SCIENTIFIC DATA

Data Descriptor: Gridded birth and pregnancy datasets for Africa, Latin America and the Caribbean

W.H.M. James[1], N. Tejedor-Garavito[1,2,3], S.E. Hanspal[1], A. Campbell-Sutton[3], G.M. Hornby[1,3], C. Pezzulo[1,2], K. Nilsen[1], A. Sorichetta[1,2], C.W. Ruktanonchai[1,2], A. Carioli[1], D. Kerr[1], Z. Matthews[4] & A.J. Tatem[1,2]

Understanding the fine scale spatial distribution of births and pregnancies is crucial for informing planning decisions related to public health. This is especially important in lower income countries where infectious disease is a major concern for pregnant women and new-borns, as highlighted by the recent Zika virus epidemic. Despite this, the spatial detail of basic data on the numbers and distribution of births and pregnancies is often of a coarse resolution and difficult to obtain, with no co-ordination between countries and organisations to create one consistent set of subnational estimates. To begin to address this issue, under the framework of the WorldPop program, an open access archive of high resolution gridded birth and pregnancy distribution datasets for all African, Latin America and Caribbean countries has been created. Datasets were produced using the most recent and finest level census and official population estimate data available and are at a resolution of 30 arc seconds (approximately 1 km at the equator). All products are available through WorldPop.

[1]WorldPop, Geography and Environment, University of Southampton, Highfield Campus, Southampton SO17 1BJ, UK. [2]Flowminder Foundation, Roslagsgatan 17, Stockholm SE-11355, Sweden. [3]GeoData, University of Southampton, Highfield Campus, Southampton SO17 1BJ, UK. [4]Division of Social Statistics and Demography & Centre for Global Health, Population, Poverty and Policy, Faculty of Social and Human Sciences, University of Southampton, Southampton SO17 1BJ, UK. Correspondence and requests for materials should be addressed to N.T.-G. (email ntg@geodata.soton.ac.uk).

Design Type	data integration objective ● source-based data analysis objective
Measurement Type(s)	population measurement
Technology Type(s)	computational modeling technique
Factor Type(s)	Data Source ● geographic location
Sample Characteristic(s)	Homo sapiens ● Anguilla ● Antigua and Barbuda ● Argentina ● Aruba ● The Bahamas ● Barbados ● Belize ● Bolivia ● Bonaire, Saint Eustatius and Saba ● Brazil ● British Virgin Islands ● Cayman Islands ● Chile ● Colombia ● Costa Rica ● Cuba ● Country of Curaçao ● Dominica ● Dominican Republic ● Ecuador ● El Salvador ● Falkland Islands ● Guyane ● Grenada ● Department of Guadeloupe ● Guatemala ● Guyana ● Haiti ● Honduras ● Jamaica ● Department of Martinique ● Mexico ● Montserrat ● Nicaragua ● Panama ● Paraguay ● Peru ● Puerto Rico ● Saint-Barthelemy ● Saint Kitts and Nevis ● Saint Lucia ● Saint Martin ● Saint Vincent and the Grenadines ● Sint Maarten ● Suriname ● Trinidad and Tobago ● Turks and Caicos Islands ● Uruguay ● Venezuela ● United States Virgin Islands ● Algeria ● Angola ● Benin ● Botswana ● Burkina Faso ● Burundi ● Cameroon ● Cape Verde ● Central African Republic ● Chad ● Comoros ● Republic of Congo ● Democratic Republic of the Congo ● Cote d'Ivoire ● Djibouti ● Egypt ● Equatorial Guinea ● Eritrea ● Ethiopia ● Gabon ● Gambia ● Ghana ● Guinea ● Guinea-Bissau ● Kenya ● Lesotho ● Liberia ● Libya ● Madagascar ● Malawi ● Mali ● Mauritania ● Mauritius ● Mayotte ● Morocco ● Mozambique ● Namibia ● Niger ● Nigeria ● Reunion Island ● Rwanda ● Saint Helena ● Sao Tome and Principe ● Senegal ● Seychelles ● Sierra Leone ● Somalia ● Republic of South Africa ● South Sudan ● Sudan ● Swaziland ● Tanzania ● Togo ● Tunisia ● Uganda ● Western Sahara ● Zambia ● Zimbabwe

Background & Summary

Accurate and detailed information on the spatial distribution and numbers of births and pregnancies is crucial for informing planning decisions related to public health[1]. The survival and health of women and their new-born babies in low income countries is a key priority, with the reduction of maternal and neonatal mortality central for meeting a number of the United Nations Sustainable Development Goals (specifically goals 3.1 and 3.2)[2]. Whilst progress has been made, there were still 303,000 maternal deaths in 2015 (ref. 3) and children in lower income countries are 14 times more likely to die during their first 28 days of life compared to their higher income counterparts. Despite this, the spatial detail of basic data on the numbers and distribution of births and pregnancies is often of a coarse resolution and difficult to obtain[4], with no co-ordination between countries and organisations to create one consistent set of subnational estimates for planning.

Whilst there are clear inequalities of maternal and neonatal healthcare between nations[5], there are also large disparities within individual countries, with growing recognition that national levels and trends could be masking important sub-national variations[6]. For example, a study in Indonesia found that under-5 mortality was nearly four times higher in the poorest fifth of the population than in the richest fifth[7], and gaps like these are more likely to occur at the sub-national level[7–9]. Although progress has been made in reducing such inequalities, there is still substantial work to be done. As such, understanding sub-national variation and inequity in health status, wealth and access to resources is increasingly being recognised as central to meeting developmental goals[10]. To understand and tackle inequalities related to maternal and neonatal health, the first step is to have a detailed knowledge of the distribution of births and pregnancies, which is known to vary substantially due to population age and sex distribution and age specific fertility rates (ASFR)[4]. These are also valuable data for subnational planning and estimation, and calculation of subnational indicators that rely on births or pregnancies as a denominator. When considering maternal and neonatal health in lower income countries, infectious disease is a major concern as pregnant women and new-borns are particularly at risk from many diseases, such as malaria[11] and HIV[12]. This issue has recently been highlighted by the Zika virus outbreak in Latin America, further intensifying the need for detailed information on the number and distribution of births and pregnancies. Currently there is a clear lack of data for such analysis, with complete and continuous datasets of numbers of births only available at the national level (e.g., United Nations Population Division[13]). Whilst sub-national datasets are readily available for some countries, their spatial detail is often coarse with differences in the recorded metrics, sampling framework and data formats meaning that it is extremely difficult to assess burden within and across multiple nations.

