

Lao People's  
Democratic Republic.

Thailand

Cambodia

Vietnam

# Sources de données digitales

Catherine Linard



UNIVERSITÉ  
DE NAMUR



# Plan

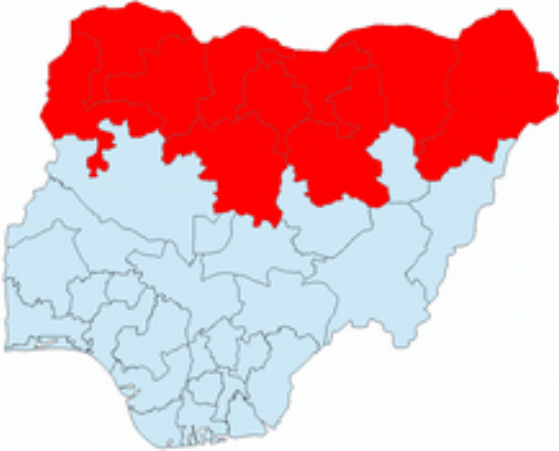
## Gridded population maps

- Human population maps for health & development
- Census data: answers and problems
- Existing gridded population datasets
- WorldPop

## Other interesting data sources

- GADM
- CIESIN
- GeoNetwork
- Worldclim

# Northern Nigeria polio vaccination planning needs



*Vaccinate as close to 100% of under 5s as possible*

- Ensure correct amount of vaccine is available
  - ***Need to know how many under 5s there are***
- Plan local vaccine needs
  - ***Need to know where under 5s are and how numbers change***
- Plan vaccinator logistics and routes
  - ***Need detailed maps of the region***



# Nigeria – 2012 OPV Costs

*OPV = \$0.15/dose*

- Estimated 2012 Population Based on 2006 Census:

**170 million**

- Estimated Target Population < 5 years old (20%):

**34 million**

- OPV Usage based on Tally Sheet Totals:

**67 million**

***33 million doses = \$4.95 million***

# Disease mapping & modelling

## Supporting data required

How many people are there and where are they?



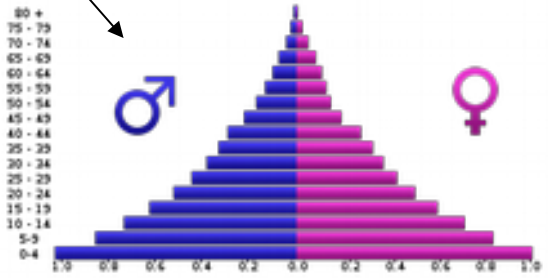
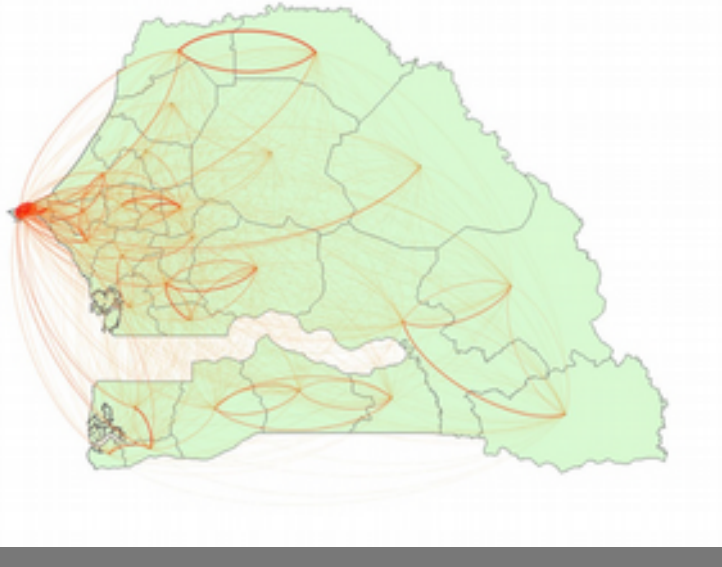
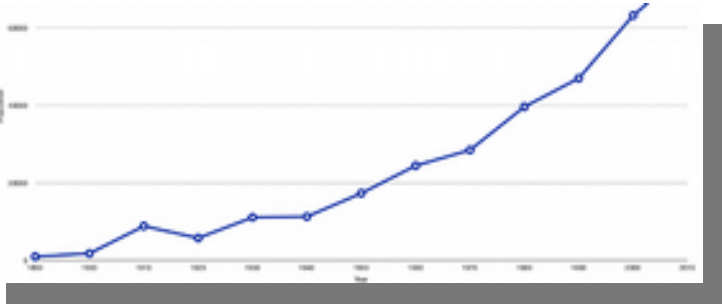
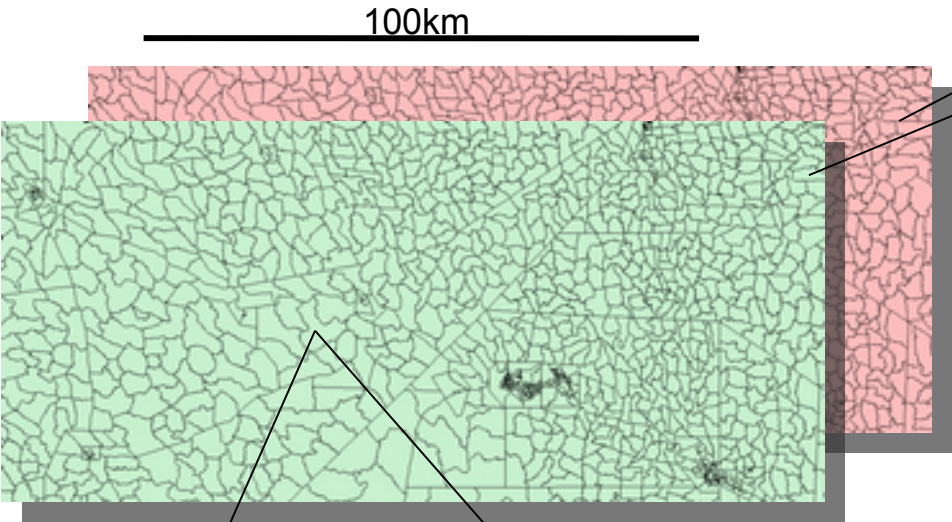
Who are these people?



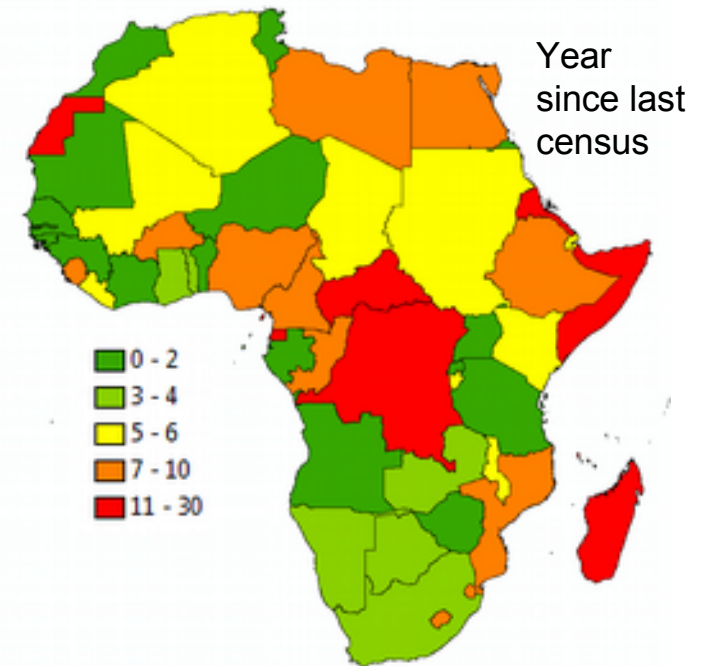
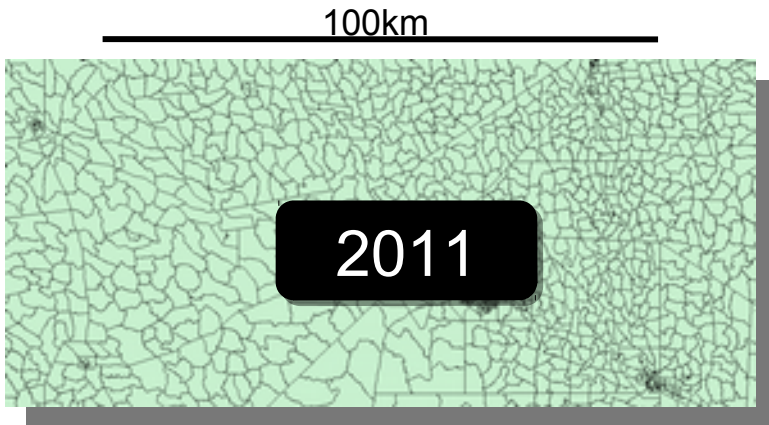
How do they move around?



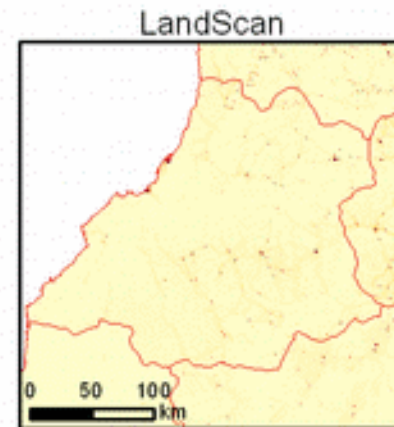
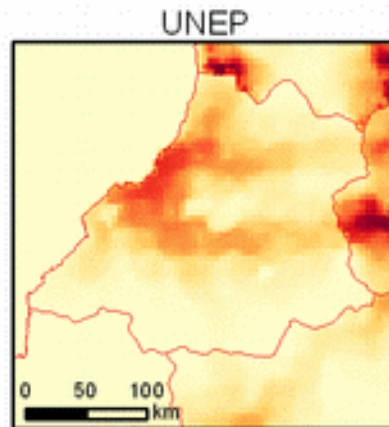
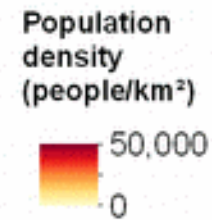
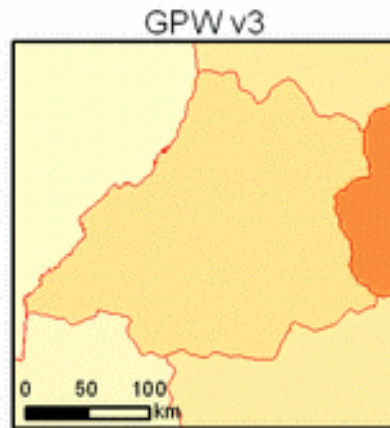
# Census data: answers



# Census data: problems



- Poor availability
- Long processing times
- Lack of matching boundary data
- Population groups missed
- Unreliable numbers
- Inconsistency in urban/rural definitions

