Global Land One km Base Elevation Project (GLOBE) topographic dataset, with a resolution of 1km, and population density.

Taking account of flooding and erosion under a range of socio-economic and climate scenarios, impacts are also dependant on the type of adaptation strategy applied and the level of protection this affords. The model incorporates two engineered adaptation strategies (henceforth 'sea-based strategies'): Without upgrade to protection and with upgrade to protection. Sea dikes are already a prominent form of defence, yet no global database of coastal defences exists. Hence, for the base year (1995), these were modelled based on a demand for safety (Tol and Yohe, 2007): where there are sufficiently high population densities (> 1 person / km^2) plus capital, and the standard of protection increases as population density increases. Initial protection levels are principally based on population density and gross domestic product per capita (GDPC), and secondly agricultural land values and a safety margin for nourishment. These protection levels are then kept constant throughout the timeframe of the study. For the upgrade in protection, the same baseline protection is assumed, but as population expands there is a higher demand for protection. Nourishment for erosion would also be undertaken.

These adaptation strategies differ from the health model's adaptation strategies, which are based on a land-based response (henceforth 'land-based strategies'), such as early warning systems or people moving to shelters. This latter strategy is based on variables from the Human Development Index (HDI), and is not linked to the DIVA model. However, both models do have common inputs of population and gross domestic product, based on the SRES scenarios of Nakicenovic and Swart (2000).

Brown S, Nicholls RJ, Lowe JA, Hinkel J (2013) Spatial variations of sea-level rise and impacts: An application of DIVA. Climate Change. (forthcoming) DOI: 10.1007/s10584-013-0925-y

Center for International Earth Science Information Network (CIESIN), Columbia University, International Food Policy Research Institute (IFPRI), World Resources Institute (WRI) (2000) Gridded Population of the World (GPW), Version 2. Palisades. CIESIN, New York

Hinkel J (2005) DIVA: An iterative method for building modular integrated models. Adv in Geosci 4:45-50

Hinkel J, Klein RJT (2009) The DINAS-coast project: Developing a tool for the dynamic and interactive assessment of coastal vulnerability. Glob Environ Change 19 (3):384-395

Nakicenovic N, Swart R (2000) Emissions scenarios. Special report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK

Peltier WR (2000a) Global glacial isostatic adjustment and modern instrumental records of relative sea level history. In: Douglas BC, Kearny MS, Leatherman SP (eds) Sea level rise; history and consequences. Academic Press, San Diego, pp 65-95

Peltier WR (2000b) ICE4G (VM2) glacial isostatic adjustment corrections. In: Douglas BC, Kearny MS, Leatherman SP (eds) Sea level rise; history and consequences. Academic Press, San Diego. On CD

Tol RSJ, Yohe GW (2007) The weakest link hypothesis for adaptive capacity: An empirical test. Glob Environ Change 17:218-227

Vafeidis AT, Nicholls RJ, McFadden L, Tol RSJ, Hinkel J, Spencer T, Grashoff PS, Boot G, Klein RJT (2008) A new global coastal database for impact and vulnerability analysis to sea-level rise. J Coast Res 24 (917-924)

Appendix S2: Implications of multiple uses of the same underlying data

Figure 1 in the main text shows how the underyling raw data was transformed and used to fit the mortality risk model. As is shown is the figure, on a number of occasions, the same underlying data enters the model in various places. For instance, population is used to standardize the mortality data, as an input into DIVA for estimating exposure, and as indicator of coping capacity in the mortality risk equation.

This problem has two origins; one general, and one particular to our mortality risk model. Firstly, this type of problem is not uncommon when using integrated assessment methods (IAMs) where outputs from one model (M_1) are used as inputs to another model (M_2). The underpinning assumption in this approach is that M_2 does not depend on a variable which is used in model M_1. This is violated in our model because population and GDP are elements of both M_1 (DIVA) and M_2 (the mortality risk model). As a result, they are present in both sides of Equation (1).

Secondly, data on which to base a model to estimate future surge mortality is extremely limited. For example, as discussed in the paper, surge-specific mortality data is not available; data for characterising vulnerability are limited to a handful of potential proxies. Consequently we put a limited set of data to multiple uses when developing the mortality risk model.

The general implications of such data 're-use' can be shown mathematically.

