Data	Source	Data year	Use	Modification	Assumptions for missing data
Division names, codes and boundaries	https://gadm.org/	version 3.6	To plot the shapefile on top of the maps and to create a gridfile that links each grid to a subarea.	The boundaries have been used to develop isogrid files that for each cell provide the country it is inside. Because R cannot deal with factors, we made up our own numbering system. The isorasters are therefore matched with a .csv file that links iso3 codes, (sub)country names and the code in the isoraster file. We created isorasters at 0.5 degree resolution for level 0 input data and 0.1 degree resolution for level 1 input data. We also created an isoraster at 0.0083333 degree resolution for level 0 input data. The boundary shapefile is used to plot the boundaries on top of the gridded results.	There are no missing data.
Population and fraction urban	Kampala Capital City Authority (KCCA)	2018	The emissions are dependent on the population size, as in an area with higher population density and the same incidence, the emissions are higher. The fraction urban is relevant, because toilet categories likely change between urban and rural areas and also where the waste goes. For example, in urban areas there is a lot less space for cover and bury than in rural areas.		There is no missing data
Fraction of population under 5 years of age	Uganda Bureau of Statistics (UBOS)	2018	For some pathogens the incidence is different for younger children compared to the rest of the population	Value for all of Uganda is used for the divisions	Value for all of Uganda is used for the divisions
HDI	http://hdr.undp.org/en/data	2018	The incidence is lower in areas with an HDI higher than the HDI boundary, compared to areas with lower HDI.		HDIs are copied from similar countries. See next tab for details
Gridded population	https://landscan.ornl.gov/	2014	We distribute subarea outputs over the urban and rural populations in the subareas. Grids with higher population density have more emissions	We create urban and rural population files from this gridded total population file. For each country we rank the population grids from high to low. The grids with highest population become urban, until the total urban population (calculated using the urban fractions) is reached. The remaining grids become rural grids. Resolution is 0.008333 degree.	There is no missing data
Toilet category fractions	Kampala Capital City Authority (KCCA) country-wide sanitation survey https://www.kcca.go.ug/uDocs/Improving%20feacal%20sludge%20management%20for%20on-site%20sanitation.pdf	2017	The way the pathogens flow through the sanitation chain and get released to the environment differs from one toilet category to the next. See further explanation below.	We aggregated the survey data over the divisions for each of the 13 JMP categories	There is no missing data
Wastewater and fecal sludge management	Kampala Capital City Authority (KCCA) country-wide sanitation survey, National Water and Sewerage Corporation (NWSC) information on location of plants	2017	The way the waste is managed determines the pathogen removal and release into the environment is explained below.		
Percent removal by wastewater treatment	Results from the Treatment Plant Sketcher tool (tools.waterpathogens.org/sketcher)	2019	The wastewater treatment removes the pathogens from the liquid part of the waste before the effluent is released into the environment. The pathogens in the solid fraction of the waste are not included in the model and not released into the environment.	We use generic information on the Kampala systems with the Treatment Plant Sketcher tool (tools.waterpathogens.org/sketcher) to determine the removal in these plants	
Incidence, shedding rate and shedding duration for Cryptosporidium and rotavirus	As explained in: Hofstra et al 2013 for Cryptosporidium and Kiulia et al 2015 for rotavirus. Based on literature data.		The incidence, shedding rate and shedding duration together determine the excretion per person in the subarea.		