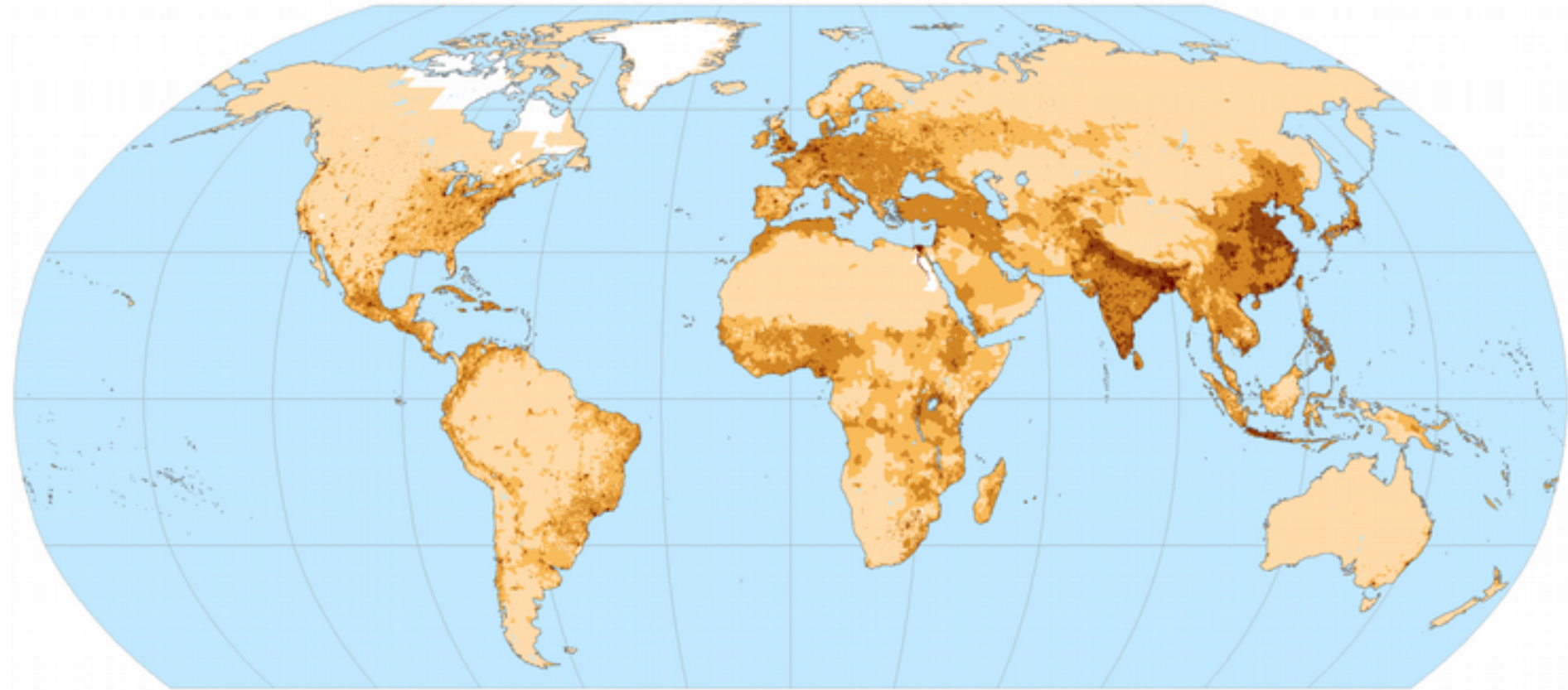


Dataset	Spatial resolution	Level of transparency in data and methods used	Last update	Distribution policy
GPW	5 km	High	2015	Open-access
GRUMP	1 km	High	2004	Open-access
UNEP	5 km	High	2000	Open-access
LandScan	1 km	Low	2012	Commercial

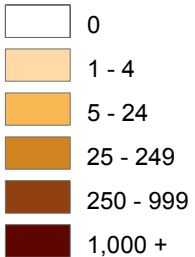


# Mapping population distributions

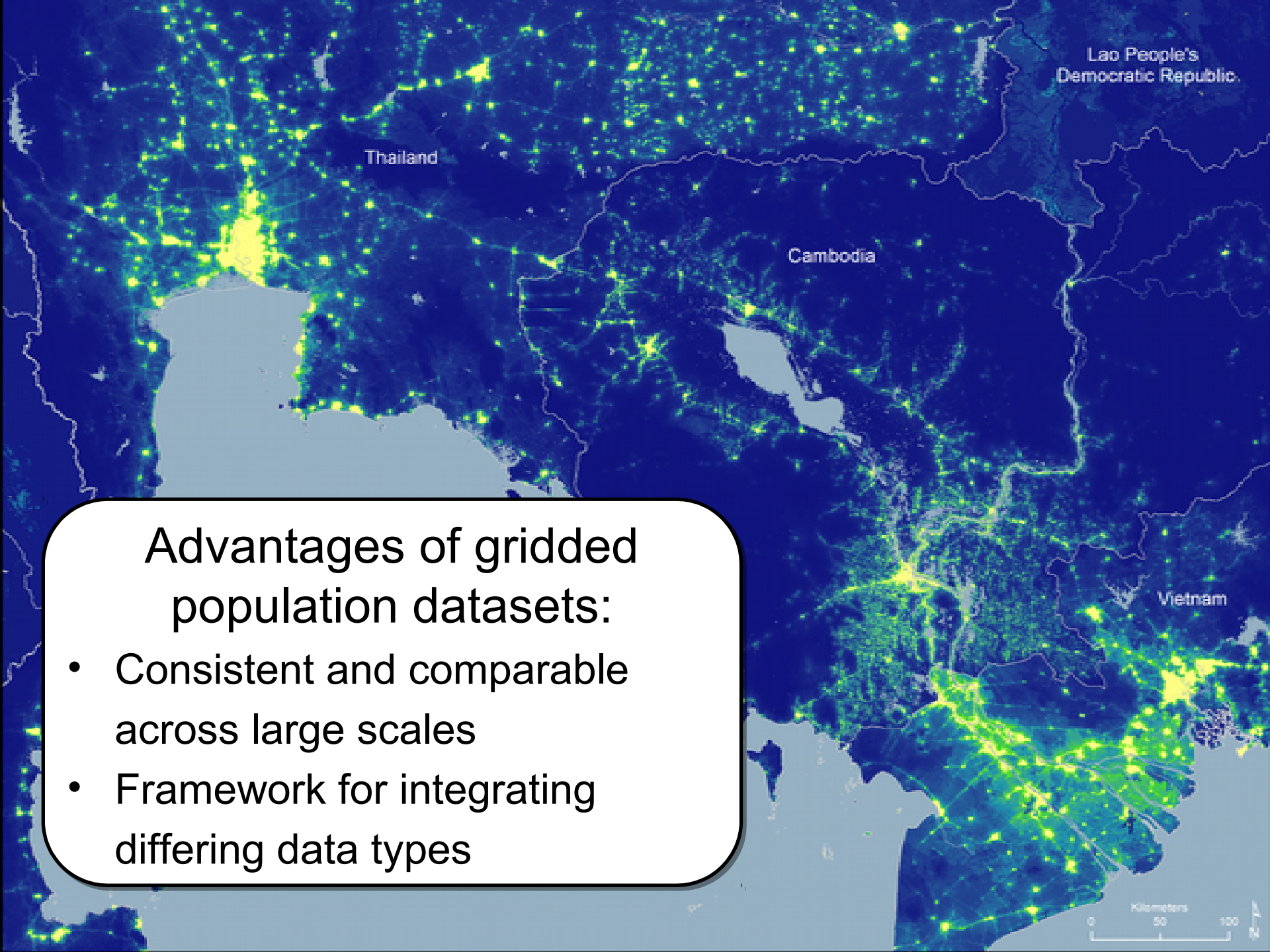
Gridded Population of the World, v3



Persons per km<sup>2</sup>



- Raster surface of population counts and densities
- Derived from ~400,000 administrative units and census data
- 2.5' resolution (approximately 4.6 km/2.9mi at the equator)



## Advantages of gridded population datasets:

- Consistent and comparable across large scales
- Framework for integrating differing data types

# World pop



[www.worldpop.org](http://www.worldpop.org)

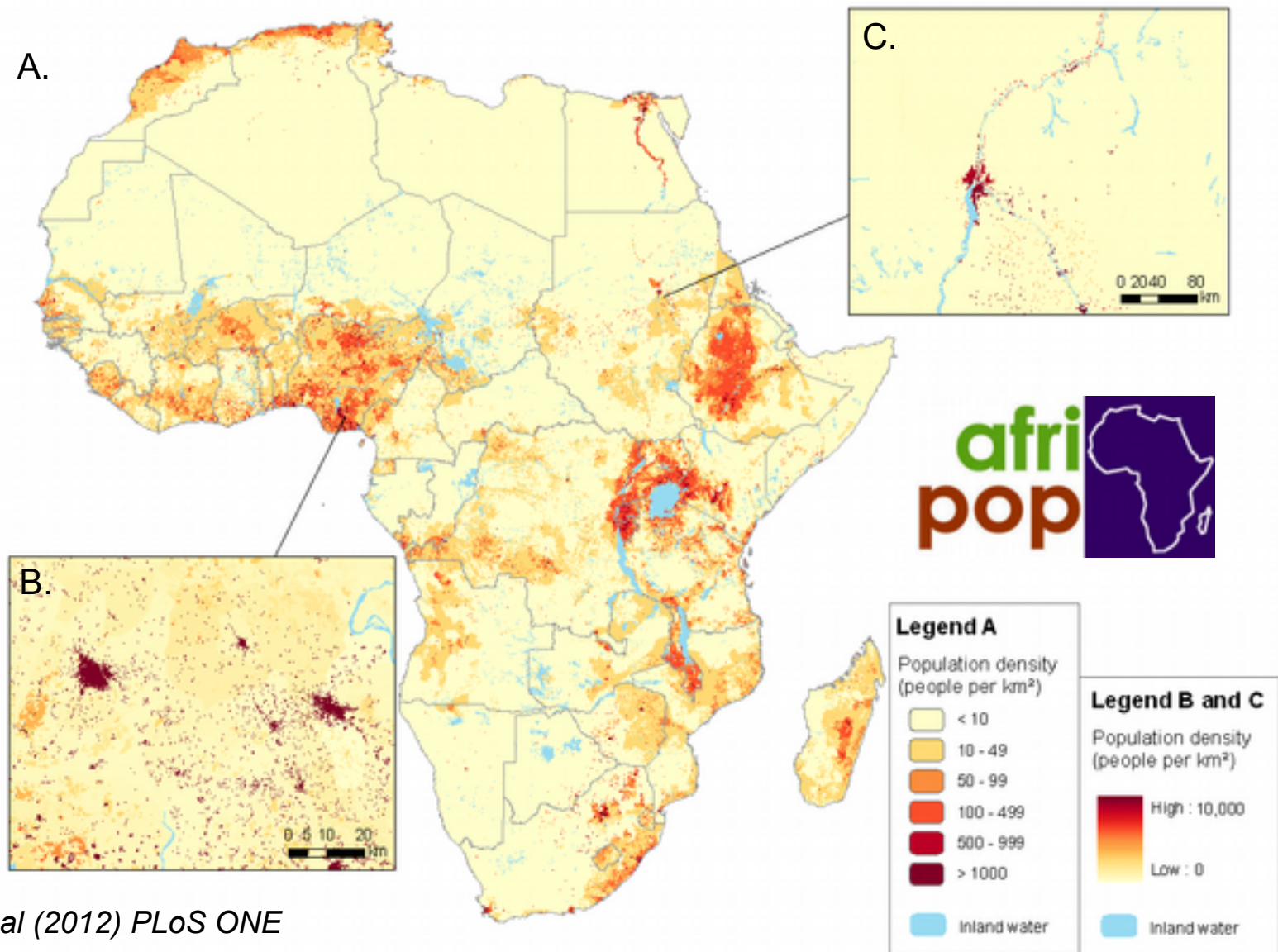
Open access archive of  
spatial demographic datasets