Equation (1) (i.e. the mortality risk model) can be written in the abstract form as:

$$f(y(x), u) = h(x, z) \tag{S2.1}$$

where (i) x (e.g. population; GDP) is present on both the left and right hand sides of Equation(1), (ii) y is meant to be an independent variable (e.g. DIVA output) but in this case it is a function of x, (iii) u and z represent variables which are not shared by the two sides of the equation, (iv) f and h are non-linear function of their arguments.

At issue is that y(x) (from the perspective of equation (1)) is unknown (i.e. a black–box).

Differentiating both sides of the Equation (1) by x gives

$$\frac{\partial f}{\partial y}\frac{dy}{dx} = \frac{\partial h}{\partial x} \tag{S2.2}$$

From the left hand side of Equation (1) we have

$$f(y) = ln\left(\frac{a}{y}\right)$$

(S2.3)

Therefore

$$\frac{\partial f}{\partial y} = -\frac{1}{y}$$

(S.24)

From the right hand side of Equation (1) in the paper

$$h(x) = b + \beta_3\, x + \beta_4 x^2$$

(S2.5)

where b here is some constant (i.e. independent of x).

Differentiating Equation (S2.5) with respect to x gives

$$\frac{dh}{dx} = \beta_3 + 2\beta_4 x$$

(S2.6)

Substituting Equations (S2.4) and (S2.6) into Equation (S2.2) gives

$$-\frac{1}{y}\frac{dy}{dx} = \beta_3 + 2\beta_4 x$$

(S.2.7)

The implicit assumption of our approach is that

$$\frac{dy}{dx} = 0$$

(S.2.8)

i.e. y is independent of x. If this is the case then

$$\beta_3 + 2\beta_4 x = 0$$

(S.2.9)

which strictly holds only when

$$x = -\frac{\beta_3}{2\beta_4}$$

(S.2.10)

From Table 1 in the main paper this means it holds for an HDI of 0.669. This suggests the model may be expected to make reasonable predictions when HDI is in the 'vicinity of this 0.669 but biases are introduced as HDI deviates from this estimate.

This HDI corresponds to the turning point (i.e. the maximum) of the parabola described by equation (1) (See Appendix S6, Figure S6.1). That is, the point at which 'development' shifts from increasing risk to decreasing it, and at this point the model is least biased. As countries move closer to the extremes (i.e. least or most 'developed'), biases increase.

Appendix S3: Standardizing EM-DAT mortality data

We standardized deaths in any given event in our data set to population in the year 2000 using:

$$\overline{m}^{(S)}_{i,2000,s} = m^{(S)}_{i,j,s} \times \frac{P_{i,2000}}{P_{i,j}} \qquad \text{(S3.1)}$$

$$\overline{m}^{(L)}_{i,2000,l} = m^{(L)}_{i,j,l} \times \frac{P_{i,2000}}{P_{i,j}} \qquad \text{(S3.2)}$$

where:

$\overline{m}^{(S)}_{i,2000,s}$ is all-cause cyclone mortality in 'small' event s in country i, standardized to the population in the year 2000.

$m^{(S)}_{i,j,s}$ is all-cause cyclone mortality in 'small' event s, in country i, in year j

$\overline{m}^{(L)}_{i,2000,l}$ is all-cause cyclone mortality in 'large' event l in country i, standardized to the population in the year 2000.

$m^{(L)}_{i,j,l}$ is all-cause cyclone mortality in 'large' event l in country i, in year j

We then used the standardized mortality estimates to find average annual, all-cause cyclone mortality at the year 2000, standardized to population in the year 2000.

$$C_{i,2000} = \frac{\sum_{s=1}^{S} \overline{m}^{(S)}_{i,2000,s}}{y^{(S)}} + \frac{\sum_{l=1}^{L} \overline{m}^{(L)}_{i,2000,l}}{y^{(L)}} \qquad \text{(S3.3)}$$

where:

$C_{i,2000}$ is average annual, all-cause cyclone mortality at the year 2000, standardized to population in the year 2000.

$y^{(S)}$ is number of years covered by the dataset of 'small' events

$y^{(L)}$ is number of years covered by the dataset of 'large' events

Table S3.1 shows event and standardized mortality data at the regional level.

Table S3.1: Baseline (*j*=2000), regional average annual number of events and all-cause cyclone mortality estimates (standardized to population in the year 2000), for 'small' (< 200 deaths) and 'large' (≥ 200 deaths) events, and for all events (i.e. small and large combined), derived from EM-DAT. See main paper for details.