This study aims to overcome the data gap identified above by producing continental scale, gridded datasets of numbers of births and pregnancies with a spatial resolution of 30 arc seconds (approximately 1 km at the equator). Advances in computational power and spatial econometric techniques, as well as the increasing availability of geo-located data, have increased the ability to produce these fine spatial resolution datasets. As such, in the framework of the WorldPop project (www.worldpop.org), and extending the approaches described by Tatem et al.[4], an open access archive of gridded birth and pregnancy distribution datasets for all African, Latin America and Caribbean (LAC) countries has been created. This process used the most recent and finest level census, census microdata, household survey data and official population estimate data available to the authors at the time of writing, alongside a range of geospatial datasets.

Methods

Gridded estimates of live births were produced for 50 Latin American and Caribbean and 58 African countries at a spatial resolution of 30 arc seconds. This was achieved by combining the latest datasets on population distribution, population age and sex structure and fertility rates in a GIS environment. Estimates of pregnancies were additionally generated using national-level estimates for stillbirths, miscarriages and abortions from the Guttmacher Institute[14]. The workflow develops the methods presented by Tatem et al.[4], using a variety of data sources to construct continent wide datasets. The process is fully automated by a Python Script, allowing the rapid processing of multiple countries and alignment to a standard grid for the production of seamless continental scale datasets. The workflow is shown in Fig. 1 and described in detail below. Maps of the data sources and date for each country and whether urban and rural ASFR estimates were available can be found in the Supplementary Figures 1 and 2 respectively.

The basis for estimation: population distributions

The population distribution forms a major component of the births and pregnancy estimation process. The WorldPop project has recently completed construction of gridded population distribution datasets for all low- and middle-income countries at a resolution of 30 arc seconds. Full details are provided on the WorldPop website (www.worldpop.org.uk) along with links describing the methods in detail[15–17]. This study uses the relevant regions of Africa (Data Citation 1) and Latin America and the Caribbean (Data Citation 2), whose total population is adjusted to match the most recent United Nations Population Division (UNPD)[18] 2015 estimates available when the population distribution datasets were produced. Figure 2a shows the gridded population distribution dataset for Bolivia as an example. To ensure data consistency, a WorldPop standard grid was used in processing; this is a gridded dataset providing ISO country codes at a resolution of 30 arc second (Data citation 3).

Calculating the proportion of women of reproductive age

Sub-national information on age and sex structure was collected, specifically women of childbearing age grouped in seven 5-years age groups (i.e., 15–19, 20–24, 25–29, 30–34, 35–39, 40–44, 45–49), as defined by the UNPD[18]. Datasets for the majority of Africa were provided by Pezzulo et al.[19] whilst datasets for the remaining African, Latin America and the Caribbean countries were assembled from a variety of sources, following the protocols defined by Pezzulo et al.[19]. Table 1 (available online only) shows the source, spatial detail (i.e., administrative unit level) and reference year used for the countries processed in this study.

With the raw data recorded and documented according to different protocols determined by national governments, the project was presented with a wide range of table data formats and schemas. Data restructuring was achieved using scripting (R 3.3.1, Python 2.7) and table processing software (Microsoft Excel 2013). The resultant standardised tables contained fields corresponding to the proportionate values of people (both sexes) in each 5-year age group, and the overall proportion of males and females in each region. Table 2 shows an example of the standardised tables, for regions of Peru.

Age and sex structure information was matched to vector geographical boundaries from the Global Administrative Areas (GADM) database[20], with the exceptions of Chile and Colombia where boundaries from the National Statistics Office were used. The extent of these boundaries was standardised to those defined by the WorldPop gridded ISO country code dataset using the Clip and Nibble tools in ArcGIS 10.3, executed as part of the Python[21] script. Figure 2b shows the distribution of females between the ages of 20 and 24 for Bolivia. Similar distributions were created for all other 5-year age groupings in the 15–49-year range.