Current focus: Central and  
South America, Africa and  
Asia

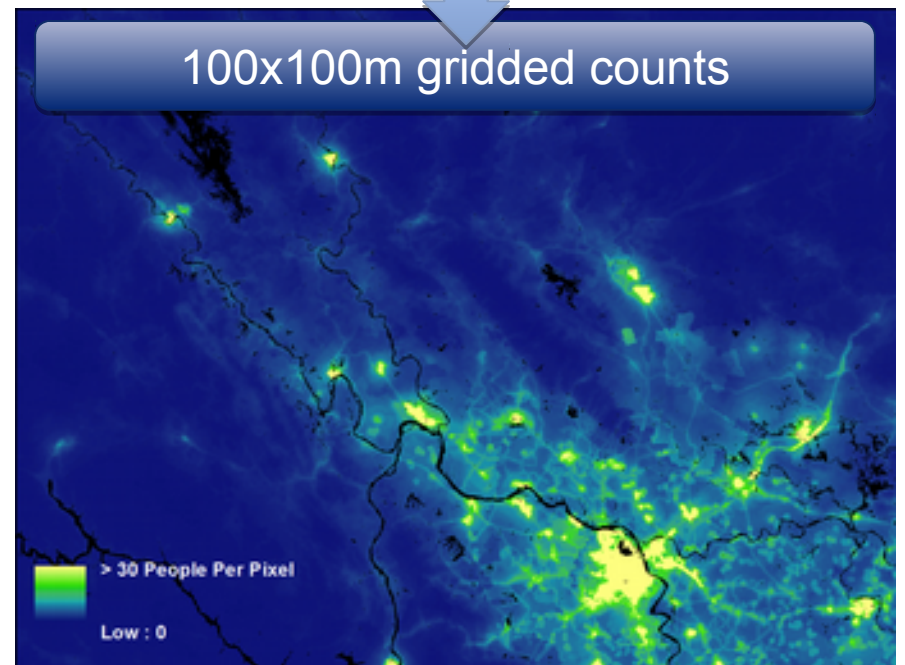
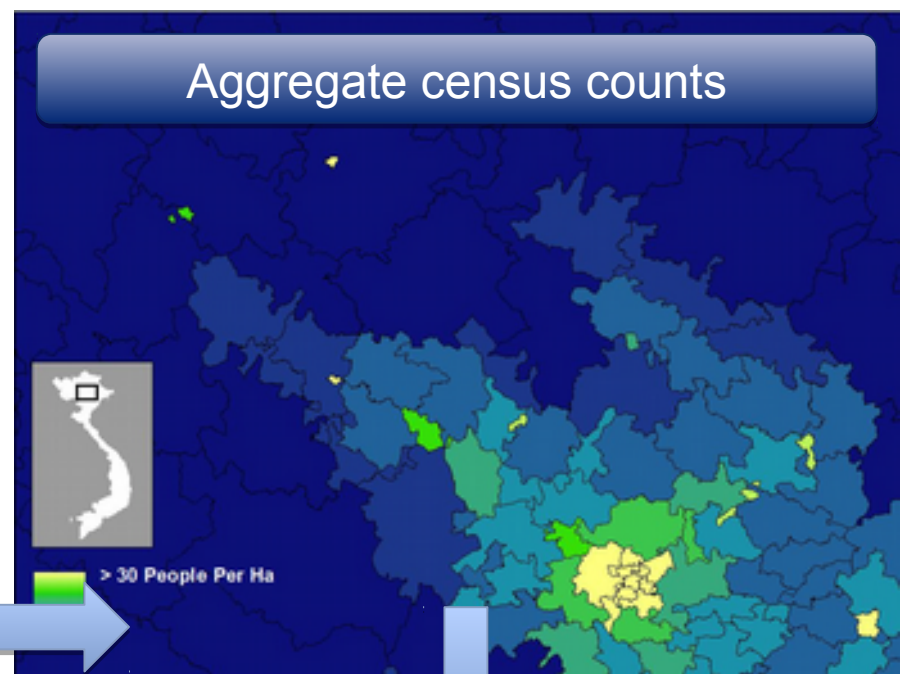
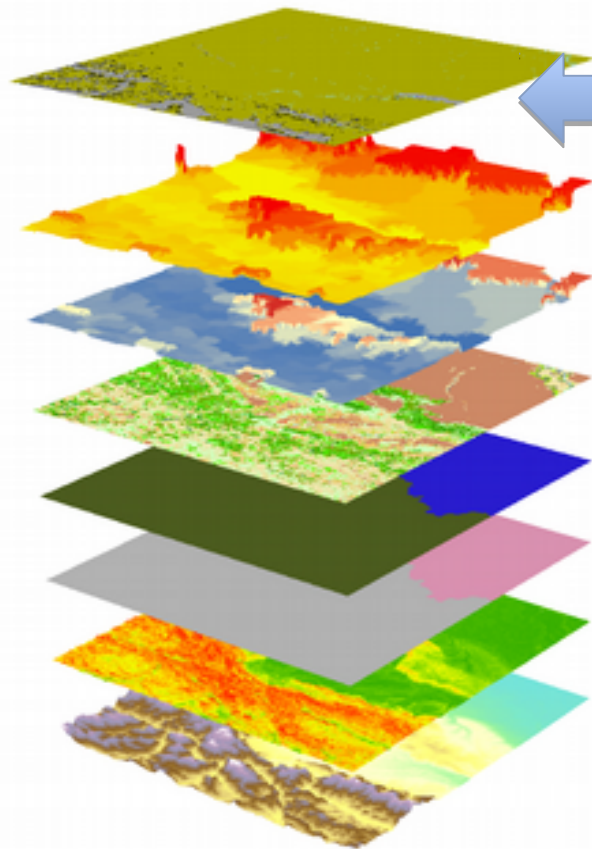
Transparent and shareable  
methods

Support development and  
health applications

# AfriPop 2010



# Top-down approach



# Intro to gridded population data

Census data linked to  
GIS administrative  
boundaries



Ancillary data e.g.  
Settlements, roads

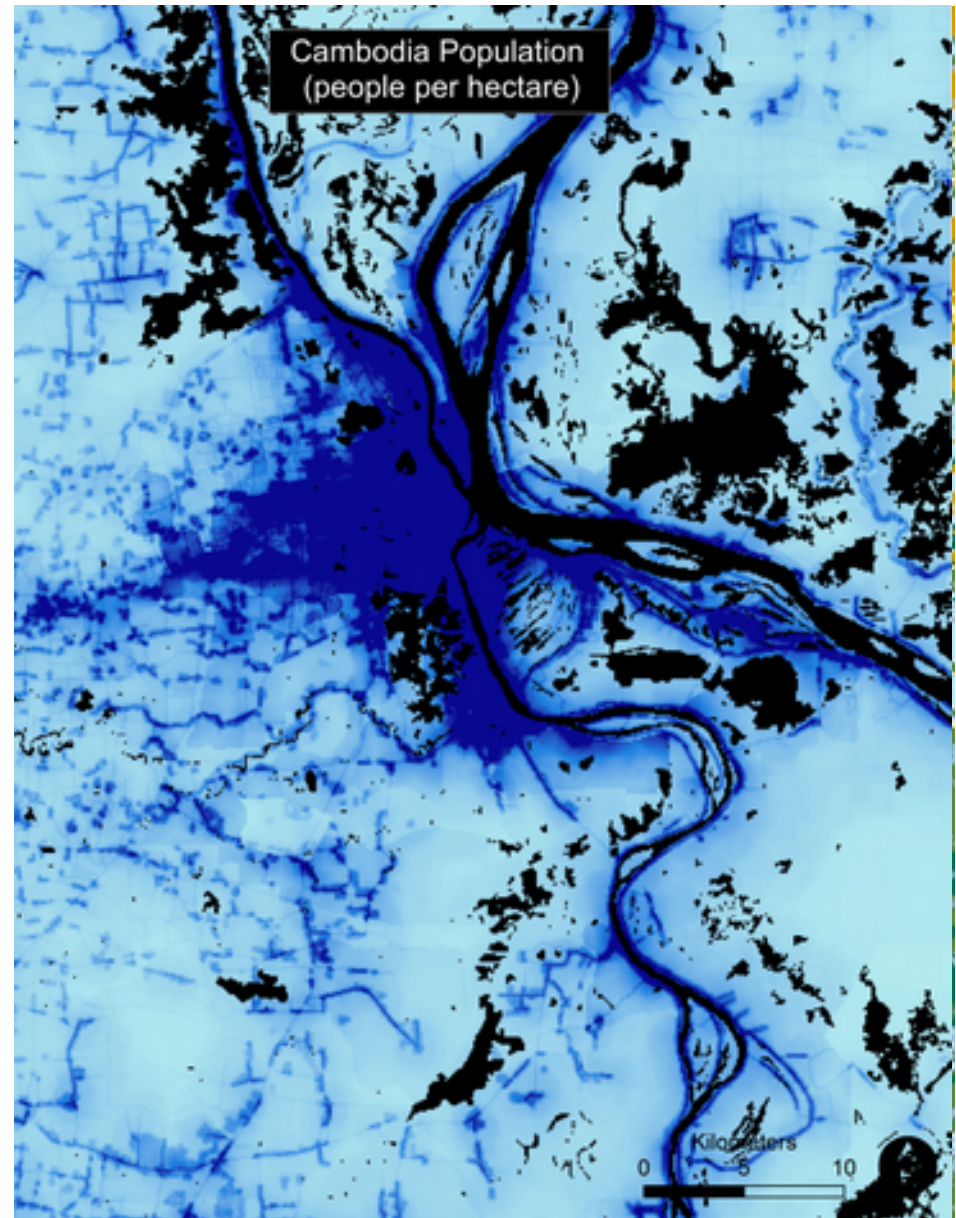


Spatial modelling rules to  
disaggregate census  
counts



Estimates of number of  
people in each grid cell

Population Density (pph)



# Open code

```
RStudio
01.1 - Data Preparation, R.r x Metadata.r x .MODIS_Opts.R x 01.3 - More Complex Random Forest Regression, R.r x
Source on Save
0 ## WP/Data/RF folder where RF represents the country name for which the
7 ## RF folder is located.
8
9 ## NOTE: It's critical that you realize that anything you specify here
10 ## will be overridden by the values read in from the Metadata.r file
11 ## so the paths may be incorrect if specified incorrectly inside that
12 ## file!
13
14 ## Configure the country abbreviation and name:
15 country <- "KHM"
16
17 ## This should be set to the folder "containing" the "RF" folder structure:
18 root_path <- "C:/Research/Population/Data/"
19 project_path <- paste(root_path, "RF/data/", country, "/", sep="")
20
21 ## Load the metadata from the file created in the country's /data folder:
22 source(paste(project_path, "Metadata.r", sep=""))
23
24 ## END: Load configuration options
25 #####
26
27
28
29 #####
30 ## BEGIN: RandomForest configuration
31
32 ## Configuration options for RandomForest mode!
33
34 ## NOTE: The following were moved to the Meta
35 ## configuration and reporting purposes:
36
```