	'Small' events		'Large' events		All events	
Region	**Events**	**Killed**	**Events**	**Killed**	**Events**	**Killed**
Asia Pacific, High Income	3.24	64.5	0.02	13.3	3.26	77.8
Asia, Central	0	0.0	0.00	0.0	0.00	0.0
Asia, East	7.67	224.0	0.20	113.5	7.86	337.5
Asia, South	2.95	157.5	0.80	20655.6	3.76	20813.1
Asia, Southeast	10.33	303.7	1.10	4311.3	11.43	4614.9
Australasia	1.43	3.6	0.00	0.0	1.43	3.6
Caribbean	8.29	53.5	0.15	173.8	8.43	227.3
Europe, Central	0.52	1.0	0.00	0.0	0.52	1.0
Europe, Eastern	0.90	3.1	0.00	0.0	0.90	3.1
Europe, Western	5.62	34.9	0.00	0.0	5.62	34.9
Latin America, Andean	0	0.0	0.02	13.2	0.02	13.2
Latin America, Central	5.48	71.1	0.34	991.8	5.82	449.1
Latin America, Southern	0	0.0	0.00	0.0	0.00	0.0
Latin America, Tropical	0.095	0.3	0.00	0.0	0.10	0.3
North Africa / Middle East	0.29	5.2	0.00	0.0	0.29	5.2
North America, High Income	2.95	32.3	0.02	42.7	2.98	75.0
Oceania	2.62	19.4	0.00	0.0	2.62	19.4
Sub-Saharan Africa, Central	0.048	0.7	0.00	0.0	0.05	0.7
Sub-Saharan Africa, East	2.24	50.9	0.07	20.6	2.31	71.6
Sub-Saharan Africa, Southern	0	0.0	0.00	0.0	0.00	0.0
Sub-Saharan Africa, West	0.048	0.0	0.00	0.0	0.05	0.0
Global total	**54.72**	**1,025.7**	**2.7**	**26,335.8**	**57.4**	**26,747.7**

Appendix S4: Estimating storm surge-attributable mortality

Based on the literature, it was assumed that in the least developed countries about 90% of cyclone mortality is attributable to surge, whereas in more developed countries it is estimated to be to be approximately 67%.

We assumed that in countries with an HDI-analogue in the lower and upper quartiles, surge accounted for 90% and 67% of cyclone deaths respectively, and, in between these quartiles the surge-attributable percent drops linearly from 90% to 67%.

Baseline storm surge-attributable mortality, $M_{i,2000}$ was estimated using the baseline Human Development Index analogue, $H_{i,2000}$ and average annual, all-cause cyclone mortality as at the year 2000, standardized to population in the year 2000, $C_{i,2000}$.

$$M_{i,2000} = \left\{ \begin{array}{cc} 0.9 \times C_{i,2000} & if\ H_{i,2000} < 0.449 \\ -0.797 \times H_{i,2000} + 1.258 \times C_{i,2000} & if\ 0.449 \le H_{i,j} \le 0.742 \\ 0.67 \times C_{i,2000} & if\ \ H_{i,2000} > 0.742 \end{array} \right\} \quad \text{(S4.1)}$$

Appendix S5: Calculating the analogue of the Human Development Index

The Human Development Index (HDI) is an indicator of development that combines data on Gross Domestic Product per capita (GDP/capita), educational attainment, and life expectancy into a single indicator of social and economic conditions. For details, see http://hdr.undp.org/en/statistics/hdi/ (accessed Nov 20th, 2012)

Most existing climate-health impact assessment work that includes socioeconomic influences on health uses GDP/capita as a proxy for social and economic conditions. Thus the use of the HDI is an advance. As the education data conventionally used in the HDI was not available as projections ('mean years of education' and 'expected years of education', for further details, see: http://hdr.undp.org/en/media/HDR_2011_EN_TechNotes.pdf (accessed Nov20th, 2012)), we used years of education at 25 years of age. Further, the HDI conventionally includes GDP/capita as Purchasing Power Parity (PPP); however, we used GDP/capita as Market Exchange Rate (MER) as this was used in the DIVA model. Hence we used a modified version of the HDI which we refer to as the HDI analogue.