Estimating fertility rates

Data on fertility was collected on a country-by-country basis to provide the most up to date and spatially detailed information. Data sources were chosen using a hierarchical approach as shown in Fig. 1, prioritising sources which included information on age specific fertility and those of the highest spatial detail. The type of fertility data used for each country is shown in Table 3 (available online only) and described in detail below. As with the age and sex structure datasets, restructuring and table processing was carried out using a variety of scripting and software packages (Python 2.7, R 3.3.1, Microsoft Excel 2013) to produce a common format and schema for each data type, as described in detail below.
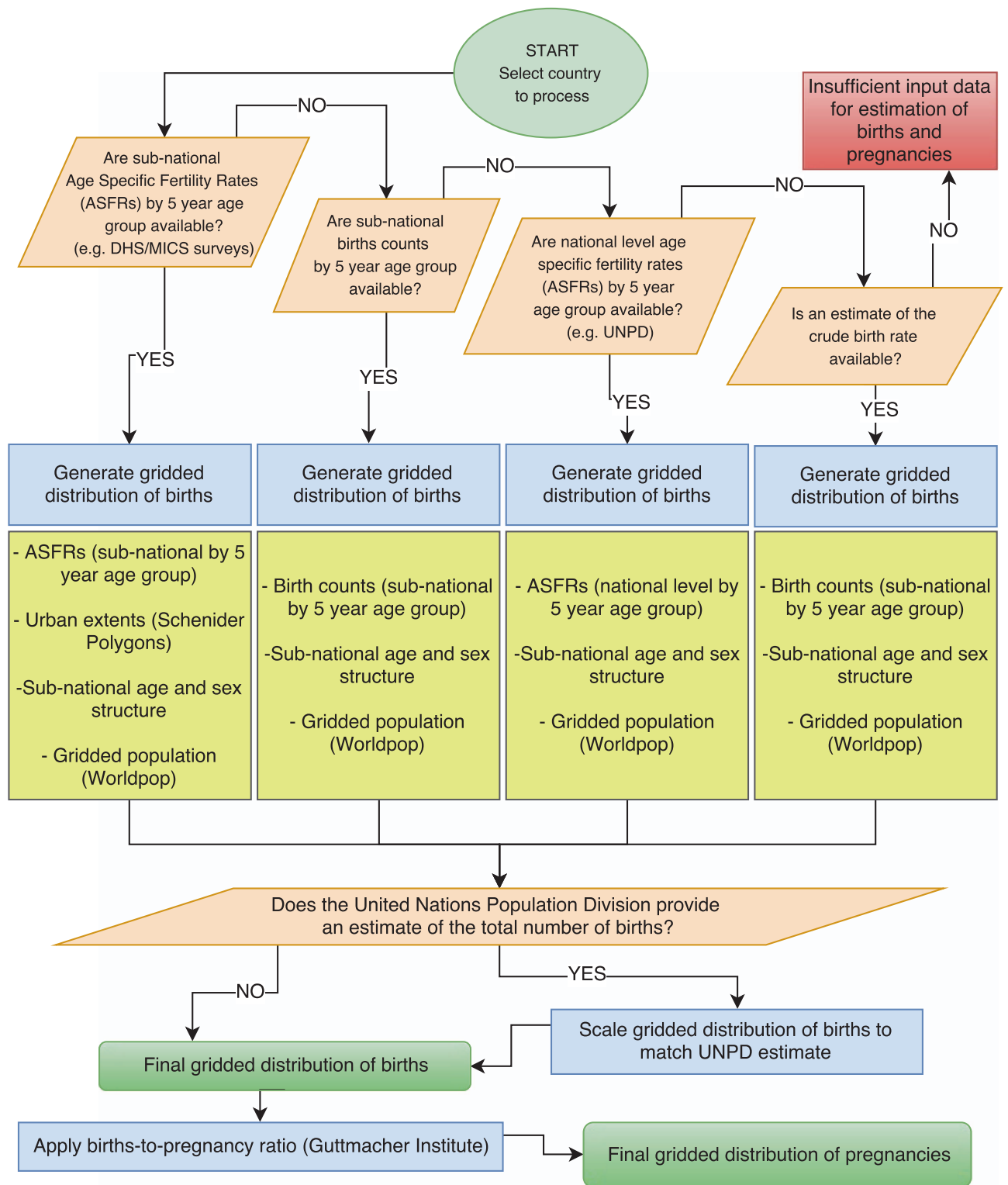
Figure 1. Schematic overview of the workflow adopted to generate gridded subnational births and pregnancy datasets. ASFR = Age Specific Fertility Rate, DHS = Demographic and Health Survey, MICS = Multiple Indicator Cluster Survey, UNPD = United Nations Population Division.

To provide the finest spatial detail of the distribution of fertility across each country, ASFRs were estimated by 5-year age groups, disaggregated sub-nationally according to the relevant Demographic and Health Surveys (DHS), Multiple Indicator Cluster Surveys (MICS) or National Statistics Agency survey regions and by urban vs rural if available. Table 3 (available online only) indicates which countries had
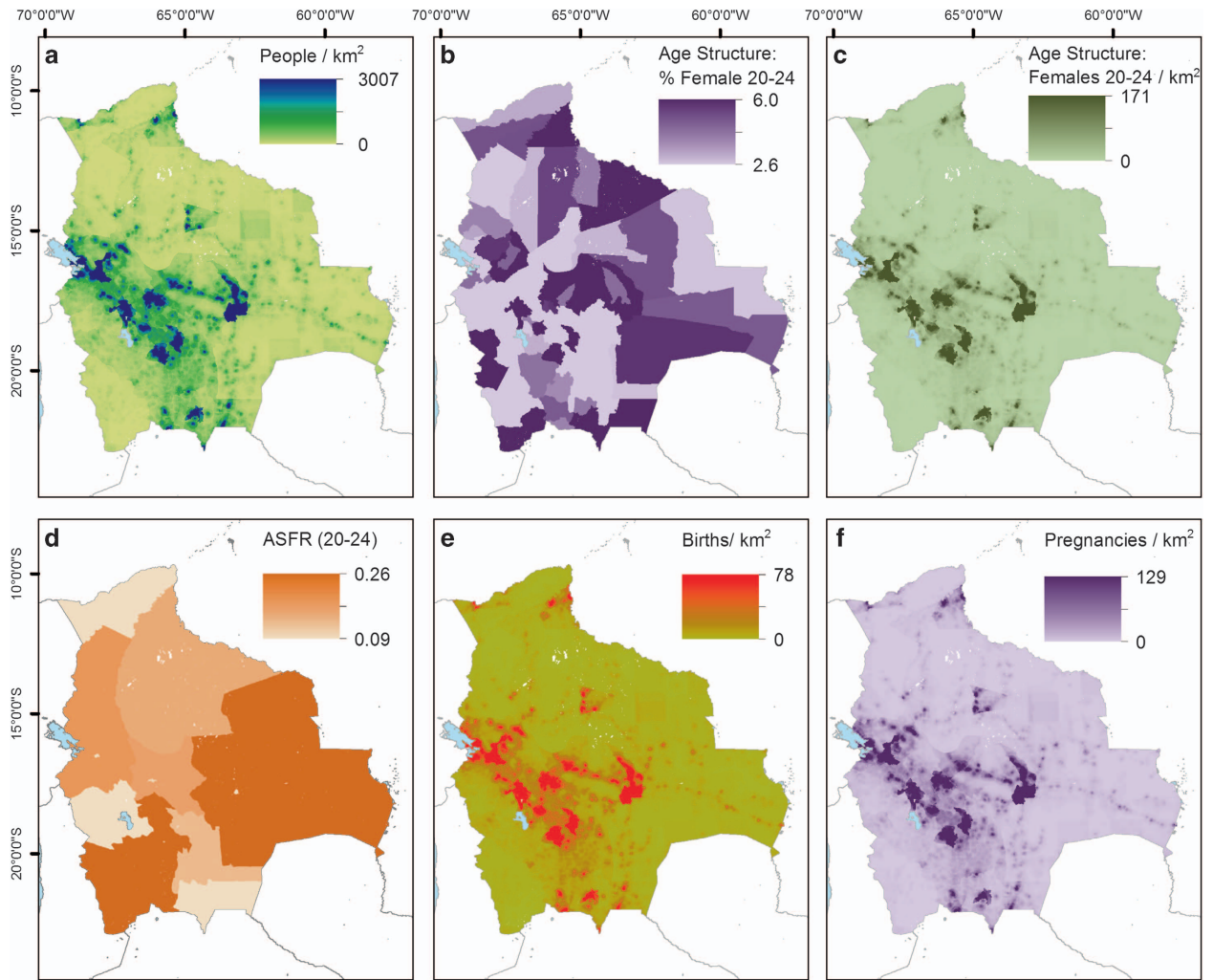
Figure 2. Examples of input, intermediate and output datasets for Bolivia. (**a**) Gridded population input (WorldPop), (**b**) Age and sex structure disaggregated by administrative unit, (**c**) Derived gridded age and sex structure, (**d**) ASFRs disaggregated by region and urban vs rural, (**e**) Gridded births output, (**f**) Gridded pregnancies output.

