Accessibility modelling in R
CGIAR-CSI Annual meeting
Amsterdam, March 2020

Access defined and its role in the United Nations SDGs

Andy Nelson &
Rolf de By
ITC
University of Twente

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

BHUTAN THIMPHU OKTOBER 28 - NOVEMBER 8 2019

Webinar and prerequisites

Webinar

- We will present the background/theory for about 20 mins and then spend the rest of the webinar going through the R script to run the accessibility analysis on Bhutan.
- Questions welcome throughout using the BigMarker Chat function.
- Our script works, but we welcome any improvements to the R script from R gurus – such as switching to terra and sf and streamlining GIS operations. (r.a.deby@utwente.nl a.nelson@utwente.nl)
- Any interest from the raster/terra/gdistance package maintainers to increase the efficiency of underlying functions for cost accumulation/cost catchment mapping?

Prerequisites

1. Install an up-to-date version of R and RStudio
2. Install these packages and their dependencies – raster, rgdal, gdistance – install.packages("raster") etc.
3. Download and unzip (or clone) the course material (data, R script and PDF) - https://github.com/gip-itc-nl/accessibility_course_bhutan
4. Download one large SRTM dataset and copy into the /inputs/ directory from point (3) http://ftp.itc.nl/incoming/accessibility/bhutan_dem_srtm.tif

- Bullet 3 is important! We have updated the script this morning …

Leave no one behind
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

BHUTAN THIMPHU OKTOBER 28 - NOVEMBER 8 2019

The *access* notion

Something "wanted, needed, required."

Ability to reach, enter or use something

Subject : a natural person or group of people, all people

Resource : when put in the hands of the subject, improves the subject's life quality

Access to a resource may be problematic and hit barriers. Access is not normally a constant and may vary over time.

Discuss time variability for a resource access situation.

Accessibility is one of a triplet

The other two requirements being:

- *Availability*
- *Utility*

All three need to be met for proper functioning of a stable system.

Severe limitations to access may lead to limitations in welfare of those affected, and the so-called *spatial poverty trap.*

Such systems are riddled with access problems because of all the logistics involved.

Accessibility mapping in R
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

A plethora of resource types

Leave no one behind
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

The basics
- Home
- Food
- Health
- Job or job tool ownership
- Opportunity to develop

Underlying principles
- Equality
- Dignity
- Freedom
- Civil rights

Requires access to
- Land
- Economy
- Health care
- Education
- Legal system
- Information

Resources can be
- Physical
- Informational
- Social

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

The *access to resource* pie chart

Accessibility mapping in R
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

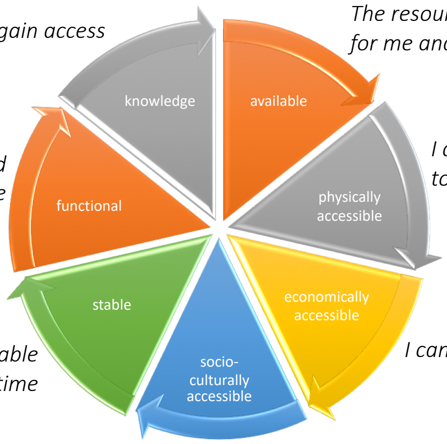
Example: think "safe drinking water" as the resource

I know how to gain access

The resource has adequate volume for me and others

The resource can be used for my purpose

I can physically reach the resource to obtain/consume it



The resource is reliable over time

I can afford the access

Society (the community, the law, religion) allows me to gain access

How can access to drinking water be constrained?

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Mark well: all constraints are relevant

Any of the existing constraints may prevent access to become reality.

All must be addressed.

Planning, thus, is needed for an integrated approach to improving access.

Constraints will be different in time, in space, in context.

Example: access to healthcare

- Not a single problem but an array of problems
- Healthcare has supply chain logistics
- Different needs for different patients (diseases, cases)

Access to

- A doctor consult
- A medicine or vaccine
- A hospital (bed)
- A treatment
- Laboratory equipment
- An insurance policy
- An ambulance

Example: access to healthcare (hospital)

Know How to reach, enter, connect with medical expertise

Volume Does hospital have capacity?

Reach Can we get there (in time)?

Money Is the service payable?

Social Does my community accept my hospital visit? Is it legal?

Open at all times Can I rely on the hospital being open?

Utility Can I expect the hospital to be helpful?

Discuss generally how these problems can be addressed.

This course: mapping *physical* accessibility

Physical accessibility is a *measure of effort* required to *reach a location*.

At regional scale, this often means *travel distance* or *travel time*.

When multiple modes of transport are available, *travel time* is a less biased, thus more useful, metric.

Also, in some situations, there may be options to *bring the resource to the subject*, instead of expecting the subjects to travel to the resource. This often becomes a mere economic equation to solve.

Discuss cases where reverse traffic is or can be the norm.

The *cost* notion

Low physical accessibility normally means that there is some form of high cost. Cost can be measured in different ways.

- Distance is an easy to use but naïve first form.
- Travel time is already better because it also measures time lost for other opportunities (*opportunity cost*). And it may represent to some extent also economic cost (transport tickets or fuel costs).

High access costs usually mean naturally lower availability and lower quality of the resources under study. For one high economic travel costs lower availability because one cannot always afford.

Modes of transport

Depends on study case and study area. Must be known for study to be sensible.

Foot, horse/donkey/elephant/camel, bike/motorbike, taxi(van), car, bus, train or boat. Plane not usually included.

Travel may (need to) be multimodal.
Walk-bus-walk is a common pattern.

Route choices

People may choose their route on the basis of:

- Economic cost
- Used vehicle; some roads may not be fit for the vehicle
- Weather and road conditions
- Time of travel
- Involved travel risk
- Urgency for the resource

In modelling route choice, *we usually assume that people will take the time-fastest route*.

Travel (transport) infrastructure

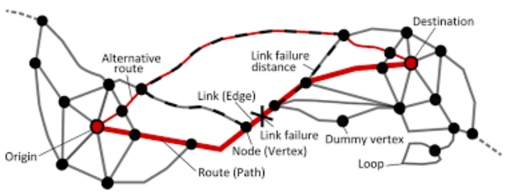
Accessibility mapping in R
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

Since we aim to understand time-efficient travel, we should know about

- Available network for mechanized travel and attainable travel speeds on it, incl. potential travel delay factors

And, where no network exists

- The terrain and its characteristics for travel on foot



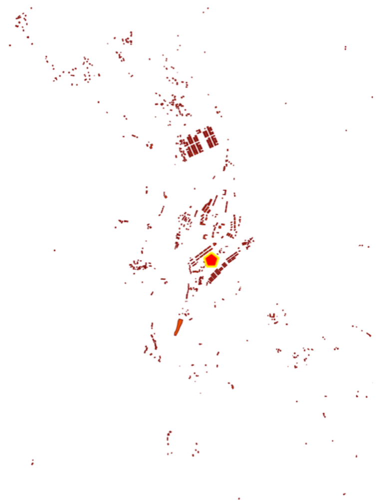
UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Route start and end points

For our type of study

- Start points are usually people's home location
- End points are the locations of where the resource can be found

Need to know both as well as possible.



The 2030 Agenda for Sustainable Development

Adopted by all United Nations Member States in 2015, provides a *shared blueprint for peace and prosperity for people and the planet*, now and into the future.

At its heart are the 17 Sustainable Development Goals (SDGs), which are an urgent call for action by all countries – developed and developing – in *a global partnership*.

They recognize that ending poverty and other deprivations must go hand-in-hand with *strategies that improve health and education, reduce inequality, and spur economic growth* – all while tackling climate change and working to preserve our oceans and forests.