The 2010 HDI report [ref] normalised each component variable using the range observed over 1980 to 2010, and we used a similar strategy to estimate baseline HDI for model fitting.

For GDP/capita as MER, we assessed the range over 1980 to 2009 (2010 data was not available at that time) in the World Bank Development Indicators (WBDI) [ref]. We excluded 'tax-haven' countries that had low population but exceptionally high GDPs. The maximum was in Norway in 2008 (US$ 42,143) and the minimum in Liberia in 1995 (US$62) [in US$ for the year 2000]. The HDI uses ln(GDP/capita) and we normalized country-level data for the year 2000 using:

$$G_i = \frac{\ln(g_i) - \ln(62)}{\ln(42143) - \ln(62)} \quad \text{(S5.1)}$$

where:

G_i is the GDP is the normalized GDP index for the year 2000 in country i

g_i is GDP/capita for the year 2000 in country i

For years of education, we used the range employed in the HDI 2010 report for the years 1980 to 2010; minimum=0 and maximum=13.2 years. While we used a different education variable than the conventional HDI, we considered this to be a reasonable range over which to normalize our data, which ranged from 0.824 to 12.25 years in the year 2000. We normalized country level data using (where the minimum years of education is 0) :

$$S_i = \frac{s_i}{13.2} \quad \text{(S5.2)}$$

where:

S_i is the normalized education index for the year 2000 in country i

s_i is average years of education at the age of 25 for the year 2000 in country i

For life expectancy at birth, we used the range reported in the 2010 HDI report for the period 1980-2010; minimum=20 years, maximum=83.2 years. Our data for the year 2000 ranged from 42.01 to 81.83 years. We normalized country-level data using:

$$L_i = \frac{l_i - 20}{83.2 - 20} \tag{S5.3}$$

where:

L_i is the normalized life expectancy index for the year 2000 in country i

l_i is life expectancy at birth for the year 2000 in country i

We then calculated the HDI analogue using:

$$H_i = \sqrt[3]{G_i \times S_i \times L_i} \tag{S5.4}$$

For projections, we normalized data for future years for each component index using the above maximums and minimums (i.e. consistently based on 1980 to 2000); this maintained a common standard over time. We rounded future HDI analogues that were greater than 1 down to 1. No future HDI analogues were equal to or less than 0.

Appendix S6: Model calibration

We based our model on the model developed by Patt et al (2010). Patt et al modelled log mortality risk (referred to as 'vulnerability' in the original paper) as a function of 'physical exposure' and 'sensitivity to exposure'. (Note that Patt et al also modelled number of people 'affected' by disasters using a similar model)

'Physical exposure' was represented as the number of disaster events over a given time period and national population. Conceptually, it was expected that as the number of events increases, coping capacity decreases, and hence mortality risk increases. For population, it was 'expected that larger countries are likely to experience disasters over a smaller proportion of their territory or population, and also benefit from potential economies of scale in their disaster management infrastructure' (Patt et al., 2010); thus as population increases, it is expected that mortality risk will decrease.

'Sensitivity to exposure' was modelled using the Human Development Index (HDI) and the fertility rate (Patt et al also tested other variables (the particular variables tested were not specified) but found they did not improve model fit). Conceptually (based on previous observations), it was expected that with 'development' mortality risk would initially increase before decreasing; thus HDI was modelled in quadratic form. Fertility $F_{i,j}$ was included as an indicator of women's empowerment, which was expected to reduce mortality risk.

To calibrate our model we tested various forms of Patt el al's (2010) general model, a selection of which are shown in Table S5.1.

All the models shown (1 to 7) had a reasonably good statistical fit ($R^2 \approx 0.4$). In all models, the standardized regression coefficients for $\log(E_i)$ and $\log(P_{i,j})$ were stable, and their significance (as indicated by p-values) was high. Additionally, the signs of the coefficients were as expected according to the underlying theory (positive for $\log(E_i)$, and negative for $\log(P_{i,j})$). As a result, we could not decide on the final model by simply comparing R^2 values. We present below the rationale for selecting the final model.

The variable $log(F)$ was consistently found to have little explanatory power and was not included in the final model.

Thus the choice of model hinged on the HDI analogue ($H_{i,j}$). Model 1 corresponds to Patt et al's final mortality risk model. The sign of the parameter for the quadratic term is as expected. However, as the relative importance of the HDI variables compared to that the other variables was low, and – more so – because the form of the fit was inconsistent with theory (i.e. the location of the peak of the parabola; not shown) we tested other forms.