the required data for estimation. ASFRs for each 5-year age group were commonly derived from DHS or MICS, with the exceptions of Aruba, Chile and Eritrea where the relevant datasets were available from the corresponding National Statistics Office.

For DHS and MICS surveys, ASFRs were estimated using a Stata program developed by Pullum[50], as discussed in Tatem et al.[4]. The program calculated the basic demographic indicator by deriving ASFRs for each of the seven 5-year age groups covering the reproductive life span from 15–49 years based on dividing the number of births to women in each age group, during a retrospective 3-years reference period, by the number of women-years during the same period. Data restructuring and table processing was carried out using R 3.3.1 and Microsoft Excel 2013 to produce a common format, as that shown in Table 4.

Datasets representing the boundaries of subnational regions (Table 3) (available online only) were assembled and the relevant ASFRs matched to them. If the ASFR data was available for urban and rural areas within the subnational regions, the MODIS 500 m Global Urban Extent dataset[51] was used to distinguish urban and rural areas and allocate the constant value within them. Figure 2d shows an example ASFR dataset for Bolivia, showing the ASFRs for one reproductive age range: the 20–24 age group by sub-region. Similar datasets were constructed for all other 5-year age groups within the 15 to 49 range.

For countries where ASFRs disaggregated sub-nationally and by urban/rural were not available, information on the spatial variation of age structured fertility was sought from vital registration systems, census records and other national sources. This was routinely in the form of births registered per

Region	t_0_4	t_5_9	t_10_14	t_15_19	......	t_60_64	t_65_plus	prop_m_t	prop_f_t
PER_Callao_Callao	0.09	0.08	0.09	0.09		0.03	0.06	0.49	0.51
PER_Cusco_Acomayo	0.12	0.14	0.14	0.08		0.03	0.08	0.49	0.51
PER_Cusco_Anta	0.10	0.12	0.14	0.10		0.03	0.08	0.50	0.50
PER_Cusco_Calca	0.11	0.12	0.14	0.10		0.02	0.06	0.50	0.50
PER_Cusco_Canas	0.12	0.14	0.14	0.09		0.03	0.08	0.50	0.50
PER_Cusco_Canchis	0.10	0.12	0.13	0.10		0.03	0.07	0.49	0.51

Table 2. Example of standardised table containing proportionate values of people by age group and sex for Peru. (Note table is for illustrative purposes only and therefore does not display all age groups).

ISO	Region	Rural/Urban	ASFR 15_19	ASFR 20_24	ASFR 25_29	ASFR 30_34	ASFR 35_39	ASFR 40_44	ASFR 45_49	Year
HTI	WEST	URBAN	0.03	0.15	0.17	0.13	0.08	0.10	0.00	2012
HTI	NORTH	URBAN	0.05	0.10	0.12	0.14	0.07	0.03	0.01	2012
HTI	CENTRAL	URBAN	0.05	0.10	0.22	0.20	0.09	0.06	0.00	2012
HTI	SOUTH	URBAN	0.03	0.10	0.14	0.17	0.07	0.07	0.02	2012
HTI	NIPPES	URBAN	0.01	0.09	0.18	0.13	0.09	0.04	0.00	2012
HTI	WEST	RURAL	0.07	0.18	0.20	0.19	0.18	0.07	0.01	2012
HTI	NORTH	RURAL	0.08	0.17	0.21	0.16	0.13	0.07	0.02	2012

Table 4. Example of standardised country table containing the ASFR values for regions in Haiti with separate entries for urban and rural areas.

Region	ISO	b_15_19	b_20_24	b_25_29	b_30_34	b_35_39	b_40_44	b_45_49	year
1	VEN	111	101	106	58	41	12	4	2012
2	VEN	113	135	109	80	55	24	7	2012
3	VEN	879	1026	788	422	194	58	13	2012
4	VEN	111	131	110	87	53	22	9	2012
5	VEN	150	152	129	106	62	30	6	2012
6	VEN	15	16	15	12	7	4	1	2012
7	VEN	20	33	21	17	9	6	4	2012
8	VEN	347	502	444	299	141	34	0	2012
9	VEN	197	193	137	78	30	6	0	2012

Table 5. Example of standardised country table containing the number of registered births per age group for regions in Venezuela.

administrative unit per 5-year age grouping. Table 3 (available online only) shows for which countries this type of data (registered births per age group) was used, with Table 5 showing an example of the standardised table format for Venezuela.