Accessibility mapping in R
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

End poverty in all its forms everywhere

1 NO POVERTY

1.4 By 2030, ensure that all men and women, in particular the poor and the vulnerable, have equal rights to economic resources, as well as **access to basic services**, **ownership and control over land and other forms of property, inheritance, natural resources, appropriate new technology and financial services, including microfinance**

Leave no one behind
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

End hunger, achieve food security and improved nutrition and promote sustainable agriculture

2 ZERO HUNGER

2.1 By 2030, end hunger and ensure **access** by all people, in particular the poor and people in vulnerable situations, including infants, **to safe, nutritious and sufficient food** all year round
2.3 By 2030, double the agricultural productivity and incomes of small-scale food producers, in particular women, indigenous peoples, family farmers, pastoralists and fishers, including through secure and equal **access to land, other productive resources and inputs, knowledge, financial services, markets and opportunities** for value addition and non-farm employment

Leave no one behind
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Et cetera!!

Leave no one behind
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

BHUTAN THIMPHU OKTOBER 28 - NOVEMBER 8 2019

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Accessibility modelling in R
CGIAR-CSI Annual meeting
Amsterdam, March 2020

Software packages used

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

BHUTAN THIMPHU OKTOBER 28 - NOVEMBER 8 2019

To be installed

- GIS to
  - Load, validate, update data
  - Display output
  - Visualize for dissemination
- Language and environment for
  - statistical computing and graphics

Also used

PostgreSQL

Version 4 pgAdmin

- Object-relational database management system
  - It is designed to handle a range of workloads, from single machines to data warehouses or Web services with many concurrent users.
- A client-side data management tool for PostgreSQL databases

Accessibility mapping in R
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

And … Even if time allows …

- iD, the online editor to OpenStreetMap
  - An online opportunity to map from images and contribute to a global community
  - No installation; use online at openstreetmap.org

- Python PL plus spatial data functions
  - A full programming language with powerful spatial libraries

QGIS

- Open Source Geographic Information System
- https://www.qgis.org/en/site/forusers/download.html

- Installation of the software
- Quick walk-through
  - Environment
  - Setup
  - Menu
  - Work-area
  - Tools and added functionality

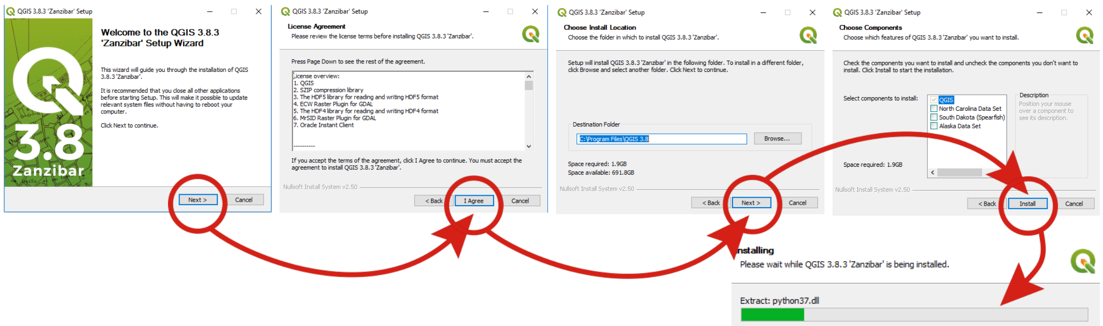
Accessibility mapping in R
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

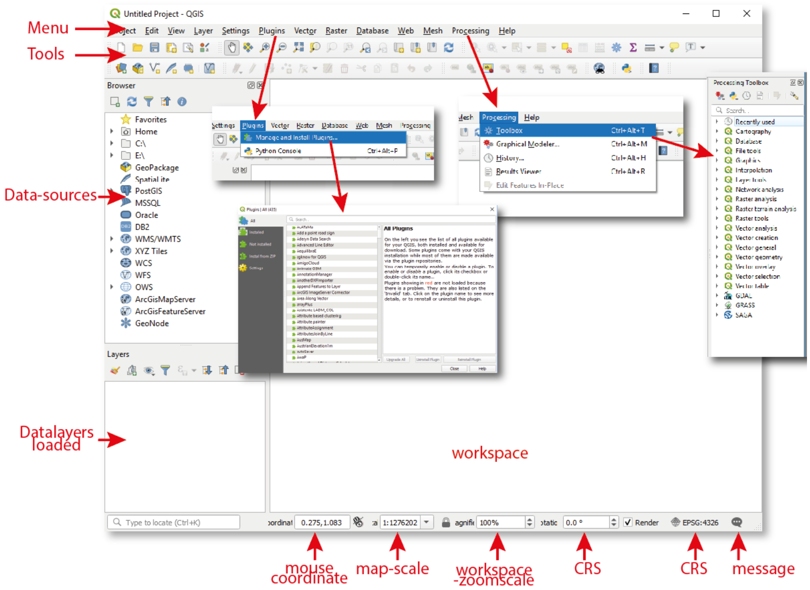
QGIS Installation

QGIS-OSGeo4W-3.8.3-1-Setup-x86_64.exe 21-10-2019 16:17 Application 464.954 KB

USB-stick/Software/QGIS-OSGeo4W-3.8.3-1-Setup-x86_64.exe



The QGIS environment



QGIS workflow, a demo

1. Set project properties
2. Load layer(s)
3. Set layer order
4. Processing, analysis
5. Visualize, symbolize
6. Output
   1. Size, orientation
   2. Insert map
   3. Add title, legend, scale, grid/graticule
   4. Add marginal info
   5. Export to appropriate file format and resolution

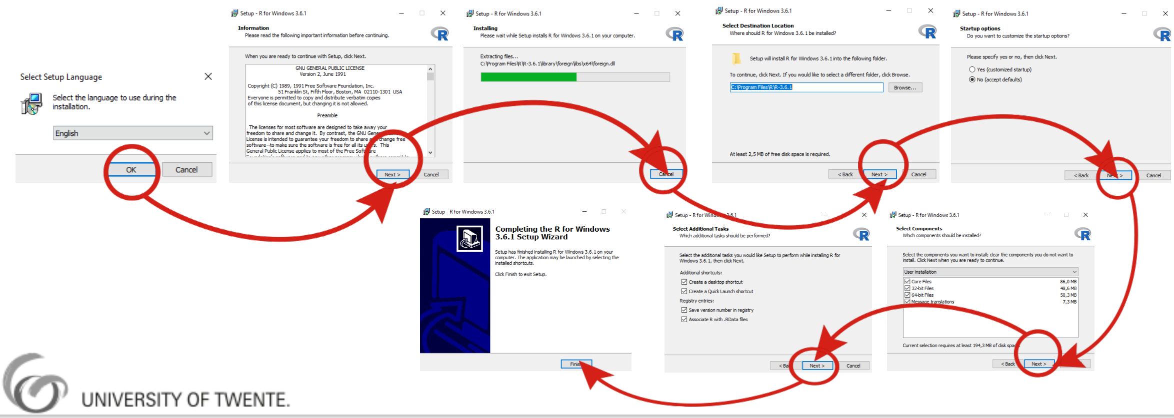
R

- The R Project for Statistical Computing
- https://cran.r-project.org/mirrors.html
  - https://cran.stat.nus.edu.sg/
- Installation of the software
- Quick walk-through
  - Environment
  - Setup
  - Menu
  - Work-area