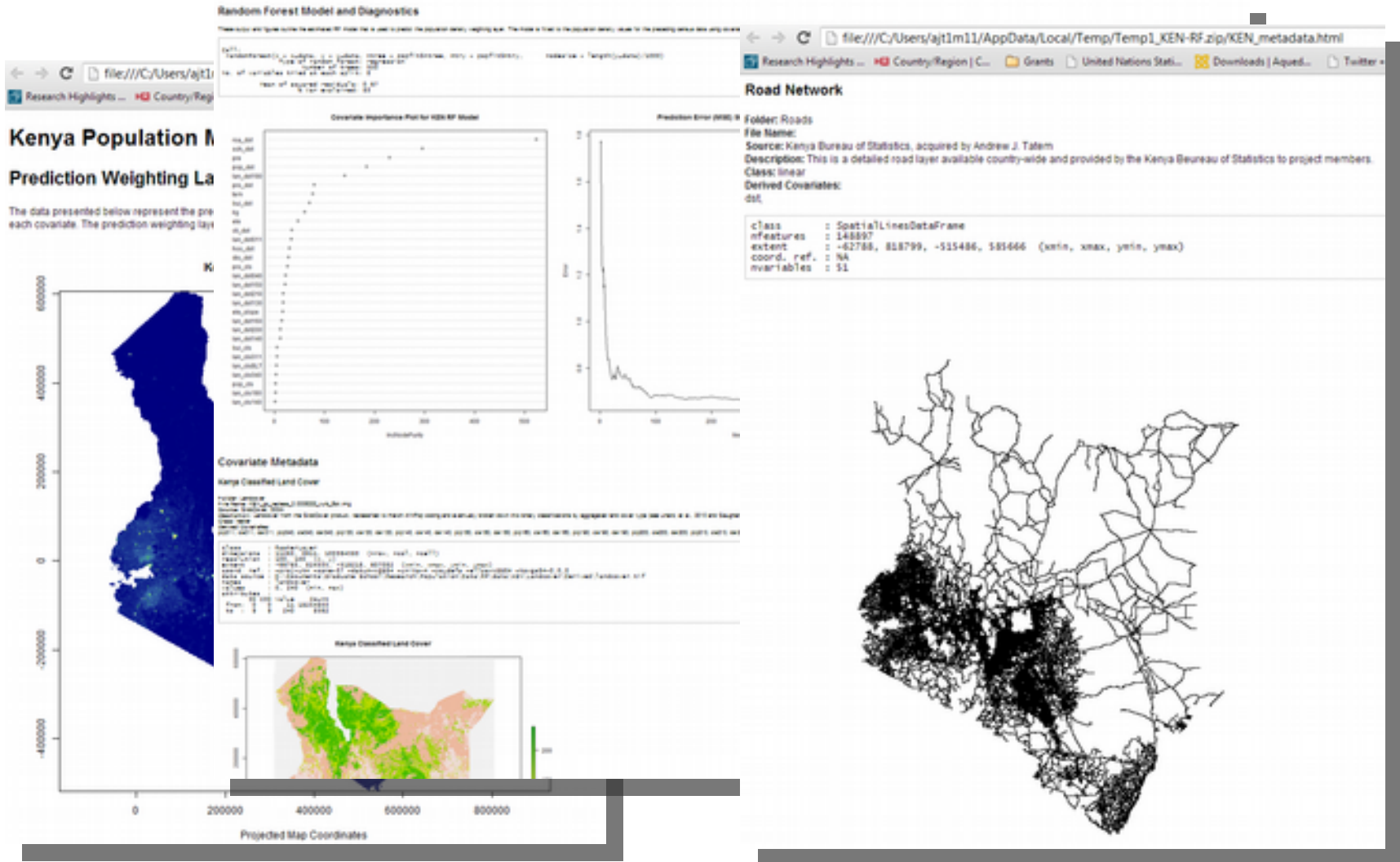


```
01.2 - Data Preparation, Python - C:/Research/Population/Data/RF/01.2 - Data Preparation, Python
File Edit Format Run Options Windows Help
## BEGIN: Set per-country configuration options
country = "KHM"
#country = "KHM_002"
#country = "KHM_002_Anc"
## Configure project and default data folders:
root_path = "C:/Research/Population/Data/"
#root_path = "C:/Users/zorrest/Research/Population/Data/"
#root_path = "D:/Research/Population/Data/"
## NOTE: This needs to be adjusted depending on whether we are in
## the Asia/Australia region or whether we are in the Africa/Americ
## region. If you are processing a large country that does not fal
## "completely" within one of the VMAP0 data regions then you should
## either merge the datasets by hand or provide alternatives to the
## following default datasets: "Roads", "Rivers", "Populated"
NGA_path = root_path + "GIS/NGA/VMAP0/v0sas_5/vmaplv0/sasaus:"
#NGA_path = root_path + "GIS/NGA/VMAP0/v0soa_5/vmaplv0/soamafr:"
#NGA_path = root_path + "GIS/NGA/VMAP0/v0noa_5/vmaplv0/noamer1:"
#NGA_path = root_path + "GIS/NGA/VMAP0/v0eur_5/vmaplv0/eurnaa1a:"
## Output projection is configured for each country based on the most
## appropriate output for distance/area considerations:
## For KHM, VMS:
intermediate_prj = "PROJCS['WGS_1984_UTM_Zone_48N',GEOGCS['GCS_WGS_1984',DATUM['
## For KEN:
#intermediate_prj = "PROJCS['WGS_1984_UTM_Zone_37N',GEOGCS['GCS_WGS_1984',DATUM[
## Should we skip processing :
skip_existing = True
## END: Set per-country con
#####
Ln 1 Col 0
```



- Feedback & improvements
- Methodological clarity
- User-specific datasets
- Platform independent