Sub-national ASFRs were calculated by dividing the number of births in each age group (e.g., Table 5) by the number of females in the corresponding group. The latter was derived from the WorldPop population distribution[15] and age-sex distributions produced following the methodology of Pezzulo et al.[19] and outlined in Table 1 (available online only).

The UNPD provides national estimates of AFSRs by 5-year age grouping for the majority of countries[18,24]. These datasets were used where subnational information was not available. As with all other datasets, the country boundaries defined by WorldPop[15] were used to define the geographical extent.
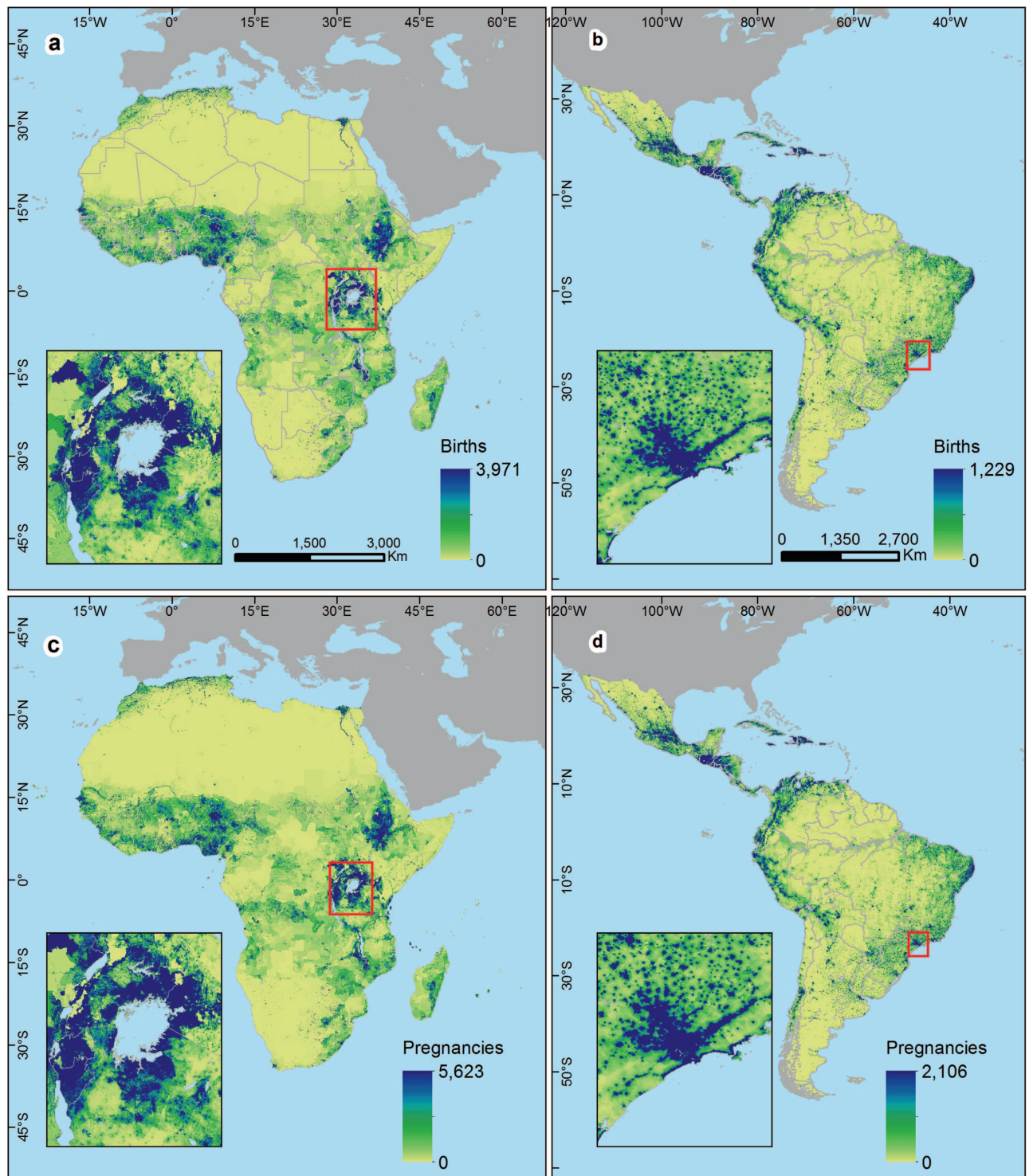
Figure 3. Estimated births and pregnancies per grid cell for Africa, Latin America and the Caribbean in 2015. The grid cell resolution is 30 arc seconds (approximately 1 km at the equator) and co-ordinates refer to GCS WGS 1984.

For 9 countries, there was no information available on age specific fertility, either sub-nationally or nationally. In these cases, crude birth rates were obtained from a variety of official sources (Table 3 (available online only)) which were subsequently matched to the appropriate GIS country boundaries supplied by WorldPop[15].