R Installation process

R-3.6.1-win.exe 21-10-2019 16:19 Application 82.697 KB

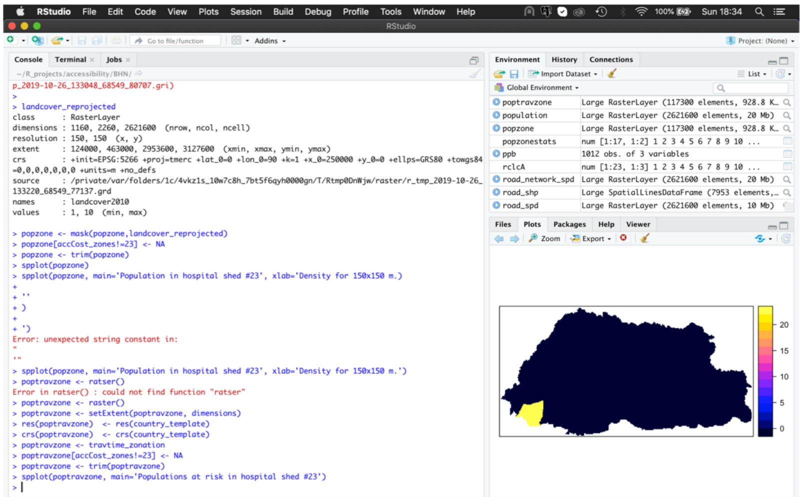
USB-stick/Software/R-3.6.1.exe



The R environment

Menu

Console & OS Terminal



Variables & Command history

Resources, Plots & Help

Accessibility mapping in R
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in todays Society
UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Some R libraries needed

Start up Rstudio, and type in the following for one-time installation of some specific packages:

```
> chooseCRANmirror(ind=1)
> install.packages("rgdal", dependencies=NA)
> install.packages("raster", dependencies=NA)
> install.packages("gdistance", dependencies=NA)
```

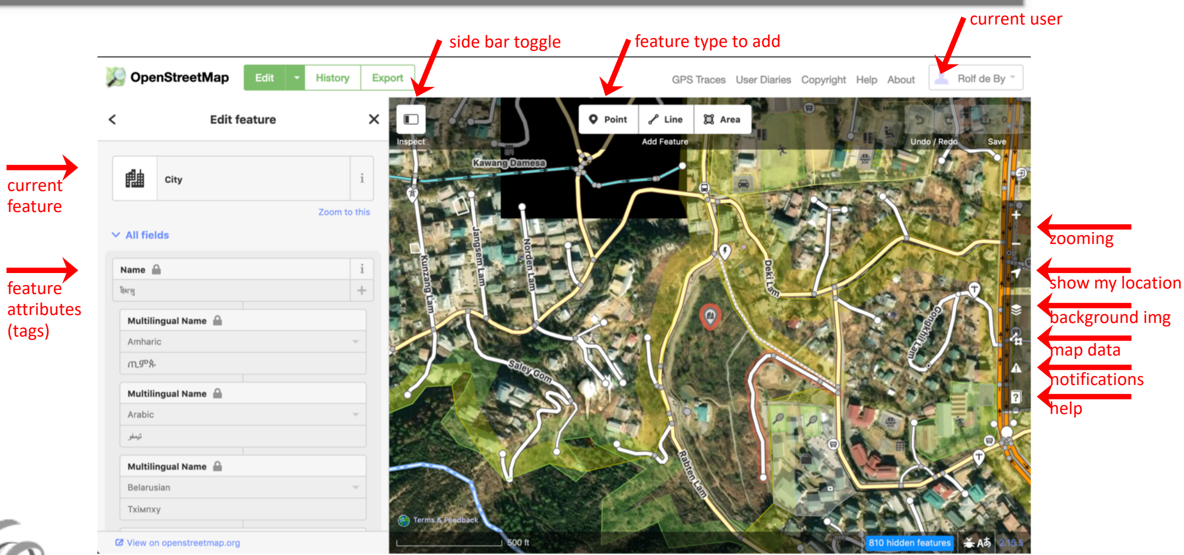
Accessibility mapping in R

Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

UNIVERSITY OF TWENTE.

ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

The iD editor to OpenStreetMap



Accessibility mapping in R

Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

UNIVERSITY OF TWENTE.

ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Accessibility modelling in R
CGIAR-CSI Annual meeting
Amsterdam, March 2020

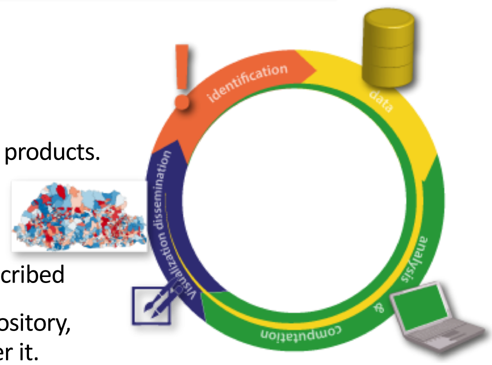
Computing physical accessibility

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

BHUTAN THIMPHU OKTOBER 28 - NOVEMBER 8 2019

Generic workflow

- **Identification module:**
  - defines the problem in detail,
  - formulates the expected result which is the basis for
  - the identification of required data as well as
  - the framework for the design and creation of the final products.
- **Data module:**
  - list the data requirements for the project.
  - Find local data and find data online
  - assessment of data usability and quality, based on described requirements.
  - Collect the data into a purposefully designed data repository,
  - curate the data and execute selection and analysis over it.
- **Analysis & computation module:**
  - find answers to the formulated questions
- **Dissemination module:**
  - Graphs and maps



Accessibility mapping in R
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Physical accessibility

Determined per resource type:
GP, clinic, hospital, school type 1 .. N, water point, bank branch, ATM, …

Determined for each location of interest: where subjects are or live

Normally, computed by identifying a ***least cost travel path*** *for each subject location:*

- determine a path to each resource
- identify the path that has lowest cost
- where it goes is that location's preferred resource location (say, that of some hospital)

Example

- A map graph that abstractly shows one subject location and many resource locations, with paths and costs.

Data needs to have

Which are the locations where subjects are/live?

Which are the locations of the resource type that we address?

Which are the travel modes that subjects may use?
(Or: by which travel mode will resource be brought to the subjects?
Can of worms opens here.)

Per travel mode, how can we determine travel cost? This requires specific data.

If multiple travel modes may be in use, how do we prioritize these? (Ie, how does the subject prioritize her/his choices?)

Common assumptions

Many subjects & subject locations: representation by raster

Smaller set of resources: possibly just vectors

Travel may be

- over an infrastructure network (roads, trains, river boats)
- across terrain (walking, skiing, swimming, lake/sea boats)

All require cost characterizations per time or distance unit.

Will ignore (for now) time dependencies

Accessibility modelling in R
CGIAR-CSI Annual meeting
Amsterdam, March 2020

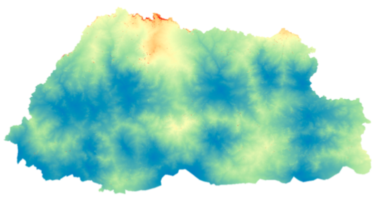
Computational workflow

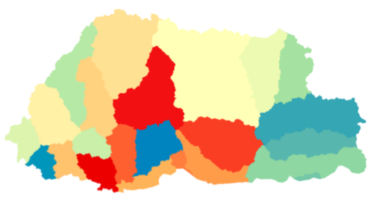
And the nitty-gritty details

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

BHUTAN THIMPHU OKTOBER 28-NOVEMBER 8 2019

Accessibility mapping in R
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society





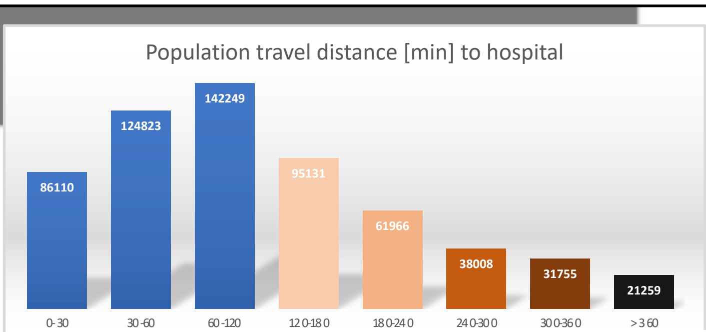
What we aim to achieve

- Understand the location-specific **travel time** to some class of facility
- Understand the location-specific facility that is closest**: the facility-sheds**

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

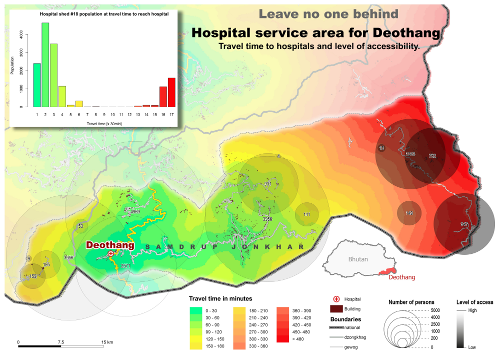
Accessibility mapping in R

Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society



And also …

Which part of the population lives **at which travel distance**?