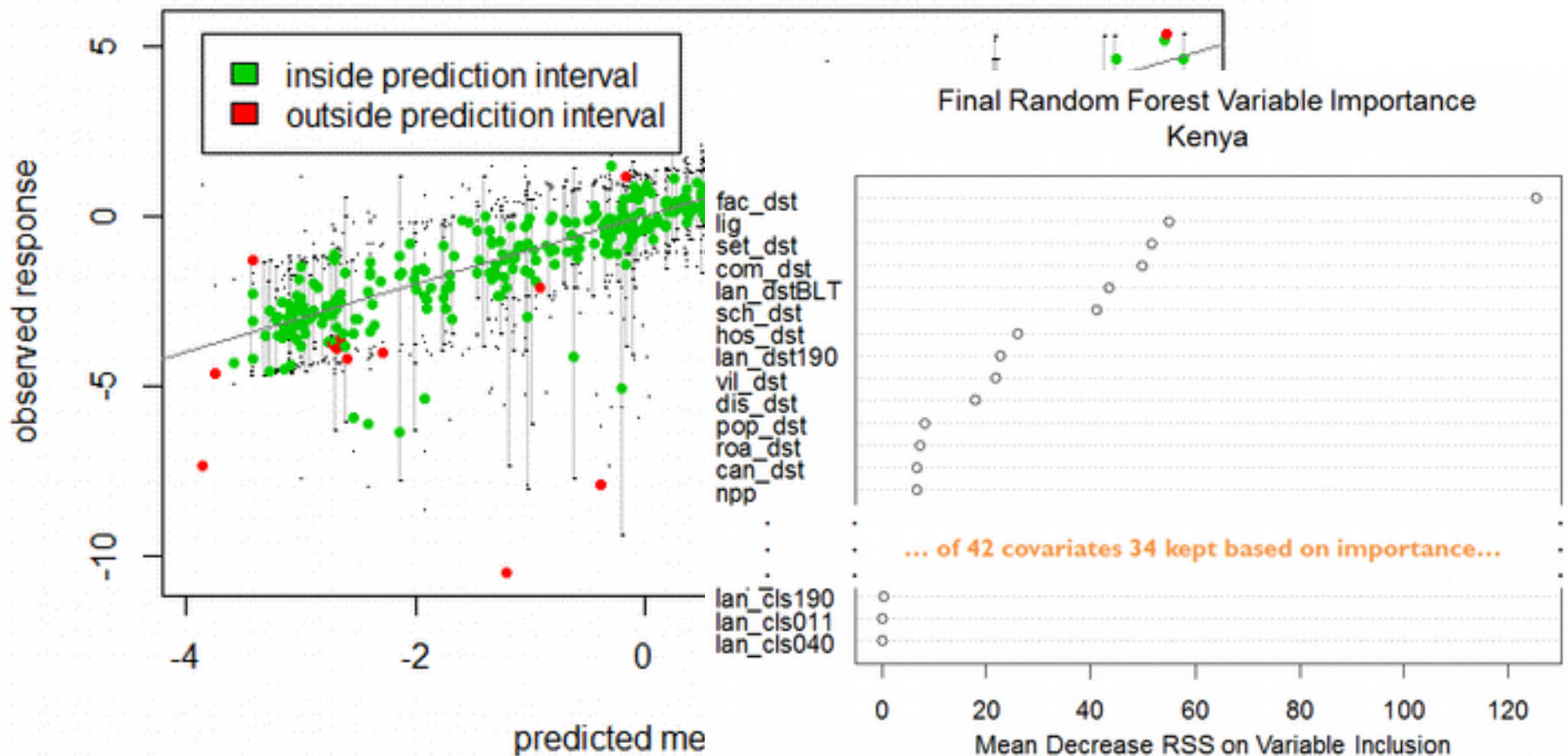
# Automated metadata



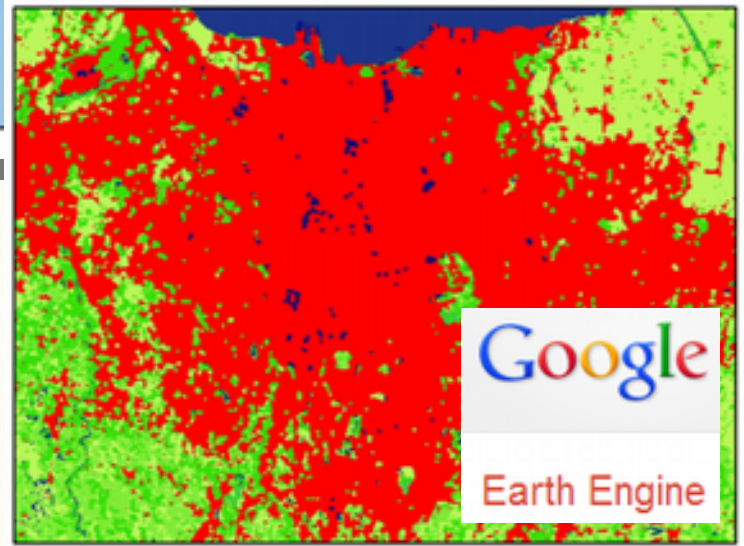
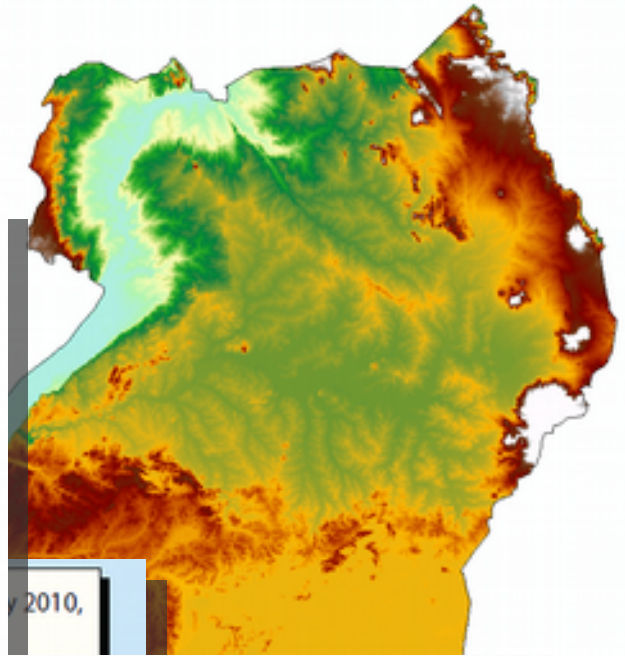
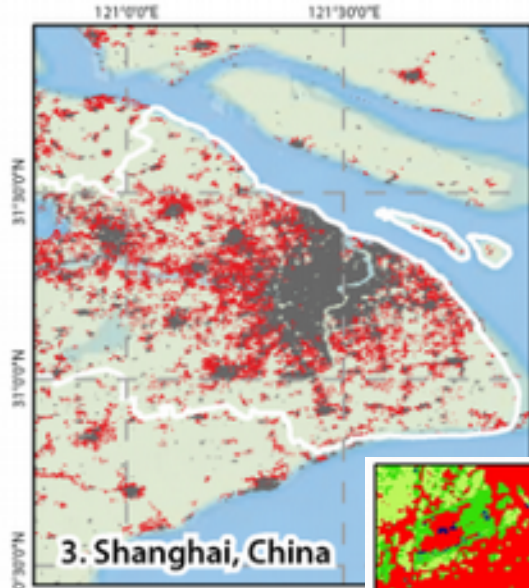
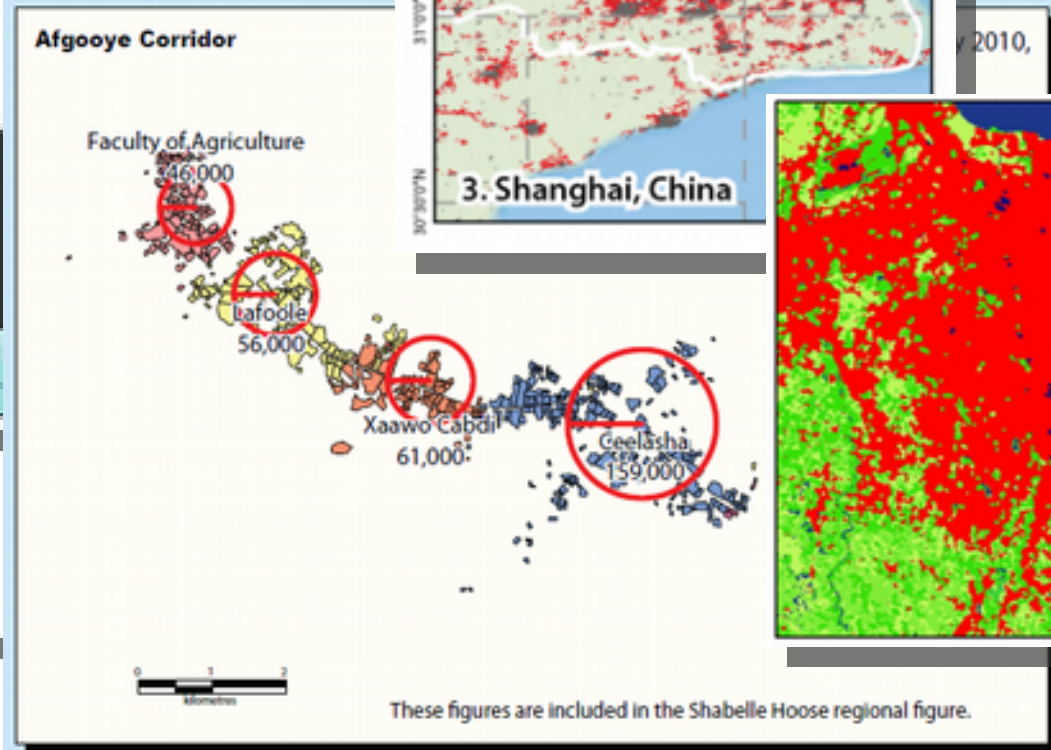
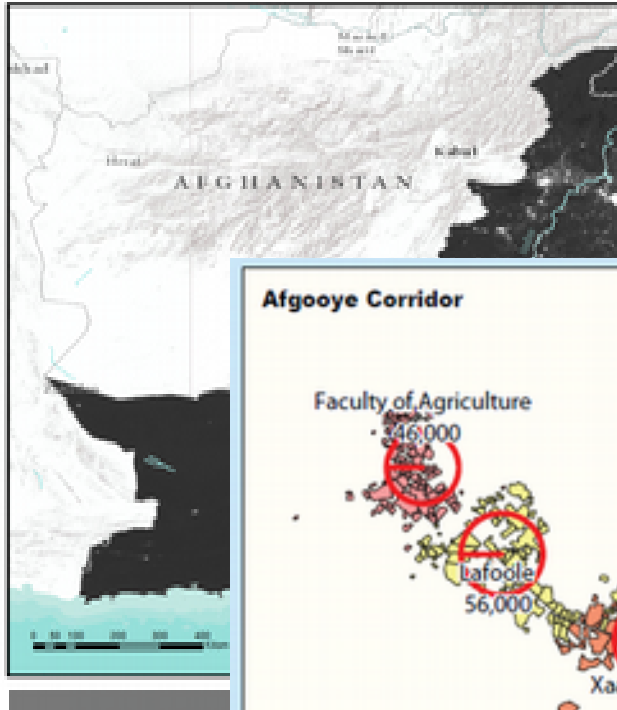


# Internal accuracy assessments

## 90 % prediction intervals on out-of-bag data



# Input covariates



# Input covariates

Type	Variable Name(s)*	Description	Cambodia Data	Vietnam Data	Kenya Data
Census		Country-specific census and scale	2008, Admin-level 3	1999, Admin-level 4	1999, Admin-level 5
Land Cover	lan_cls011, lan_dst011	Cultivated terrestrial lands	MDA Landcover, 30m	MDA Landcover, 30m	GlobCover, 300m
	lan_cls040, lan_dst040	Woody / Trees			
	lan_cls130, lan_dst130	Shrubs			
	lan_cls140, lan_dst140	Herbaceous			
	lan_cls150, lan_dst150	Other terrestrial vegetation			
	lan_cls160, lan_dst160	Aquatic vegetation			
	lan_cls190, lan_dst190	Urban area			
	lan_cls200, lan_dst200	Bare areas			
	lan_cls210, lan_dst210	Water bodies			
	lan_cls230, lan_dst230	No data, cloud/shadow			
	lan_cls240, lan_dst240	Rural settlement			
	lan_cls250, lan_dst250	Industrial area			
		lan_clsBLT, lan_dstBLT	Built, merged urban/rural class		
Continuous					
Raster-Format	Npp	Annual NPP, 2010	MODIS 17A3	MODIS 17A3	MODIS 17A3
	Lig	Lights at night	Suomi VIIRS-Derived	Suomi VIIRS-Derived	Suomi VIIRS-Derived
	Tem	Mean temperature, 1950–2000	WorldClim/BioClim	WorldClim/BioClim	WorldClim/BioClim
	Pre	Mean precipitation, 1950–2000	WorldClim/BioClim	WorldClim/BioClim	WorldClim/BioClim
	Ele	Elevation	HydroSHEDS [31]	HydroSHEDS[31]	HydroSHEDS[31]
	ele_slope	Slope	HydroSHEDS-Derived	HydroSHEDS-Derived	HydroSHEDS-Derived
Converted					
Vector-Format	roa_dst	Distance to roads	World Food Programme	OSM [39]	Kenyan Bureau of Stat.
	riv_dst	Distance to rivers/streams	World Food Programme	VMAPO hydro/watcrsl	VMAPO hydro/watcrsl
	pop_cls, pop_dst	Generic populated places	VMAPO merged <sup>f</sup>	VMAPO merged <sup>f</sup>	Tatem, et al. [46]
	wat_cls, wat_dst	Water bodies	World Food Programme	OSM [39]	VMAPO hydro/inwaters
	pro_cls, pro_dst	Protected areas	WDPA, IUCN [47]	WDPA, IUCN [47]	WDPA, IUCN [47]
	can_cls, can_dst	Canals	World Food Programme		
	com_cls, com_dst	Communities	World Food Programme		
	dis_cls, dis_dst	District seats	World Food Programme		
	cit_cls, cit_dst	Cities		OSM [39]	
	ham_cls, ham_dst	Hamlets		OSM [39]	
	vil_cls, vil_dst	Villages	KHM Census 2008	OSM [39]	
	sub_cls, sub_dst	Suburbs		OSM [39]	
	tow_cls, tow_dst	Towns		OSM [39]	
	poi_cls, poi_dst	Populated Points		OSM [39]	
	rai_cls, rai_dst	Railways		OSM [39]	
	fac_cls, fac_dst	Generic health facilities	HR-COD [48]		
	cli_cls, cli_dst	Health clinics			Noor, et al. [44]
	dsp_cls, dsp_dst	Dispensaries	HR-COD [48]		Noor, et al. [44]
	hos_cls, hos_dst	Hospitals			Noor, et al. [44]
	sch_cls, sch_dst	Schools	HR-COD [48]		Kenya Open Data [49]
set_cls, set_dst	Settlement points	World Food Programme			
bui_cls, bui_dst	Built land cover			Tatem, et al. [46]	

# Input satellite-derived covariates

Type	Variable Name(s)*	Description	Cambodia Data	Vietnam Data	Kenya Data
Census		Country-specific census and scale	2008, Admin-level 3	1999, Admin-level 4	1999, Admin-level 5
Land Cover	lan_cls011, lan_dst011	Cultivated terrestrial lands	MDA Landcover, 30m	MDA Landcover, 30m	GlobCover, 300m
	lan_cls040, lan_dst040	Woody / Trees			
	lan_cls130, lan_dst130	Shrubs			
	lan_cls140, lan_dst140	Herbaceous			
	lan_cls150, lan_dst150	Other terrestrial vegetation			
	lan_cls160, lan_dst160	Aquatic vegetation			
	lan_cls190, lan_dst190	Urban area			
	lan_cls200, lan_dst200	Bare areas			
	lan_cls210, lan_dst210	Water bodies			
	lan_cls230, lan_dst230	No data, cloud/shadow			
	lan_cls240, lan_dst240	Rural settlement			
	lan_cls250, lan_dst250	Industrial area			
	lan_clsBLT, lan_dstBLT	Built, merged urban/rural class			
	Continuous				
Raster-Format	Npp	Annual NPP, 2010	MODIS 17A3	MODIS 17A3	MODIS 17A3
	Lig	Lights at night	Suomi VIIRS-Derived	Suomi VIIRS-Derived	Suomi VIIRS-Derived
	Tem	Mean temperature, 1950–2000	WorldClim/BioClim	WorldClim/BioClim	WorldClim/BioClim
	Pre	Mean precipitation, 1950–2000	WorldClim/BioClim	WorldClim/BioClim	WorldClim/BioClim
	Ele	Elevation	HydroSHEDS [31]	HydroSHEDS[31]	HydroSHEDS[31]
	ele_slope	Slope	HydroSHEDS-Derived	HydroSHEDS-Derived	HydroSHEDS-Derived

# Données démographiques WorldPop

- Données démographiques disponibles en accès libre et gratuit
- Basé sur un processus de “machine learning” (=modélisation)
- Intègre des données de recensement, d'enquêtes, des images satellitaires et des données SIG
- WorldPop fournit notamment des estimations du nombre/de la densité de personnes résidant dans des mailles de 100x100m (2000-2020)
- Aussi : structures d'âge, naissances, grossesses
- Les cartes sont accompagnées de métadonnées

# Applications possibles

- Epidémiologie spatiale
- Gestion des catastrophes
- Modèles d'accessibilité
- Transport
- Aménagement du territoire
- Etudes d'impact environnemental

etc.