Name	Description	Resolution	Files Format	University of Southampton DOI
Births in Latin America and the Caribbean	Estimated live births per grid cell for 2015 for LAC for 50 countries	3 arc seconds	GeoTIFF	10.5258/SOTON/WP00529
Births in Africa	Estimated live births per grid cell for 2015 for Africa for 58 countries	3 arc seconds	GeoTIFF	10.5258/SOTON/WP00528
Pregnancies in Latin America and the Caribbean	Estimated pregnancies per grid cell for 2015 for LAC for 50 countries	3 arc seconds	GeoTIFF	10.5258/SOTON/WP00527
Pregnancies in Africa	Estimated pregnancies per grid cell for 2015 for Africa for 58 countries	3 arc seconds	GeoTIFF	10.5258/SOTON/WP00526

Table 6. **Name and description of datasets available for Africa and Latin America and Caribbean countries.**
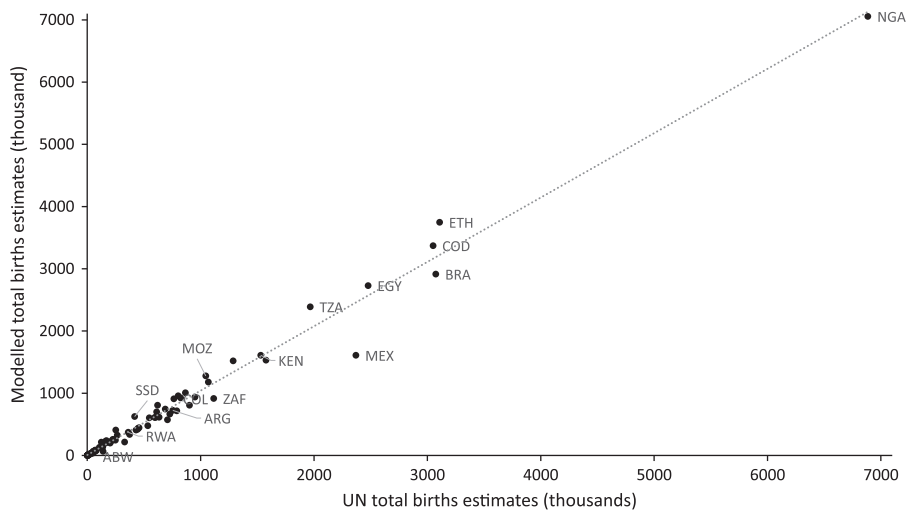
Figure 4. **Total number of births per country estimated by this study plotted against the corresponding UNPD estimate.**

Estimating the number of births from fertility, population and age structure

For countries where measures of age specific fertility were available, the distribution of live births was estimated by multiplying the number of females in each age group (e.g., Fig. 2c) by the corresponding ASFR gridded dataset (e.g., Fig. 2d) or value (in the case of national ASFR estimates). The resultant seven age specific gridded datasets were summed to generate an estimate of total births. For countries where fertility was expressed simply as a crude birth rate (Table 3 (available online only)), the births distribution was calculated by multiplying the crude birth rate by the initial 30 arc second UNPD adjusted WorldPop population grid for 2015 (ref. 15).

Finally, for each country, the distribution was scaled to match the UNPD estimate of the total number of births[18]. For countries where the UNPD does not provide an estimate of the total number of births, the initial total was used. An example of the final distributed births gridded dataset for Bolivia is shown in Fig. 2e, with results for Africa, Latin America and the Caribbean shown in Fig. 3 and Supplementary Table 1.

Estimating the distribution of pregnancies

The Guttmacher institute has published country specific estimates of the number of stillbirths, miscarriages and abortions at the national level[14]. These estimates for 2014 were integrated with UNPD national estimates on numbers of live births[18] to construct a ratio between numbers of births and pregnancies. This ratio was applied to the live births distribution to generate an estimate of the distribution of pregnancies. For countries not covered by the Guttmacher dataset, the nearest suitable geographical country value was used. An example of the final distributed pregnancies gridded dataset for Bolivia is shown in Fig. 2f whilst results for the entire Africa, Latin America and the Caribbean are shown in Fig. 3.

Code availability

The Python code developed for production of the births and pregnancies datasets is publicly and freely available through Figshare[52]. The code consists of a Python programming language script (version 2.7; www.python.org) and relies on the ArcGIS 10.4.1 ArcPy site package for performing GIS specific spatial operations. The script is internally documented to both explain its purpose (including a description of the GIS-specific spatial operations it performs) and, when required, guiding the user through its customisation.

Data Records

The high-resolution births and pregnancies datasets described in this article referring to the 108 countries listed in Table 3 (available online only) are publicly and freely available through the WorldPop Repository (http://www.worldpop.org.uk/data/). A collection of these datasets has been compiled for the births for LAC (Data Citation 4) and Africa (Data Citation 5) and pregnancies for LAC (Data Citation 6) and Africa (Data Citation 7), as described in Table 6.

Technical Validation

All data collected, assembled and used were (i) already validated by the corresponding data collector, owner and/or distributor, and (ii) further checked, in the framework of this project. The gridded 5-year age and sex count datasets constructed for Latin America and the Caribbean (e.g., Fig. 2c) were verified following the protocol outlined in Pezzulo et al.[19], who compiled and assessed similar datasets for Africa and Asia. Briefly, this comprised of summing all the layers into a single dataset (representing the total numbers of people for all age and sex groups at the grid cell level) and then subtracting it from the corresponding WorldPop continental gridded population count dataset to make sure that the country totals matched the UNPD estimates for the year in question. All fertility rates used in this study were checked, on a county-by-country basis, to make sure they were within reasonable ranges. Additionally, for countries where additional sources of fertility data were available, estimates were produced using all available sources to compare the adjusted total births. These results showed that, although differences may be observed at the grid cell level, the totals at the administrative unit level are very similar. Endeavours were made to assemble the most recent, reliable and spatially detailed data at the time of writing. However, additional input from readers who may have knowledge or access to more recent and/ or better datasets are welcome for improving future iterations of the outputs.

The accuracy and quality of fertility estimates from survey data such as those provided by the DHS, have been assessed in several reports, by testing the quality of the birth history data in a large number of countries. These checks were mainly aimed to identify potential omission and displacement of births, potential displacement of births, or misreporting of date of birth[53,54]. Overall, although a number of issues were identified for some countries, these studies found that most estimates were either good or of acceptable quality. Furthermore, outcomes from Pullum and Becker[53] show that in general the latest DHS surveys are less prone to issues like incomplete birthdates, omissions and displacement of births and deaths. Similarly, a more recent report from Pullum and Staveteig[55], exploring the quality and consistency of age and date reports in DHS surveys, demonstrates that DHS data is constantly evaluated to improve its quality.