UNIVERSITY OF TWENTE.

ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Leave no one behind

Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

Workflow data needs

- Target facilities
- Human population
- Impact factors in human travel
  - Traffic network infrastructure
  - Terrain specifics (where no infra exists)

- E.g. Hospital locations
- Where do (how many) people live?
- Roads, railroads, paths
- Land cover
- Elevation & slope
- Study area delineation

UNIVERSITY OF TWENTE.

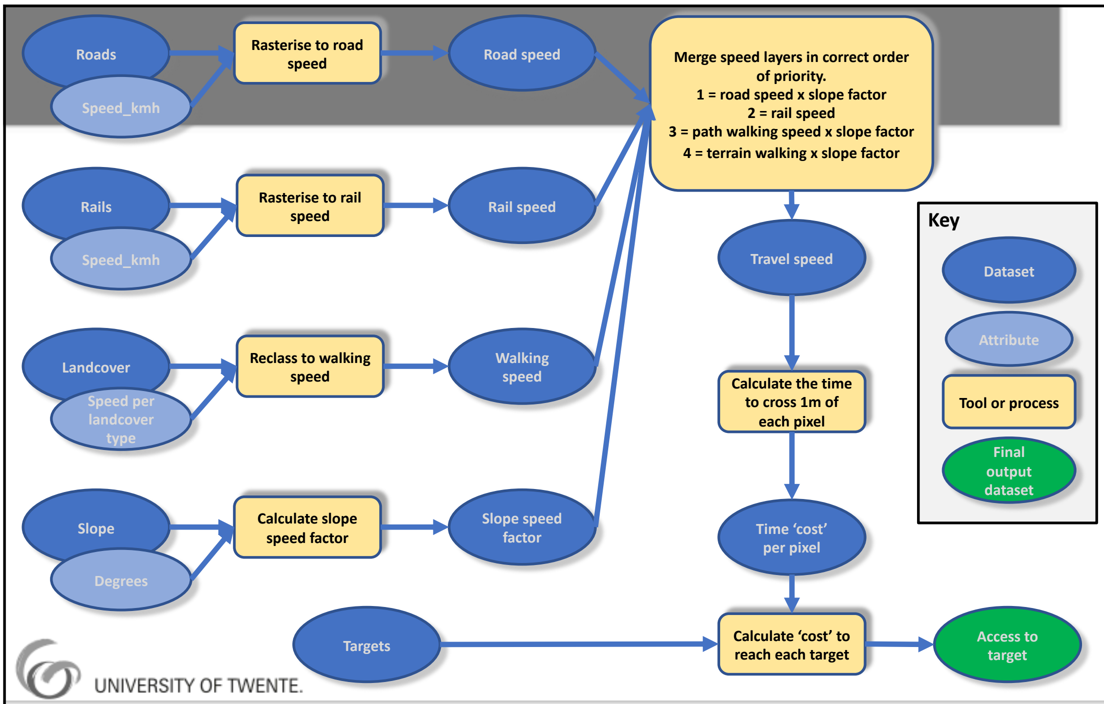
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

The workflow schematically

- Mixed travel modes means mix of vector and raster data
- We need to make them fit
- We need to prioritize one mode over another
- Eventually, this will be a raster-based analysis
- A cell value will denote *"the cost of traversing the cell's area"*

Leave no one behind
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION



Leave no one behind
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Refere

UNIVERSITY OF TWENTE.

ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Leave no one behind

Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

References

1 Global map of travel time for 2015 to cities of 50,000 people or more

Paper - A global map of travel time to cities to assess inequalities in accessibility in 2015 https://www.nature.com/articles/nature25181

Data & code – Accessibility to cities https://malariaatlas.org/research-project/accessibility_to_cities/

Google Earth Engine layers
https://explorer.earthengine.google.com/#detail/Oxford%2FMAP%2Faccessibility_to_cities_2015_v1_0 &
https://explorer.earthengine.google.com/#detail/Oxford%2FMAP%2Ffriction_surface_2015_v1_0

2 Global maps of travel time for 2015 to cities and ports of different sizes

Paper - A suite of global accessibility indicators https://www.nature.com/articles/s41597-019-0265-5

Data & code - Travel time to cities and ports in the year 2015 https://doi.org/10.6084/m9.figshare.7638134.v3

3 The excellent gdistance package. Gdistance and Google Earth Engine are the ONLY platforms for correct travel time calculations on a "sphere"

Package - Distances and Routes on Geographical Grids https://cran.r-project.org/package=gdistance

Paper - Distances and Routes on Geographical Grids https://doi.org/10.18637/jss.v076i13

Note that you cannot compare the year 2000 (https://forobs.jrc.ec.europa.eu/products/gam/) and year 2015 travel time maps (refs #1 and #2); the methods and data inputs are completely different.

UNIVERSITY OF TWENTE.

ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Accessibility mapping in R

Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

The main computing steps, 1

1. Determine the common denominators for your various data sets (crs, resolution, delineation, date of publication)
2. Prepare input data sets
   1. Slope from elevation
   2. Rasters from vector data
   3. Study area mask
   4. Reproject to common crs
3. Define maximum travel speeds for types of road, land cover, slope and elevation
4. Compute maxspeed rasters for all input categories
5. Compute terrain walking speed accounting for slope
6. Compute networked speed raster accounting for slope

The main computing steps, 2

7. Merge maxspeed rasters and prioritize higher speeds over lower.
8. Convert [*km/h*] to [*min/m*], thus creating the **friction surface**
9. For each target facility (out of *N*), determine travel time from each location to it, obtaining *N* travel time rasters. [Special function!]
10. Determine the minimum travel time per location by comparing between travel times to all facilities.
    OUTPUT 1
11. Determine facility closest per location
    OUTPUT 2
12. Chart total population per travel time bucket
    OUTPUT 3

The main computing steps, 3

13. Determine similar results per hospital service area, depending on policy needs

Additional uses of such workflows

- Validate your input data: inexplicable patterns may point at wrong inputs
  - A gender balance map
  - A youth/matures balance map
  - Map the #people per building
  - Map extreme travel distances
- What-if scenarios
  - Constructing extra facilities: what if we build two new hospitals? And where should they go?
  - Natural disaster scenarios: what if a road gets blocked by landslides?