# residential population count and composition mapping

## Demographics



Maternal and newborn health



THE WORLD BANK  
Working for a World Free of Poverty

Poverty mapping

## Dynamics



THE WORLD BANK  
Working for a World Free of Poverty



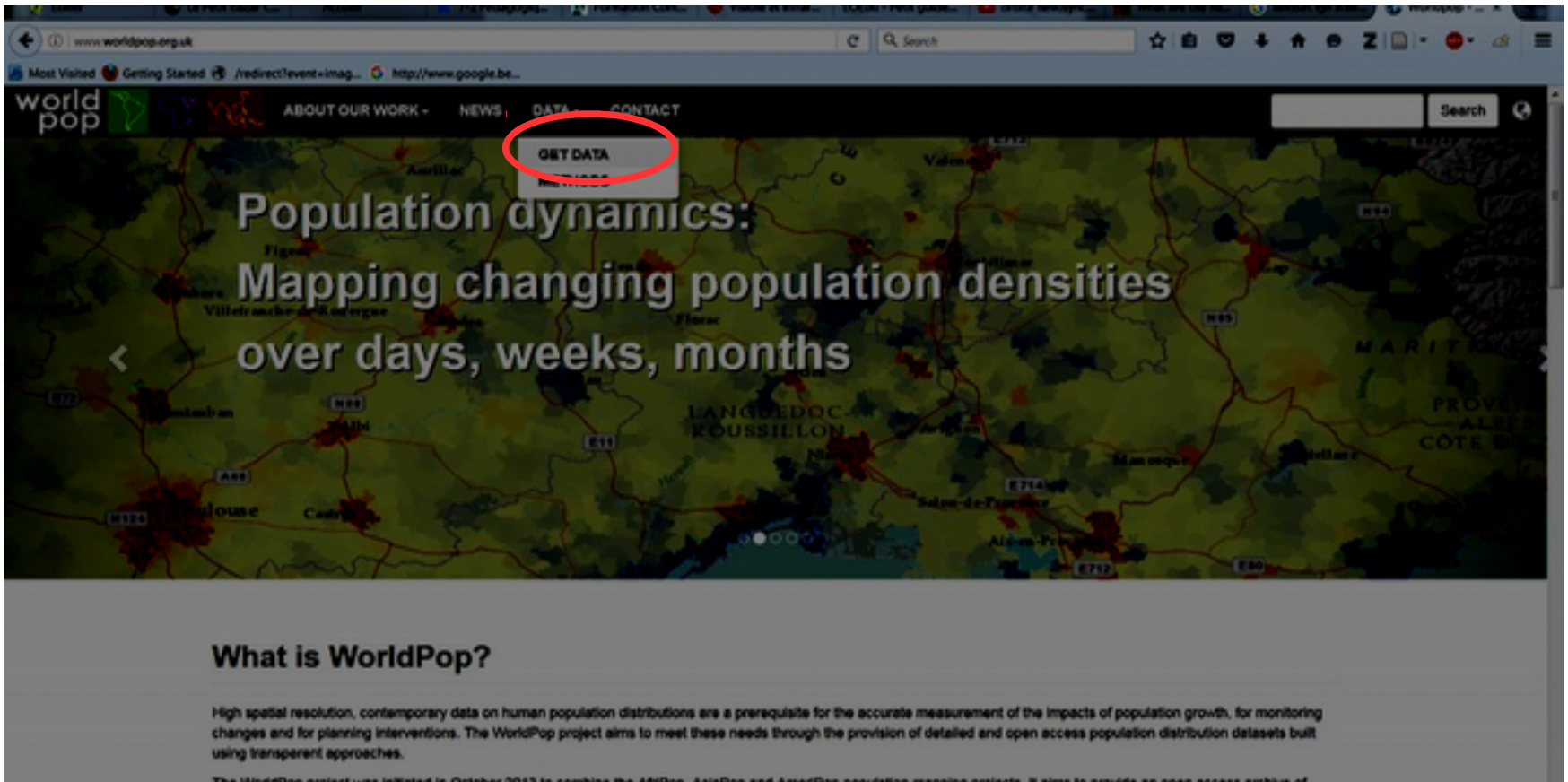
Earth Engine

Urbanization and settlement growth



Population mobility

# Télécharger des données



The image shows a screenshot of the WorldPop website. The browser address bar displays 'www.worldpop.org.uk'. The navigation menu includes 'ABOUT OUR WORK -', 'NEWS', 'DATA', and 'CONTACT'. The 'DATA' link is highlighted with a red circle, and a 'GET DATA' button is also highlighted with a red circle. The main content area features a map of France with the text 'Population dynamics: Mapping changing population densities over days, weeks, months'. Below the map, there is a section titled 'What is WorldPop?' with a paragraph of text.

www.worldpop.org.uk

Most Visited Getting Started /redirect?event+imag... http://www.google.be...

world pop

ABOUT OUR WORK - NEWS DATA CONTACT

GET DATA

Population dynamics:  
Mapping changing population densities  
over days, weeks, months

What is WorldPop?

High spatial resolution, contemporary data on human population distributions are a prerequisite for the accurate measurement of the impacts of population growth, for monitoring changes and for planning interventions. The WorldPop project aims to meet these needs through the provision of detailed and open access population distribution datasets built using transparent approaches.

The WorldPop project was initiated in October 2013 to combine the AfrPop, AsiaPop and AmerPop population mapping projects. It aims to provide an open access archive of

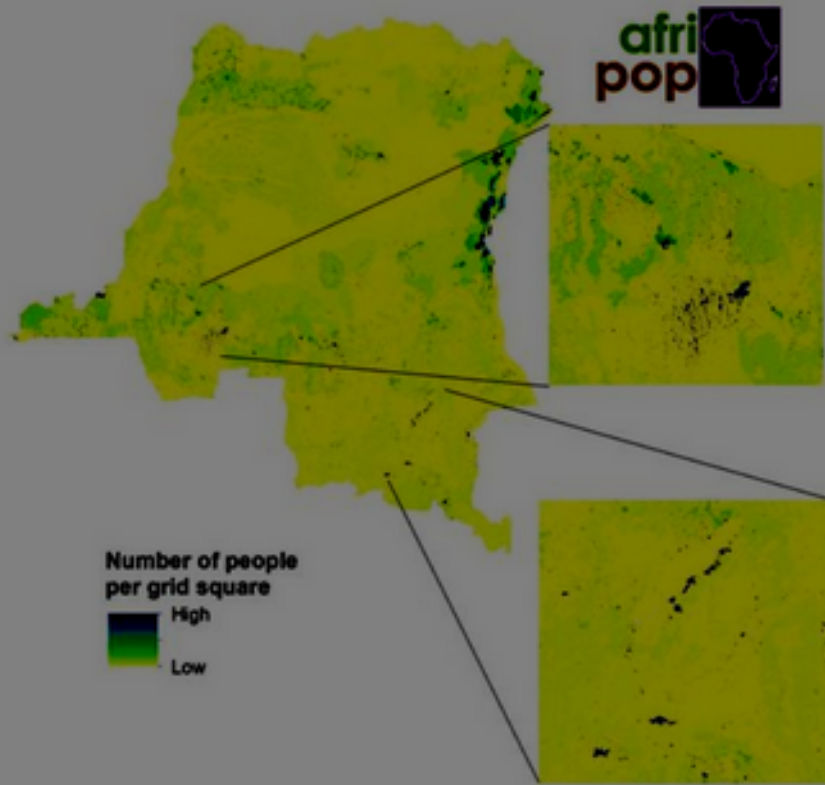


# Lire les métadonnées

## Summary

Selected Data : Africa > Democratic Republic of the Congo > Population

### Democratic Republic of the Congo



AfriPop ([www.afripop.org](http://www.afripop.org)) dataset details

**DATASET:** Alpha version 2010 and 2015 estimates of numbers of people per grid square, with national totals adjusted to match UN population division estimates (<http://esa.un.org/wpp/>) and remaining unadjusted.

**REGION:** Africa

**SPATIAL RESOLUTION:** 0.000833333 decimal degrees (approx 100m at the equator)

**PROJECTION:** Geographic, WGS84

**UNITS:** Estimated persons per grid square

**MAPPING APPROACH:** Land cover based, as described in: Linard, C., Gilbert, M., Snow, R.W., Noor, A.M. and Tatem, A.J., 2012, Population distribution, settlement patterns and accessibility across Africa in 2010, PLoS ONE, 7(2): e31743.

**FORMAT:** Geotiff (zipped using 7-zip (open access tool): [www.7-zip.org](http://www.7-zip.org))

**FILENAMES:** Example - AGO10adjv4.tif = Angola (AGO) population count map for 2010 (10) adjusted to match UN national estimates (adj), version 4 (v4). Population maps are updated to new versions when improved census or other input data become available.

**DATE OF PRODUCTION:** January 2013

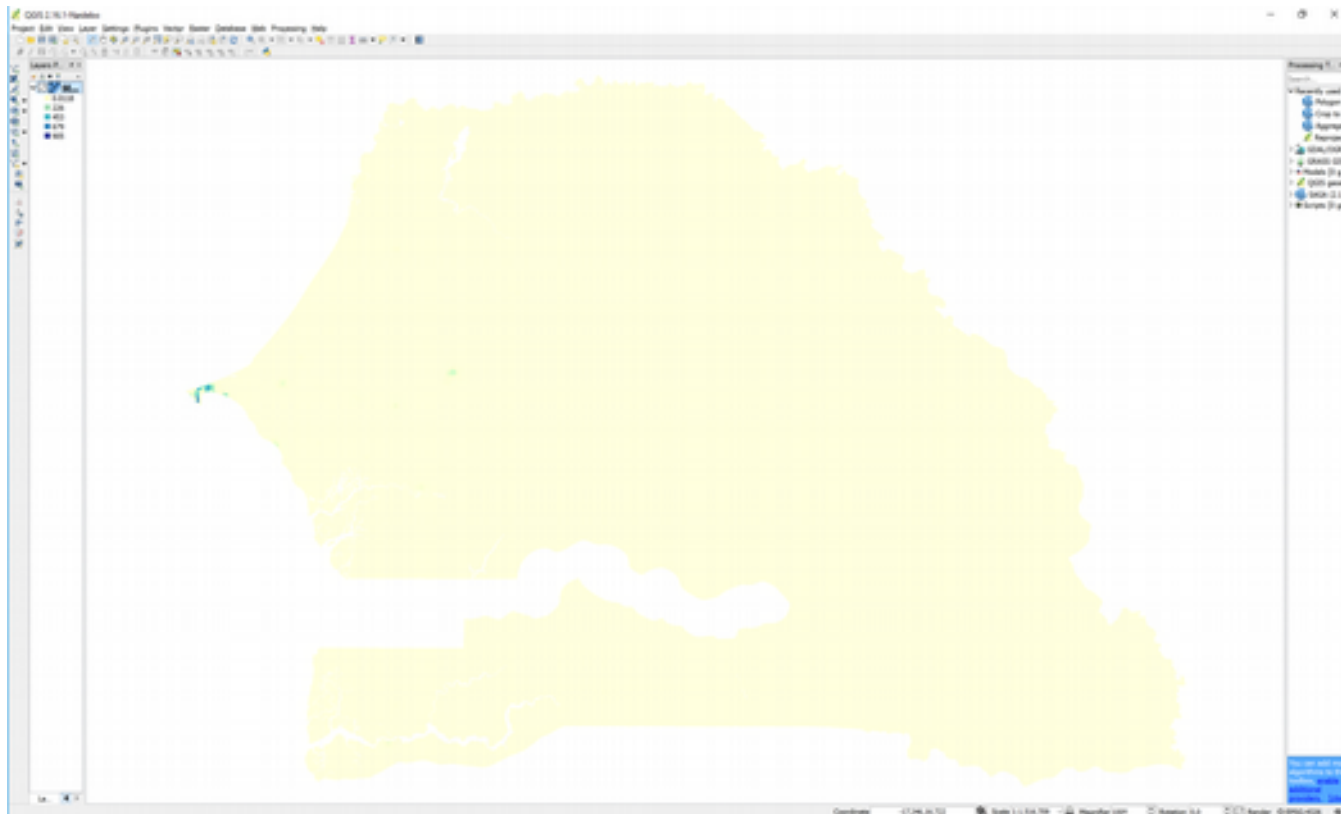
Dataset construction details and input data are provided here: [www.afripop.org](http://www.afripop.org) and here: <http://www.plosone.org/article/info:doi/10.1371/journal.pone.0031743>