Modelled estimates of total number of births per country prior to adjustments (to match UNPD estimate) were also plotted against the UNPD estimates[18] to assess the size of differences obtained through using subnational data sources. Figure 4 shows the correlation for the 95 countries for which UNPD provides an estimate, with a corresponding R^2 value of 0.982. Analysis was not possible for the 9 countries for which the UNPD does not provide an estimate, although these make up a very small proportion (0.01%) of the total births across the whole study area.

Usage Notes

The datasets presented here can be used both to (i) support applications measuring sub-national metrics of maternal and new-born health and (ii) to inform planning decisions. However, considering that they represent modelling outputs generated using ancillary covariates for producing the underlying WorldPop population distribution datasets, to avoid circularity, they should not be used to make predictions or explore relationships about any of those ancillary datasets[56]. Thus, before using the births and pregnancies datasets in correlation analyses against factors which are included in the construction of the population distribution datasets (e.g., correlating birth distribution with land-cover), ideally the population modelling process should be re-run using the WorldPop-RF code[57] with the applicable covariates removed.

References

1. Ruktanonchai, C. et al. Equality in maternal and newborn health: modelling geographic disparities in utilisation of care in five East African countries. PLoS ONE 11, e0164519 (2016).
2. United Nations. Transforming Our World: The 2030 Agenda for Sustainable Development, A/RES/70/1 (United Nations, 2016).
3. World Health Organisation. Trends in maternal mortality: 1990 to 2015, WHO/RHR/15.23 (WHO, 2016).
4. Tatem, A. et al. Mapping for maternal and newborn health: the distributions of women of childbearing age, pregnancies and births. International Journal of Health Geographics 13, 2 (2014).

5. Graham, W. *et al.* The geography of maternal death. in *Maternal and Infant Deaths: Chasing Millennium Development Goals 4 and 5 Royal College of Obstetricians and Gynaecologists Study Group* (eds Neilson J., Norman J. & Kehoe S.) 3–28 (Cambridge University Press, 2010).
6. Victora, C. *et al.* The contribution of poor and rural populations to national trends in reproductive, maternal, newborn, and child health coverage: analyses of cross-sectional. *The Lancet Global Health* **5,** e402–e407 (2017).
7. Victora, C. *et al.* Applying an equity lens to child health and mortality: more of the same is not enough. *The Lancet* **362,** 233–241 (2003).
8. United Nations. *The state of the world's midwifery: analysis of the sexual, reproductive, maternal, newborn and adolescent health workforce in East & Southern Africa.* (United Nations Population Fund (UNFPA) East and Southern Africa Regional Office (ESARO), 2017).
9. Armstrong, C. E. *et al.* Subnational variation for care at birth in Tanzania: is this explained by place, people, money or drugs? *BMC Public Health* **16,** 795 (2016).
10. United Nations and Social Council. *Progress towards the Sustainable Development Goals*, Report of the Secretary-General (United Nations Economic and Social Council, 2016).
11. Dellicour, S., Tatem, A. J., Guerra, C. A., Snow, R. W. & ter Kuile, F. O. Quantifying the Number of Pregnancies at Risk of Malaria in 2007: A Demographic Study. *PLoS Medicine* **7,** e1000221 (2010).
12. Joint United Nations Programme on HIV/AIDS (UNAIDS). *Global report: UNAIDS report on the global AIDS epidemic 2010* (UNAIDS, 2010).
13. United Nations. *World Population Prospects. Key findings and advance tables. 2017 revision*, Working Paper No. ESA/P/WP/248 (Department of Economic and Social Affairs Population Division, 2017).
14. Sedgh, G., Singh, S. & Hussain, R. Intended and Unintended Pregnancies Worldwide in 2012 and Recent Trends. *Studies in Family Planning* **45,** 301–314 (2014).
15. Lloyd, C., Sorichetta, A. & Tatem, A. High resolution global gridded data for use in population studies. *Sci. Data* **4,** 170001 (2017).
16. Stevens, F., Gaughan, A., Linard, C. & Tatem, A. Disaggregating census data for population mapping using random forests with remotely-sensed and ancillary data. *PLoS ONE* **10.2,** e0107042 (2015).
17. Tatem, A. WorldPop, open data for spatial demography. *Sci. Data* **4,** 170004 (2017).
18. United Nations. *World Population Prospects: The 2017 Revision, DVD edition* (United Nations Department of Economic and Social Affairs Population Division, 2017).
19. Pezzulo, C. *et al.* Sub-national mapping of population pyramids and dependency ratios in Africa and Asia. *Sci. Data* **4,** 170089 (2017).
20. Hijmans, R. *et al.* Global Administrative Areas http://www.gadm.org/ (2017).
21. Rossum, G. V. & Boer, J. D. Interactively testing remote servers using the Python programming language. *CWi Quarterly* **4,** 283–304 (1991).
22. *Central Bureau of Statistics-Aruba. Age Specific Fertility Rates 1990-2000* (Aruba, 2016).
23. ICF International. Demographic and Health Surveys. https://dhsprogram.com/data/available-datasets.cfm (1994-2016).
24. United Nations. *World Fertility Data 2008*, POP/DB/Fert/Rev2008 (United Nations Department of Economic and Social Affairs Population Division, 2008).
25. Ministerio de Salud Republica Argentina. *Estadicias Vitales 2012* (Argentina, 2013).
26. CBS- Statistics Netherlands. *Trends in the Caribbean Netherlands 2016* (The Netherlands, 2016).
27. Department of Statistics of the Ministry of Finance of The Bahamas. *Vital Statistics Report 2013* (The Commonwealth of the Bahamas, 2013).
28. Institut National de la Statistique et des Etudes Economiques France. *Comparateur de territoire, Commune de Saint-Barthélemy*, Report No. 50450 (France, 2013).
29. Instituto Brasileiro de Geografia e Estatística. *Estatísticas do Registro Civil. Tabela 2609* (Brazil, 2016).
30. Statistics Botswana. *Vital Statistics Report 2014* (Botswana, 2014).
31. Instituto Nacional de Estadísticas e Informacion de Salud de Chile. *Estadísticas de Natalidad y Mortalidad, Chile 2013* (Chile, 2013).
32. Instituto Nacional de Estadística y Censos-Costa Rica. *Nacimiemtos* (Costa Rica, 2011).
33. United Nations International Children's Fund (UNICEF). Multiple Indicator Cluster Survey (MICS) http://www.childinfo.org/mics4_surveys.html (2011-2014).
34. Curaçao Central Bureau of Statistics. *Statistical Yearbook 2013* (Curaçao, 2013).
35. Economics and Statistics Office Government of the Cayman Islands. *Population and vital statistics* (Cayman Islands, 2015).
36. Instituto Nacional de Estadísticas-Ecuador. *Estadística de Nacimientos* (Ecuador, 2012).
37. National Statistics Office (NSO) [Eritrea] and Fafo Institute For Applied International Studies. *Eritrea Population and Health Survey 2010* (The State of Eritrea and Fafo Institute for Appled International Studies, 2013).
38. Falkland Islands Government. *Statistical year book 2014* (Falkland Islands, 2014).
39. Government of Jamaica. *Vital statistics* (Jamaica, 2013).
40. The World Bank. *World Development Indicators* (The World Bank, 2014).
41. Instituto Nacional de Estadística y Geografía-Mexico. *Nacimientos registrados* (Mexico, 2015).
42. Instituto Nacional de Estadística y Censos-Panama. *Estadísticas Vitales, Volumen II - Nacimientos Vivos y Defunciones Fetales: Año 2012* (Panama, 2012).
43. Gobierno de Puerto Rico - Departamento de Salud. *Informe anual de estadisticas vitales* (Puerto Rico, 2010).
44. Dirección General de Estadística Encuestas y Censos-Paraguay. *Anexo 3: Nacidos vivos inscriptos por año de nacimiento* (Paraguay, 2016).
45. Statistics Office Corporate Policy and Planning Unit (CPPU)-Saint Helena. *Statistical Yearbook 2012/13* (Saint Helena, 2013).
46. Instituto Nacional de Estadística-Uruguay. *Estadicias Vitales* (Uruguay, 2015).
47. Statistical Office Central Planning Division. *St. Vincent and the Greenadines: Population and Vital Statistics* (St. Vincent and the Greanadines, 2013).
48. Instituto Nacional de Estadística-Venezuela. *Demograficos - Estadísticas Vitales* (Venezuela, 2016).
49. Statistics South Africa. *Recorded Live Births* (South Africa, 2013).
50. Pullum, T. Program to produce ASFRs, TFR, GFR for specific windows of time, with covariates. *Pers. Comm* (2012).
51. Schneider, A., Friedl, M. & Potere, D. A new map of global urban extent from MODIS satellite data. *Environmental Research Letters* **4,** 044003 (2009).
52. James, W. H. M. *et al.* Source code for: *Gridded birth and pregnancy datasets for Africa, Latin America and the Caribbean.* Figshare https://dx.doi.org/10.6084/m9.figshare.5464147.v1 (2017).
53. Pullum, T. & Becker, S. Evidence of omission and displacement in DHS birth histories. Report No. 11 (ICF International, 2014).
54. Schoumaker, B. *Quality and consistency of DHS fertility estimates, 1990 to 2012*, Report No. 12 (ICF International, 2014).
55. Pullum, T. & Staveteig, S. *An assessment of the quality and consistency of age and date reporting in DHS surveys, 2000-2015*, Report No. 19 (ICF International, 2017).