Common denominators

Extent
Territory of Bhutan

Due to computation constraints on our hardware:

- 150 x 150 m grid
- crs:5266

Elevation data
SRTM, 1 arcsec grid in crs:4326

Land cover
ICIMOD 2010, 30m UTM grid at crs:32645

All **Bhutanese data**
- Infrastructure: roads, paths, houses
- Administrative boundaries
- Census records
  as vector, crs:5266

OpenStreetMap
vector data; crs:4326

Leave no one behind
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

BHUTAN THIMPHU OKTOBER 28 - NOVEMBER 8 2019

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Accessibility modelling in R
CGIAR-CSI Annual meeting
Amsterdam, March 2020

The workflow details

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

BHUTAN THIMPHU OKTOBER 28 - NOVEMBER 8 2019

Step 1

Determine the common denominators for your various data sets (crs, resolution, delineation, date of publication)

See above

The script code has some needed, but ugly, initialization steps to
- Make robust for different operating systems
- Have one or two useful functions
- Define some constants and templates

Step 2

Prepare input data sets

1. *Slope from elevation*
2. *Rasters from vector data*
3. *Study area mask*
4. *Reproject to common crs*

Key part of the code:

```
# reproject landcover and elevation rasters
landcover_reprojected <-
  projectRaster(landcover, country_template,
    crs = '+init=EPSG:5266', res = 150)
slope_reprojected <-
  projectRaster(bhutan_dem_srtm_slope, country_template,
    crs = '+init=EPSG:5266', res = 150)
```

Step 3

Define maximum travel speeds for types of road, land cover, slope and elevation

This is mostly input data preparation

For instance, table *osm_roads* and table *landcover_class_speeds.csv:*

id	from	to	becomes	description
0	0	1	0.5	Snow and glacier
1	1	2	3	Shrub
2	2	3	4	Grassland
3	3	4	4	Barren area
4	4	5	3	Mixed forest
5	5	6	3	Conifer forest
6	6	7	3	Broadleaved forest
7	7	8	4	Agriculture
8	8	9	5	Urban
9	9	10	0.5	Water body

Step 4

Compute maxspeed rasters for all input categories

We show how landcover affects walking the terrain

Key part of the code:

```
# See example data file, it must be a 5-colum
#  (id,from,to,becomes,description) file
lcs = read.table("inputs/landcover_class_speeds.csv",
                 colClasses=c("integer","integer","integer","numeric",
                              "character"),
                 header=TRUE, sep=",", row.names=1)

# apply the reclassification to get travel speeds per lc class
landcover_speed <
                  reclassify(landcover_reprojected,lcs,right=FALSE)
```

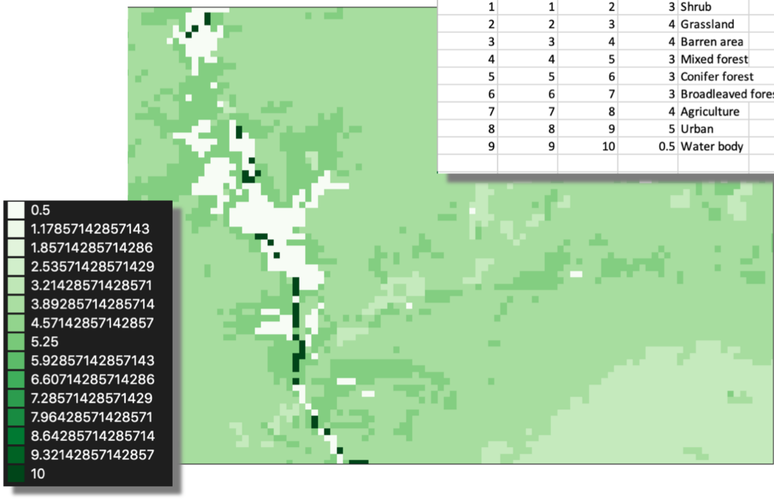
UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION
Leave no one behind
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society
BHUTAN THIMPHU OKTOBER 28 - NOVEMBER 8 2019

Step 4 results

The data file

And some lc example

id	from	to	becomes	description
0	0	1	0.5	Snow and glacier
1	1	2	3	Shrub
2	2	3	4	Grassland
3	3	4	4	Barren area
4	4	5	3	Mixed forest
5	5	6	3	Conifer forest
6	6	7	3	Broadleaved forest
7	7	8	4	Agriculture
8	8	9	5	Urban
9	9	10	0.5	Water body



Leave no one behind
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society
BHUTAN THIMPHU OCTOBER 28 - NOVEMBER 8 2019

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Step 5

Compute terrain walking speed accounting for slope

$$W = 6e^{-3.5\left|\frac{dh}{dx}+0.05\right|}$$

$$\frac{dh}{dx} = S = \tan\theta$$

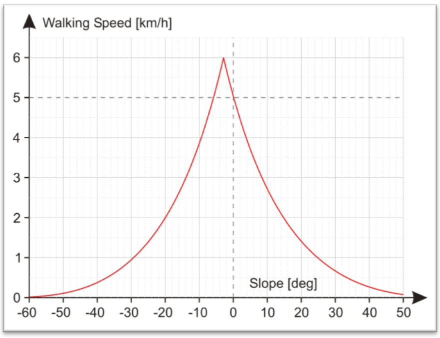
where

W = walking velocity [km/h][2]
dh = elevation difference,
dx = distance,
S = slope,
θ = angle of slope (inclination).

Bit of theory

Steep terrain slows down our ability to traverse the terrain by foot.
Tobler (1993) provides an equation to adjust walking speed by slope.

We can apply this equation to the slope map to compute a walking velocity. If we then divide that by a base walking speed of 5km/hr, we get a slope factor that we can apply to each landcover class speed.



Leave no one behind
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society
BHUTAN THIMPHU OCTOBER 28 - NOVEMBER 8 2019

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Step 5

Compute terrain walking speed accounting for slope

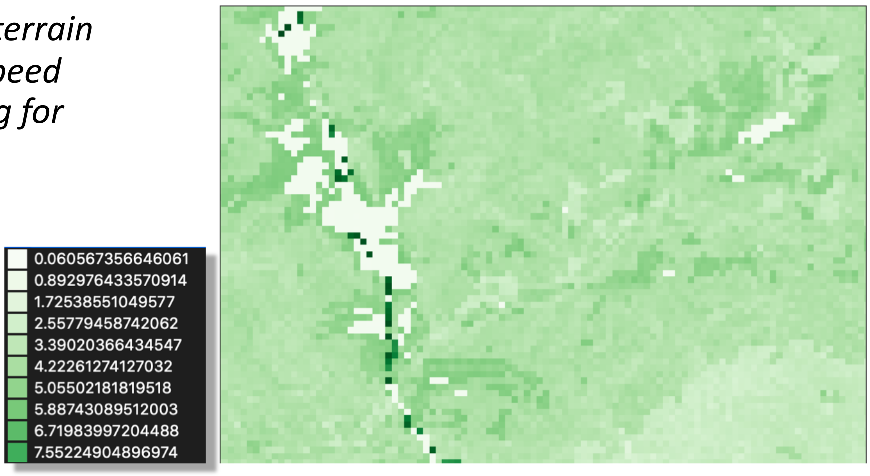
Key part of the code:

```
# slope for walking;
# Tobler's walking speed is given by W = 6e^-3.5|tan(slope)+0.05|
# 6 km/h is maximum speed at a slight downward angle;
# 0.05 rad is the optimal angle for best horizontal speed;
# -3.5 is a factor of decay of maximally attainable speed as
# slope angle changes.
# But we use -0.4 as Bhutan's slopes are very steep,
# leading to unrealistic decay under -3.4
slp_walk <-  6 * exp(-0.4 * abs(tan(slope_reprojected*pi/180) + 0.05))

# divide by base speed of 5km/h to get speed factor for
# slope-adjusted walking speeds over landcover
terrain_walk_spd <- landcover_speed * slp_walk/5.0
```


Step 5 results

Compute terrain walking speed accounting for slope [km/hr]



UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION
Leave no one behind
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

Step 6

Compute networked speed raster accounting for slope

This means all networked travel modes, but ***we show car travel only***

Key part of the code:

Again Tobler's equation, but now experimentally defined for car travel over slopes.

```
# slope for car driving
# Experimental Tobler's car speed is given by
# W = 50e^-0.4|tan(slope)+0.12|
# Original -2.4 gave far too much decay.
slp_car <-  50 * exp(-0.4 * abs(
                     tan(slope_reprojected*pi/180) + 0.12))
road_shp <- readOGR(filepath("inputs/osm_roads.shp"),
                     integer64="allow.loss")
# transform to crs of the country raster template
road_shp <- spTransform(road_shp, crs(country_template))
```

UNIVERSITY OF TWENTE.

ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Leave no one behind

Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

Step 6

Continued …

Key part of the code:

```
# assign raster cell values using the maxspeed field for roads
road_spd <- rasterize(x=road_shp, y=road_spd,
                     field="maxspeed", fun=max)

# if slope needs to be in play for road and footpaths speeds,
# here is where that happens
sloped_road_spd <- road_spd * slp_car / 50.0
```

UNIVERSITY OF TWENTE.

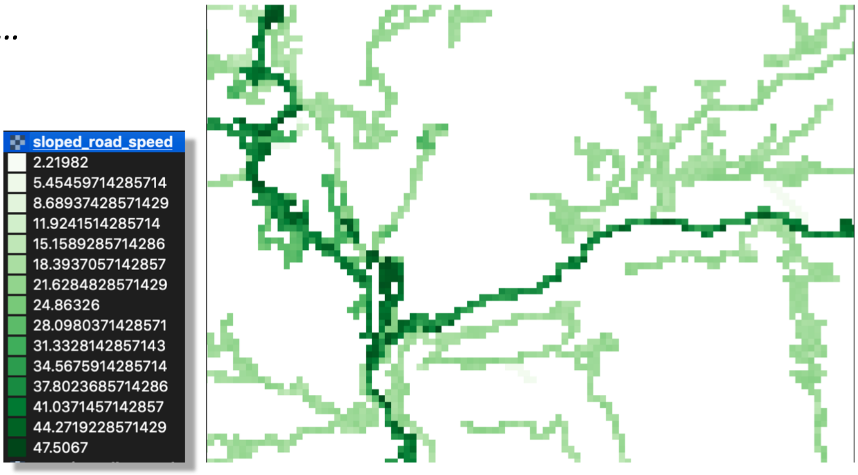
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Leave no one behind

Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

Step 6 results

Continued …



Step 7

Merge maxspeed rasters and prioritize higher speeds over lower

Key part of the code:

```
# merging the various (road, footpath, ...) *network*
# speed rasters; ensure that you prioritize properly
road_network_spd <- merge(sloped_road_spd
                          sloped_footpath_spd)
# Merge the speed components; road takes priority over rail
# over ship travel, over walking the terrain ...
merged_spd <- merge(road_network_spd, terrain_walk_spd)
```

Leave no one behind
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Step 7

*The raster manual on **merge()***

126 *merge*

merge *Merge Raster* objects*

Description

Merge Raster* objects to form a new Raster object with a larger spatial extent. If objects overlap, the values get priority in the same order as the arguments, but `NA` values are ignored (except when `overlap=FALSE`). See subs to merge a `Raster*` object and a `data.frame`.

Step 8

*Convert [km/h] to [min/m], thus creating the **friction surface***

Observe the inversion in units

Resistance to conductance

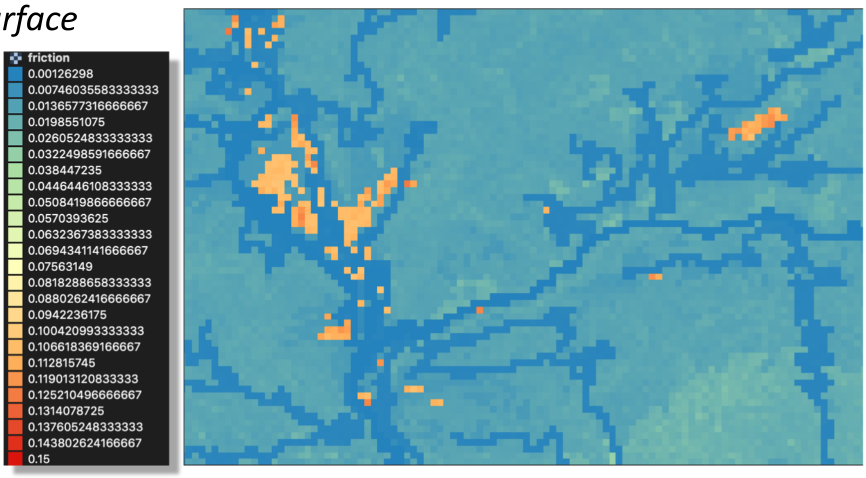
Key part of the code:

```
# convert speed in km per hr to travel time in minutes per metre
# THIS IS THE FRICTION SURFACE
friction <- 1.0 / (merged_spd * 1000 / 60.0 )
```

Step 8 results

Friction surface

[min/m]



Step 9

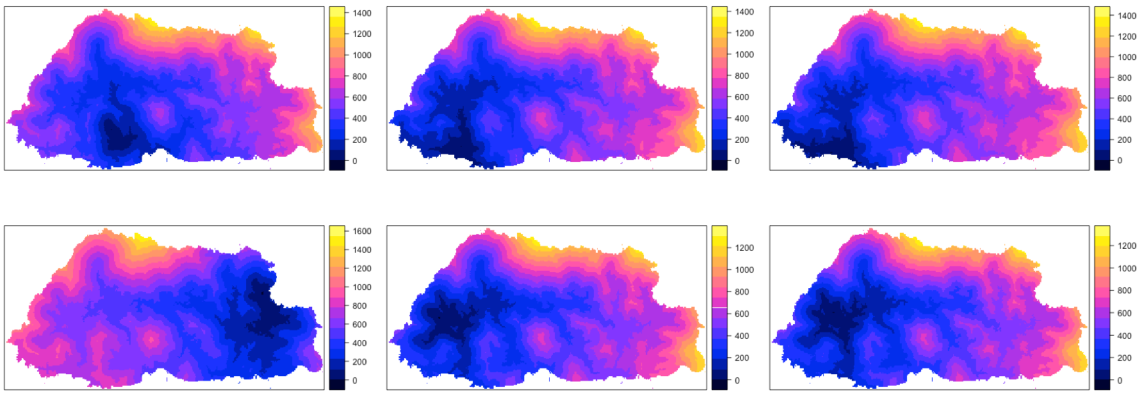
For each target facility (out of N), determine travel time from each location to it, obtaining N travel time rasters.

[min]

Key part of the code:

```
# run accumulative cost for each target
# (targets from shapefile (3 columns, [ID X Y])
for(i in 1:length(targets))
{
  # grab the X and Y coord
  t  <- cbind(targets@coords[i,1],targets@coords[i,2])
  # add accumulated cost to the stack
  accCost_stack <- stack(accCost_stack, accCost(T.GC, t))
}
```

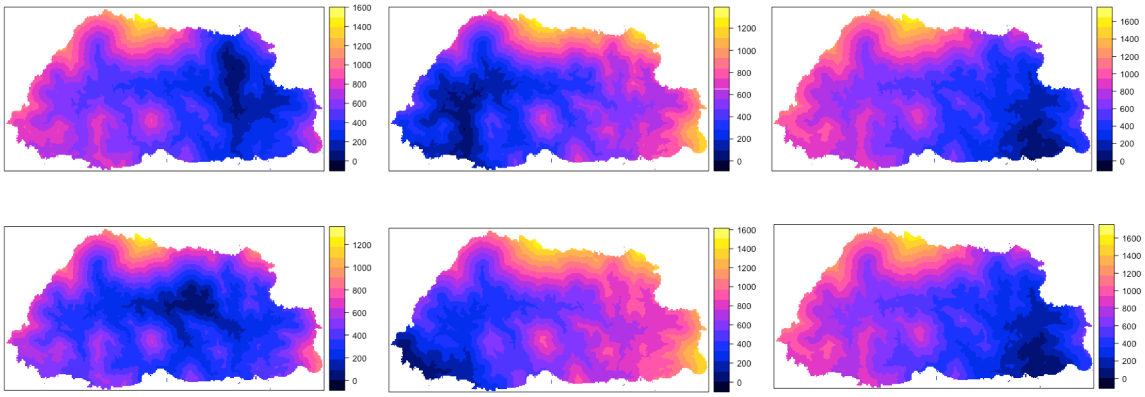
Step 9 results (for 23 hospitals)