# Noms de fichier - exemple

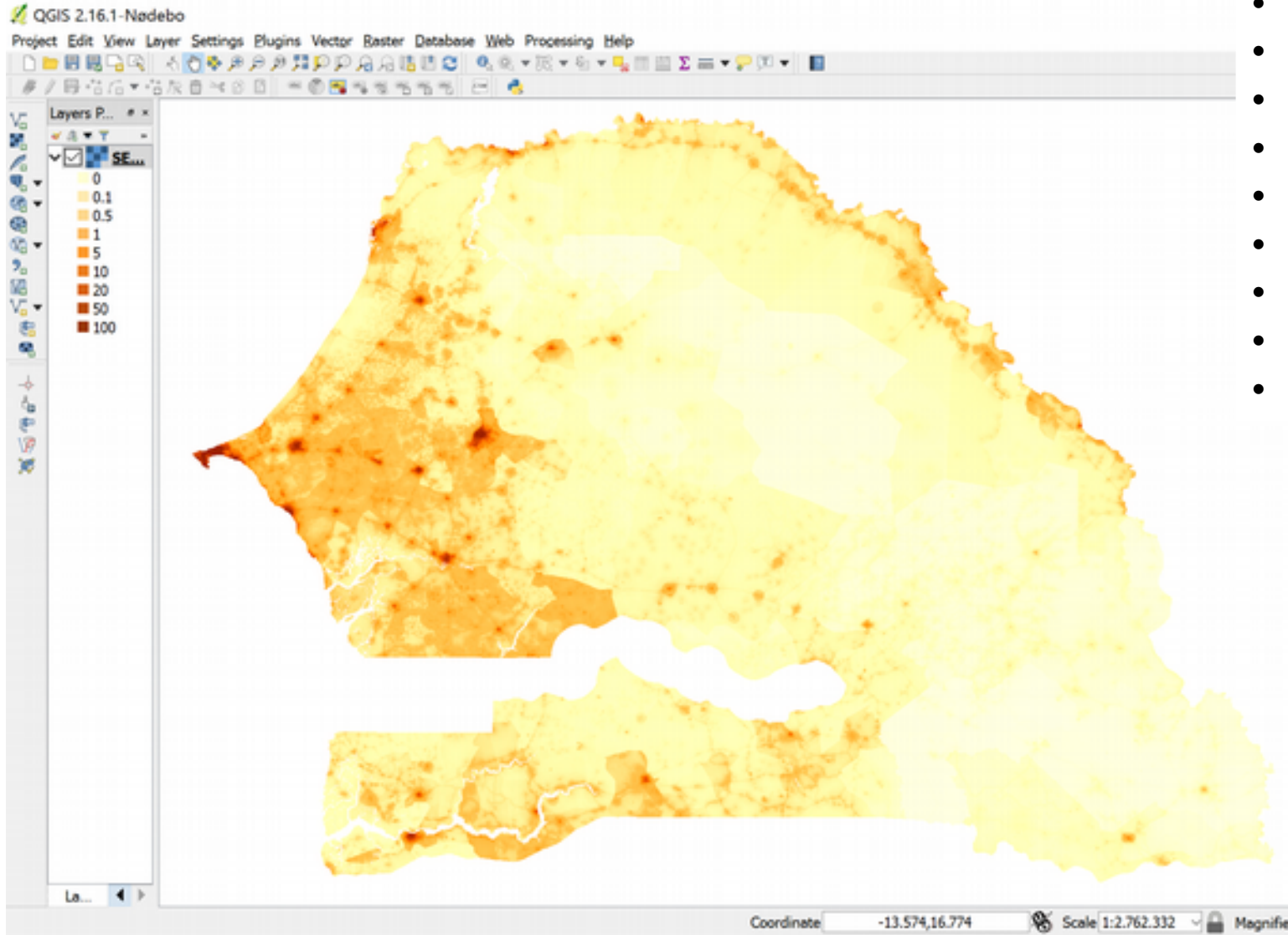
- SEN\_ppp\_v2b\_2015\_UNadj.tif =
  - Sénégal (SEN)
  - population par pixel (ppp) ou population par hectare (pph)
  - année (2015)
  - adjusted to match UN national estimates (UNadj)
  - version 2b (v2b)

# WorldPop dans QGis

- La distribution asymétrique des valeurs (voir histogramme) rend la visualisation difficile avec une palette de couleur par défaut dans QGis (attention: utilisez la fonction « load min/max values » pour voir les valeurs min et max correctes)



# Palette de couleur personnalisée



9 classes:

- 0
- 0 – 0,1
- 0,1 – 0,5
- 0,5 – 1
- 1 – 5
- 5 – 10
- 10 – 20
- 20 – 50
- 50 – 100

# Plan

## Gridded population maps

- Human population maps for health & development
- Census data: answers and problems
- Existing gridded population datasets
- WorldPop

## Other interesting data sources

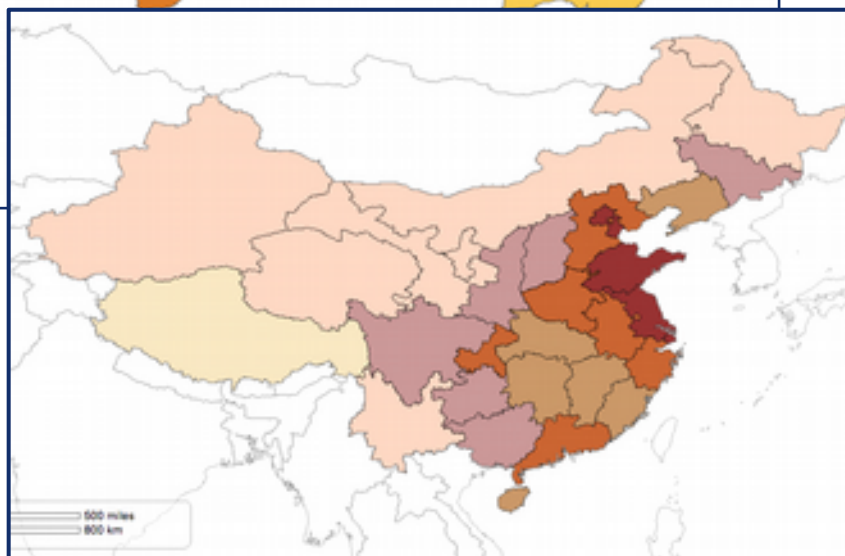
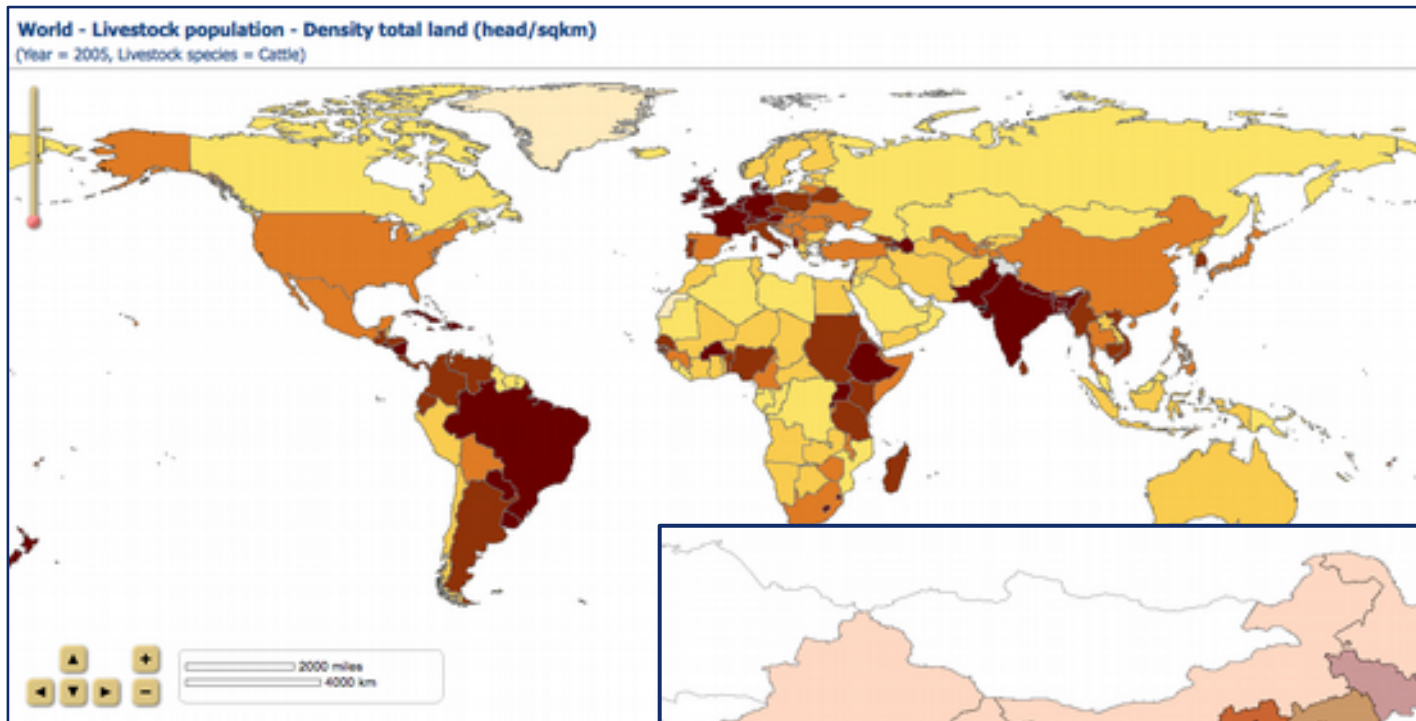
- CIESIN
- GeoNetwork
- GADM
- Worldclim

# Autres sources de données: quelques exemples

- <https://gadm.org/>: divisions administratives
- [www.ciesin.org](http://www.ciesin.org): données environnementales et sociales variées
- <http://www.fao.org/geonetwork>: données GIS
- [www.worldclim.org](http://www.worldclim.org): données climatiques