Table S6.1: Standardized regression coefficients[a], p-values (in brackets), and adjusted R^2 obtained during model calibration for a selection of models.

Variable[b]	Model 1[c]	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
log(E)	0.24 (0.001)	0.23 (0.001)	0.24 (0.001)	0.25 (0.001)	0.24 (0.001)	0.25 (0.001)	0.25 (0.001)
log(P)	-0.64 (0.0001)	-0.64 (0.0001)	-0.63 (0.0001)	-0.62 (0.0001)	-0.64 (0.0001)	-0.65 (0.0001)	-0.63 (0.0001)
log(H)	-0.20 (0.432)	0.22 (0.001)					
$[\log(H)]^2$	-0.44 (0.055)		-0.25 (0.0001)				
H				1.28 (0.001)	0.18 (0.007)		1.25 (0.001)
H^2				-1.10 (0.004)		0.15 (0.03)	-1.08 (0.003)
log(F)	-0.004 (0.97)			0.027 (0.803)			
Adjusted R^2	0.42	0.41	0.42	0.42	0.39	0.38	0.43

[a] The standardized regression coefficients quantify the change in the LHS of the equation relative to its standard deviation when a given RHS variable is changed by one standard deviation; the greater the absolute value of the statistic, the more responsive the LHS is to the variable. For example, for model 1, 0.24 means that a 1 standard deviation change in $E_i$is associated with a 0.24 standard deviation change in log mortality risk.

[b] Variables are: E is average annual events; P is population; H is the Human Development Index analogoue, F is Total Fertility Rate. See main text for details.

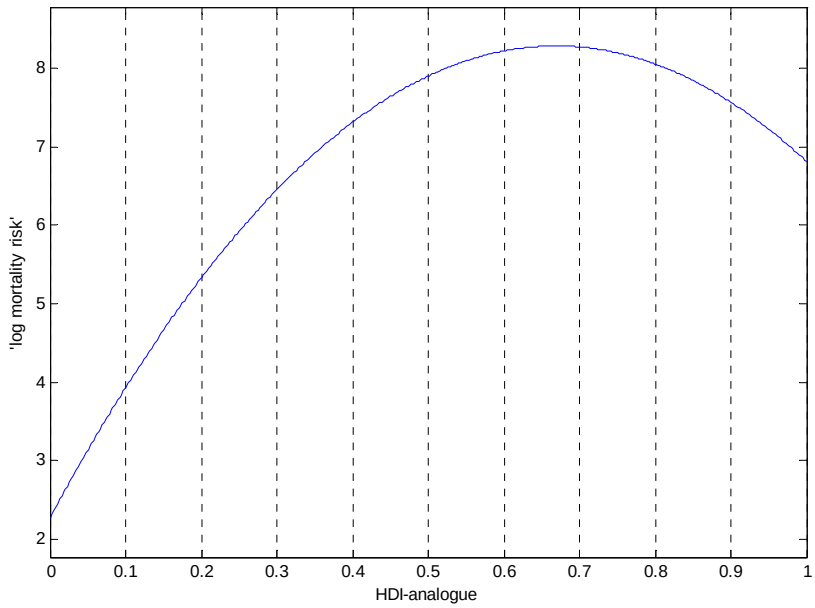
[c] For all models, the LHS of the regression equation is the LHS of equation (1) in the main paper; i.e. log mortality risk.

Model 2 suggests mortality risk monotonically increases as $H_{i,j}$ increases (i.e. as countries develop); this does not seem plausible. Model 3 supports a quadratic fit; however, $H_{i,j}$ is again of relatively low importance, and, as for model 1, the form of the fit is inconsistent with theory (i.e. location of parabola peak).

Model 4 used $H_{i,j}$ in unlogged form. This gave a good fit and the relative importance of $H_{i,j}$ was considerably higher. In models 5 and 6 we included either the linear or quadratic terms for the unlogged form of $H_{i,j}$ but found the importance of $H_{i,j}$ was greatly diminished in both. Thus we retained both the linear and quadratic terms in Model 7.

In terms of the theoretical fit of model 7, the sign of the quadratic term is negative as expected. Further, the form of the model in relation to the location of the peak of the parabola is as expected. Figure S6.1 shows a plot of the Model 7 as a function of the HDI-analogue.