56. Balk, D. *et al.* Determining global population distribution: methods, applications and data. *Advances in parasitology* **62,** 119–156 (2006).
57. Stevens, F. *et al.* Source code for: *WorldPop-RF, Version 2b.1.1. Figshare* https://doi.org/10.6084/m9.figshare.1491490.v3 (2015).

Data Citations

1. WorldPop. *University of Southampton* https://dx.doi.org/10.5258/SOTON/WP00004 (2017).
2. WorldPop. *University of Southampton* https://dx.doi.org/10.5258/SOTON/WP00138 (2017).
3. WorldPop. *University of Southampton* https://dx.doi.org/10.5258/SOTON/WP00530 (2017).
4. WorldPop. *University of Southampton* https://dx.doi.org/10.5258/SOTON/WP00529 (2017).
5. WorldPop. *University of Southampton* https://dx.doi.org/10.5258/SOTON/WP00528 (2017).
6. WorldPop. *University of Southampton* https://dx.doi.org/10.5258/SOTON/WP00527 (2017).
7. WorldPop. *University of Southampton* https://dx.doi.org/10.5258/SOTON/WP00526 (2017).

Acknowledgements

Author Contributions

W.H.M.J. drafted the manuscript. W.H.M.J., N.T.G, S.E.H., A.C.S., K.N., A.C., C.P., G.M.H., D.K., C.W.R and A.S., acquired and assembled the raw data. W.H.M.J. produced the final datasets and technical validation of the data. N.T.G, S.E.H., K.N., A.C., C.P., G.M.H., C.W.R and A.S., A.J.T. and Z.M. aided drafting the manuscript. All authors read and approved the final version of the manuscript.

Additional information

Tables 1 and 3 are only available in the online version of this paper.

Supplementary information accompanies this paper at http://www.nature.com/sdata

Competing interests: The authors declare no competing interests.

How to cite this article: James, W. H. M. *et al.* Gridded birth and pregnancy datasets for Africa, Latin America and the Caribbean. *Sci. Data* 5:180090 doi: 10.1090/sdata.2018.90 (2018).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.