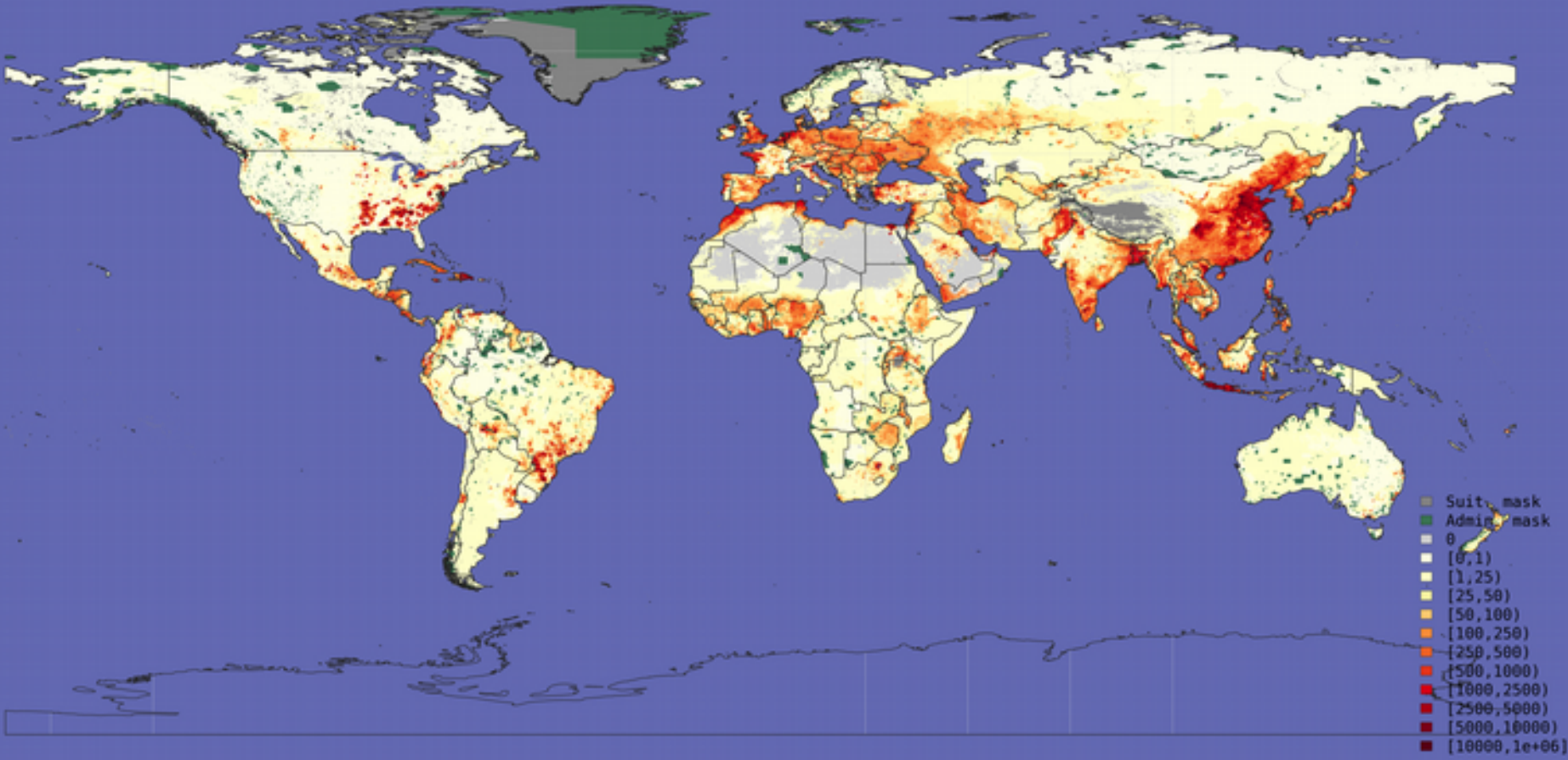
# Gridded livestock of the World (accessible via geonetwork)





# GLW 3.0 runs: chicken

Predicted data: polygon-corrected (weighted) predictions

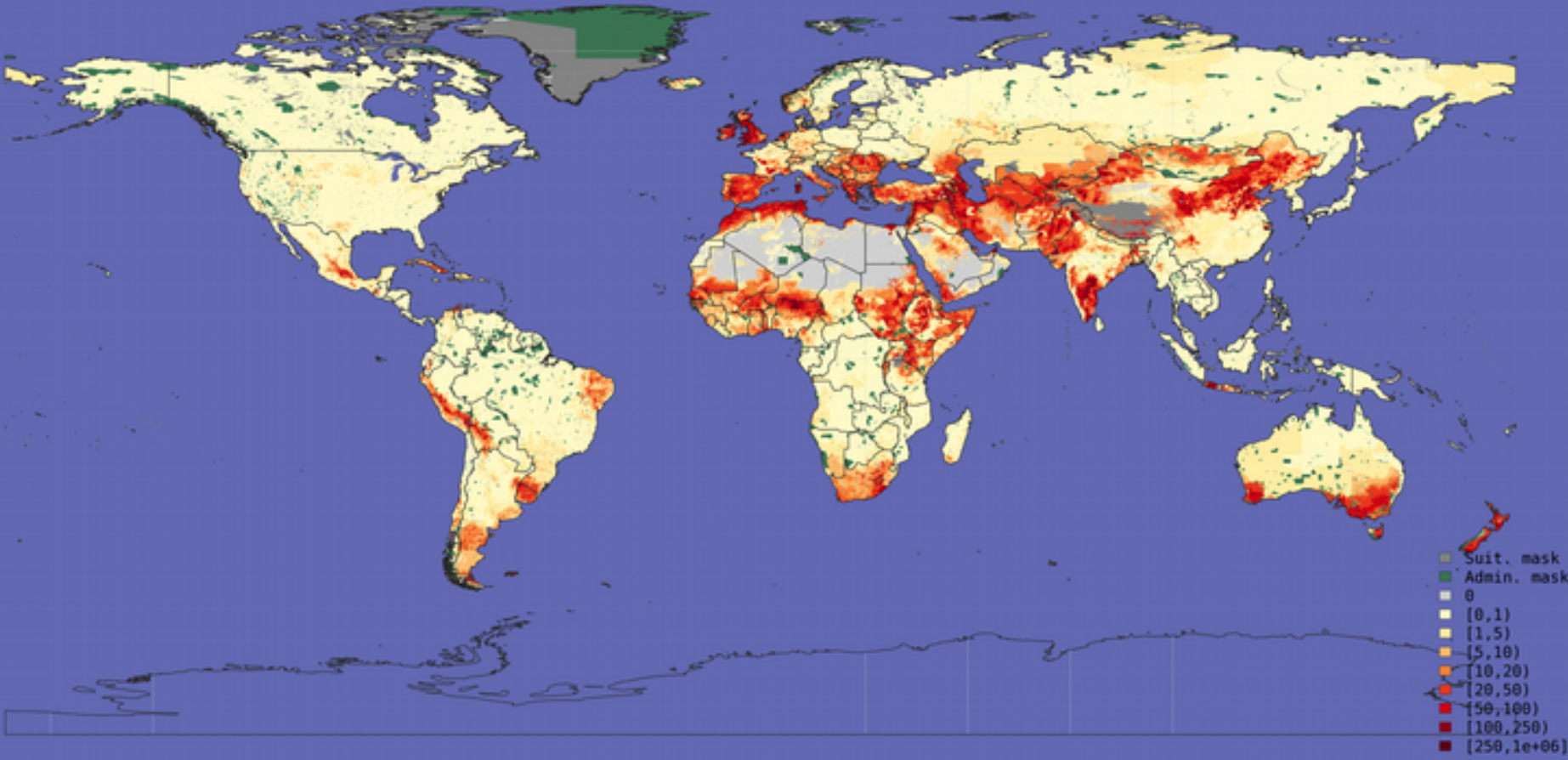






# GLW 3.0 runs: sheep

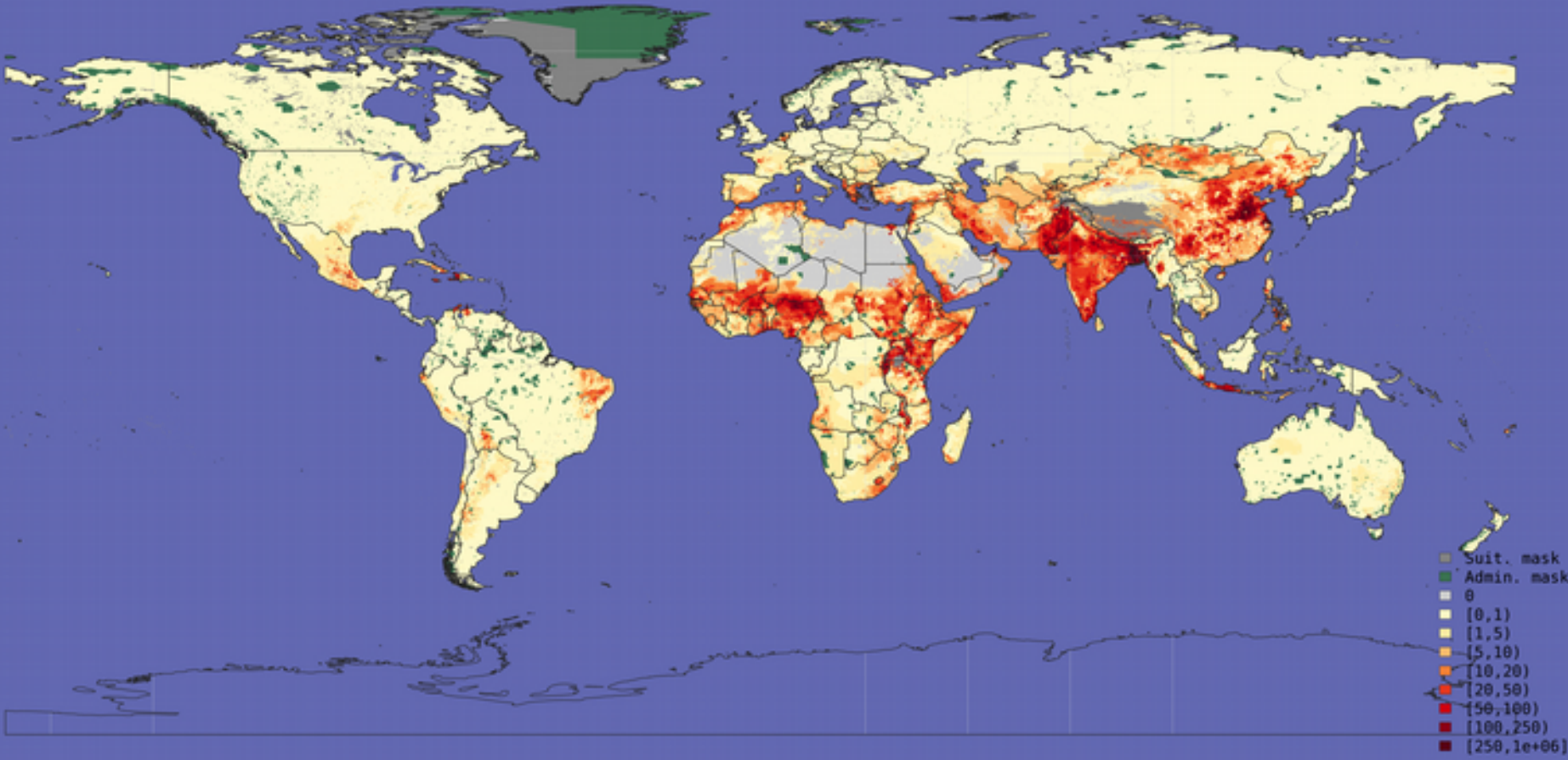
Predicted data: polygon-corrected (weighted) predictions





# GLW 3.0 runs: goats

Predicted data: polygon-corrected (weighted) predictions



Choisissez 1 ou 2 base de données qui vous intéresse et essayez de répondre aux questions suivantes:

- Thématique des données (climat, sols, population, etc.)?
- Organisme(s) producteur(s) (nom, nationalité, type d'organisme)?
- Méthode de collecte (terrain, compilation, experts, etc.)?
- Type de données (raster, vecteur, tabulaire, etc.)?
- Région(s) couverte(s)?
- Échelle / résolution spatiale?
- Période couverte?
- Méta-données disponibles?
- Pouvez-vous télécharger les données? Visualiser?