Figure S6.1: Plot of 'log mortality risk' (y-axis) as a function of the HDI-analogue (x-axis) holding the events and population variables constant, using Model 7.



The plot suggests the model matches the theory. 'Log mortality' risk initially increases with the HDI-analogue, peaking when it is between 0.6 and 0.7, and then begins to decline. We note that the location of the peak corresponds to the peak in Patt et al;s model (2010, see Figure 1 B), which peaks when HDI ≈ 0.65.

Further, also corresponding to Patt et al, the lowest mortality risk is in the least developed countries rather than in the most developed. This may appear counterintuitive. However, the curve shows *risk*, not actual mortality, and how that risk is translated into absolute mortality depends on how many people are exposed and *under what conditions*.

Our interpretation of the curve is that in the 'most developed' countries (i.e. HDI approaching 1), a combination of sea- and land-based strategies of adaptation reduce the number of people actually exposed to flooding, but this means when people are exposed, it is likely to be associated with a severe event (as protection prevents exposure to milder events). In other words, if you are exposed, the risk of death would be expected to be reasonably high.

In contrast, in 'less developed' countries, limited protection means people are likely to be exposed to events of all intensities, meaning exposure is likely to be high, but the average risk of mortality among the exposed is lower (as much exposure is due to milder events).

Thus we chose Model 7 as the mortality risk equation (equation 1).

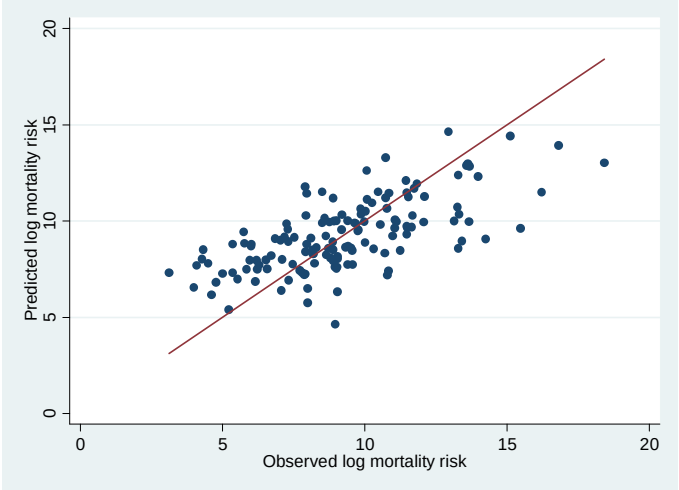
Patt AG, Tadross M, Nussbaumer P, Asante K, Metzger M, Rafael J, Goujon A, Brundrit G (2010) Estimating least-developed countries' vulnerability to climate-related extreme events over the next 50 years. Proceedings of the National Academy of Sciences 107 (4):1333-1337. doi:10.1073/pnas.0910253107

Appendix S7: Model Goodness-of-Fit

As described in the paper, no independent data were available for model validation. Consequently, we compared 'observed' (i.e. based on the data used to fit equation (1)) and predicted (i.e. as estimated by equation (1)) log mortality risk and mortality. Although this method provides only a weak from of validation (as the data used for validation were also used in the model fitting) it nevertheless provides estimates of the expected errors in mortality risk and mortality estimates.

Figure S7.1 shows the correlation between 'observed' log mortality risk (x-axis) and fitted log mortality risk (y-axis). The correlation co-efficient is 0.67, which suggests the model may predict log mortality risk at baseline reasonably well. This does not, however, indicate whether or not it will make reasonable predictions of future mortality.

Figure S7.1: Scatter plot of observed log mortality risk (x-axis) and predicted mortality risk (y-axis)



For mortality, the correlation coefficient for 'observed' mortality versus predicted mortality was 0.08, suggesting the model is an unreliable predictor of even the mortality data used to calibrate the model. Table S7.1 shows observed and predicted mortality and quantifies the error in the prediction for the countries with the five highest average annual mortalities at baseline. Clearly, mortality predictions are very inaccurate, both relatively and absolutely.

Table S7.1 Observed and predicted average annual mortality for the countries with the five highest average annual mortalities at baseline, and associated errors.

Country	'Observed' mortality	Predicted mortality	Error ratio, as a percent[1]	Absolute error[2]
Bangladesh	17,295	561	3	-16,734
Myanmar	2,906	77	3	-2,829
India	1,411	317	22	-1,094
Philippines	758	1,497	197	739
Vietnam	279	8,720	3125	8,441

[1] 'Error ratio' is calculated as: $\frac{predicted\ mortality}{observed\ mortality} \times 100$. For example, an 'error ratio' of 3% indicates that predicted mortality is 3% of observed mortality; 100% indicates perfect prediction.

[2] 'Absolute error' is calculated as: $predicted\ mortality - observed\ mortality$. Negative errors indicate underestimates and positive errors indicate overestimates.

The poor prediction of mortality compared to the prediction of log risk mortality is however not entirely unexpected. This is because the model was fitted in the logarithmic mortality space rather than in the mortality space. Without loss of generality, and to ease algebraic processing, we will use a simple conceptual logarithmic model to illustrate the relationship between the error in the log mortality model and the resulting error in the mortality model.

Denote by ε the error in the fitted log mortality model:

$$\varepsilon = \hat{z} - z \tag{S7.1}$$

where z and $\hat{z}$ are respectively the 'observed' and estimated $\log_e$ number killed. In general, this error, which can be positive (overestimate) or negative (underestimate), is a function of z. The true (v) and the estimated $(\hat{v})$ number killed are given respectively by:

$$v = exp(z) \tag{S7.2}$$

and

$$\hat{v} = exp(\hat{z}) \tag{S7.3}$$

From equations (S7.2) and (S7.3),

$$\frac{\hat{v}}{v} = \frac{exp(\hat{z})}{exp(z)} = exp(\hat{z} - z) \tag{S7.4}$$

Using equation (S7.4), the error in estimating the actual number killed resulting from the error in the model estimating $\log_e$ number killed is:

$$\delta = \hat{v} - v = v\ exp(\varepsilon) - v = v(exp(\varepsilon) - 1) \tag{S7.5}$$

We deduce from equation (S6.5) that for large positive errors ε, $\delta \approx v\ exp(\varepsilon)$, i.e. the errors are exponentiated and are unbounded. For large negative errors ε, $\delta \approx -v$, i.e. the errors are bounded by -100%.

Table S7.2 shows predicted and observed average annual mortality and the associated errors when aggregated to the regional level. Again, it is clear that mortality predictions are naturally also very poor.

Table S7.2: Observed and predicted average annual mortality and associated errors, aggregated to regional level.

Region	'Observed' mortality	Predicted mortality	Error ratio, as a percent[1]	Absolute error[2]
Asia Pacific, High Income	52.1	14.4	28	-37.7
Asia, Central	0.0	4.1	-	4.1
Asia, East	270.9	2,001.5	739	1,730.6
Asia, South	18,731.8	879.9	5	-17,851.9
Asia, Southeast	4,052.1	10,441.5	258	6,389.3
Australasia	2.4	25.3	1,038	22.9
Caribbean	196.8	353.4	180	156.6
Europe, Central	0.7	12.5	1,855	11.8
Europe, Eastern	2.3	104.0	4,508	101.7
Europe, Western	23.4	143.0	611	119.6
Latin America, Andean	10.1	14.9	148	4.8
Latin America, Central	375.2	72.3	19	-302.9
Latin America, Southern	0.0	21.4	-	21.4
Latin America, Tropical	0.3	8.7	3,427	8.4
North Africa / Middle East	50.2	60.8	121	10.6
North America, High Income	3.7	161.4	4,410	157.7
Oceania	16.7	36.6	220	19.9
Sub-Saharan Africa, Central	0.7	2.1	315	1.4
Sub-Saharan Africa, East	64.4	232.5	361	168.1
Sub-Saharan Africa, Southern	0.0	0.9	-	0.9
Sub-Saharan Africa, West	0.0	48.2	-	48.2
Global total	**23,853.8**	**14,639.2**		

[1] 'Error ratio' is calculated as: $\frac{predicted\ mortality}{observed\ mortality} \times 100$. For example, an 'error ratio' of 28% indicates that predicted mortality is 28% of observed mortality; 100% indicates perfect prediction.

[2] 'Absolute error' is calculated as: $predicted\ mortality - observed\ mortality$. Negative errors indicate underestimates and positive errors indicate overestimates.