Leave no one behind
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society
BHUTAN THIMPHU OCTOBER 28 - NOVEMBER 8 2019

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

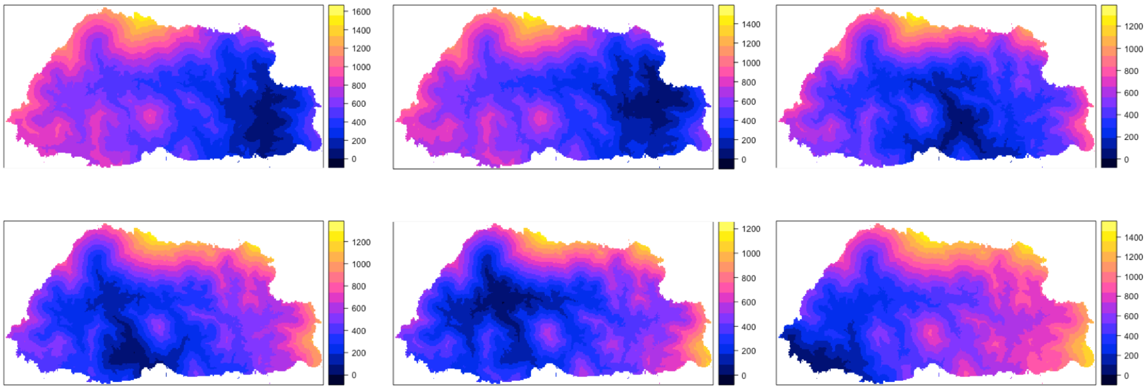
Step 9 results



Leave no one behind
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society
BHUTAN THIMPHU OCTOBER 28 - NOVEMBER 8 2019

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

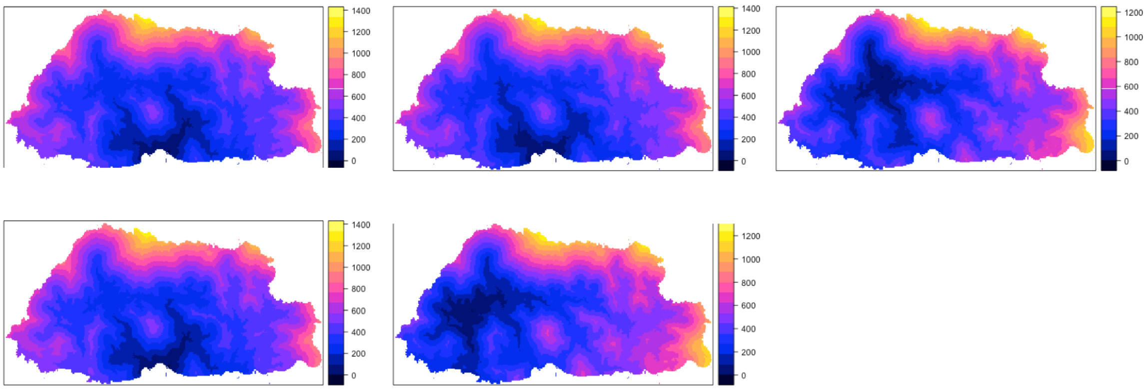
Step 9 results



Leave no one behind
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Step 9 results



Leave no one behind
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Step 10

Determine the minimum travel time per location by comparing between travel times to all facilities.

Key part of the code:

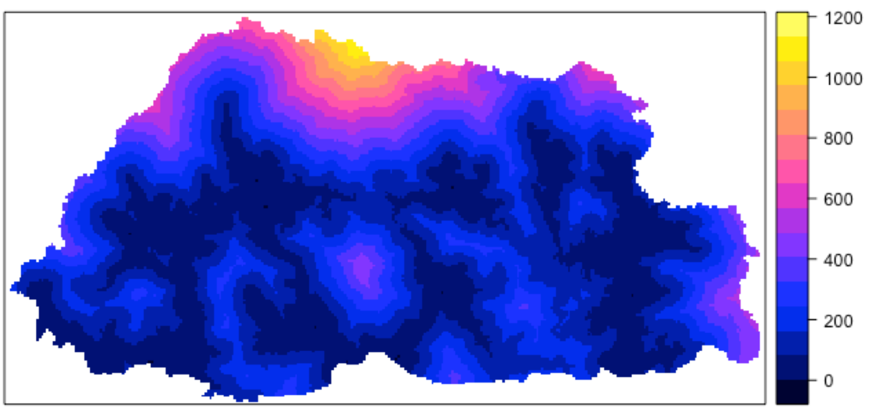
```
# make a new raster based on minimum value
accCost_min <- min(accCost_stack)
```

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Leave no one behind
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

Step 10 results

In [min] of travel time



UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Leave no one behind
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

Step 11

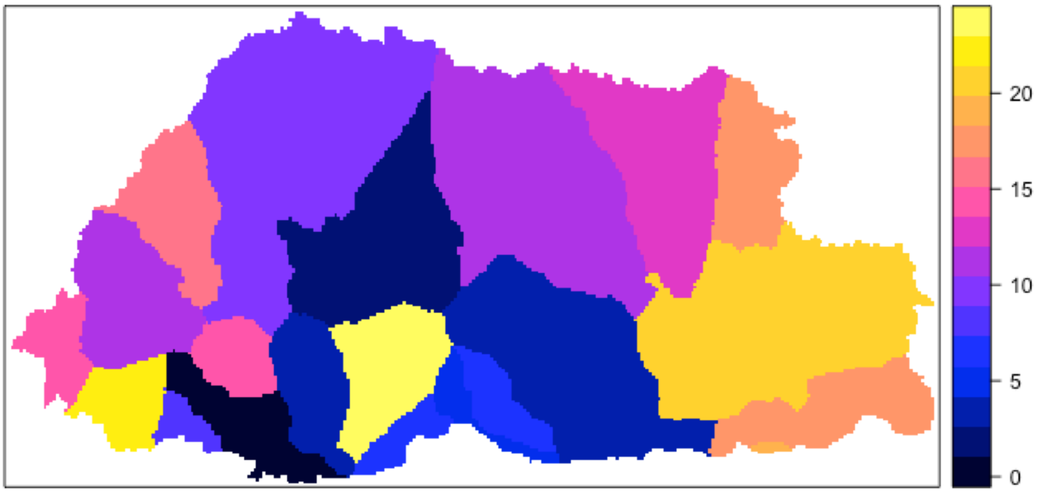
Determine facility closest per location.

Key part of the code:

```
# make a reclass table to assign target IDs to layer IDs.
m <- c(1, 2, targets@data$gid[1])
for(i in 2:length(targets))
{  m <- c(m,i, i+1, targets@data$gid[i])  }
rclcA <- matrix(m, ncol=3, byrow=TRUE)

# apply the reclassification to get allocation zone IDs
# - THIS IS THE COST ALLOCATION ZONE MAP
accCost_zones <- reclassify(accCost_minID,rclcA,right=FALSE)
```


Step 11 result



Step 12

Chart total population per travel time bucket

Key part of the code:

```
# first get the building centroids; we represent buildings
# by centroids (more trivial assignation to raster cells)
building_shp <-
    readOGR(filepath("inputs/buildingcentroid.shp"),
    integer64="allow.loss")
# assign count of buildings to raster
buildings <-
    rasterize(x=building_shp,y=buildings,field="gid",
    fun='count',background=0)
# obtain the census tracts:
censustracts_shp <-
   readOGR(filepath("inputs/censustract.shp"),
   integer64="allow.loss")
```

UNIVERSITY OF TWENTE.

ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Leave no one behind
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

Step 12

Continued …

Key part of the code:

```
censustracts <-
   rasterize(x=censustracts_shp, y=censustracts,
   field='id', fun='last')
ppb = read.table("inputs/censustract_paxperbuilding.csv",
           colClasses=c("integer","integer","integer","numeric"),
           header=TRUE, sep=",", row.names=1)
censustracts_housing <-
         reclassify(censustracts, ppb, right=FALSE)
population <- as.integer(censustracts_housing*buildings)
```

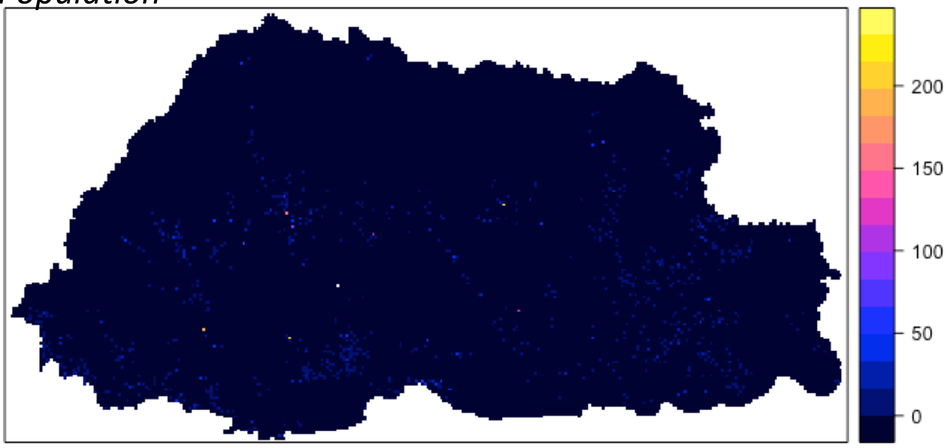
UNIVERSITY OF TWENTE.

ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Leave no one behind
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

Step 12 results midway

Population



Leave no one behind
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Step 12

zone	minutes
0	0
1	30
2	60
3	90
4	120
5	150
6	180
7	210
8	240
9	270
10	300
11	330
12	360
13	390
14	420
15	450
16	480
17	1440

Chart total population per travel time bucket

Continued code:

```
# Check Bhutan popsize
cellStats(population,stat='sum')
# Read the breaks from traveltime_zones csv file:
tt_zones = read.table(
                "inputs/traveltime_zonebreaks.csv",
                colClasses=c("integer","integer"),
                header=TRUE, sep=",", row.names=1)
# Determine the breaks vector
breaks <- tt_zones$minutes
```

Leave no one behind
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Step 12

Chart total population per travel time bucket

Continued code:

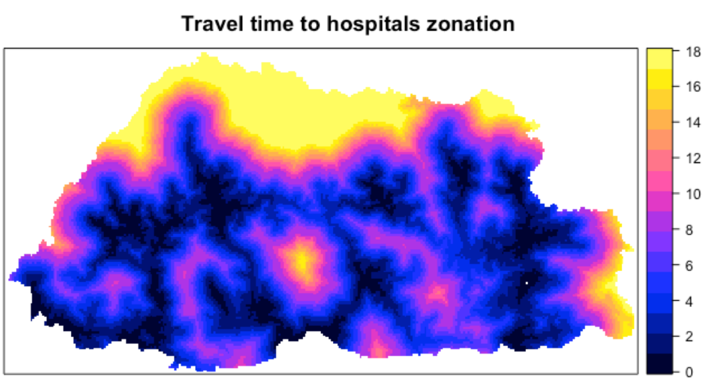
```
# determine traveltime zones spatially by breaks provided
travtime_zonation <- cut(accCost_min,breaks)
spplot(travtime_zonation,
        main='Travel time to hospitals zonation')
```


Step 12 results midway

Leave no one behind
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

Chart total population per travel time bucket (zone)

Mark: numbers are zones, not hours!



E.g., zone 14 represents travel times between 390 and 420 minutes, ie, 6.5 to 7 hrs.

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Step 12

Chart total population per travel time bucket

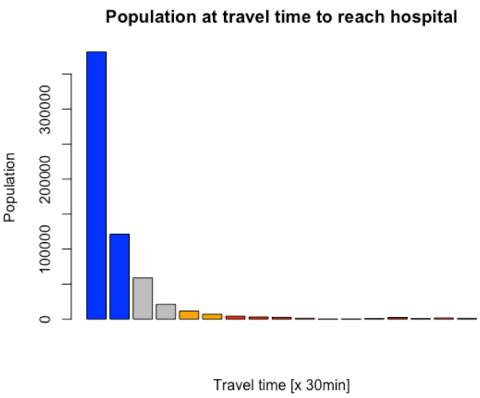
Continued code:

```
# determine traveltime zonal statistics
zonestats <- zonal(population,travtime_zonation,fun='sum')

barplot(t(zonestats)[2,],
        main='Population at travel time to reach hospital',
        names.arg=plotxnames, xlab='Travel time [x 30min]',
        ylab='Population', col=colors)
```


Step 12 results midway

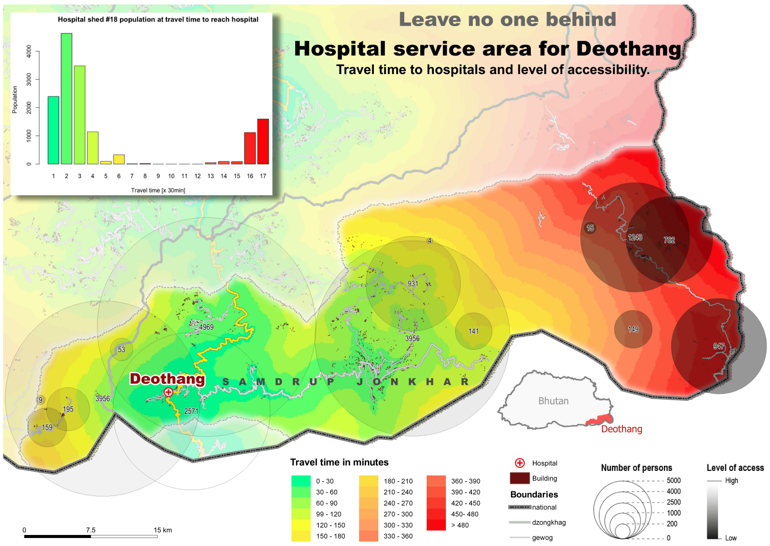
Chart total population per travel time bucket of 30min



Zone 1 holds 381,557 people.

Step 13

Determine similar results per hospital service area and map results



Leave no one behind
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Accessibility modelling in R
CGIAR-CSI Annual meeting
Amsterdam, March 2020

Validating and calibrating travel times

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

BHUTAN THIMPHU OKTOBER 28 - NOVEMBER 8 2019

What is *validation*?

Show or prove that something that you have measured or made is true to life, is valid. Get something confirmed and approved.

Such requires knowing the truth, to allow comparison. Can be done by

- Direct measurement (remember fieldwork?)
- Cross-checking with those who do or who know
- Trusted resources

Very different cost factors involved in such validation efforts.

Purposes

- Evidence and confidence that your method works, can be trusted and thus can be used
- Accumulate the proof basis of the method, which in turn allows to advocate/advertise its use

*Questions regarding travel times are common to **Living Standards Measurement Surveys** promulgated by the World Bank and hence data are available for many countries. Despite their popularity, reported travel times contain serious inaccuracies, biases and limitations. Respondents may misreport times (a.k.a. recall bias) and idiosyncratic household conditions (i.e. disabilities, physical fitness, schedules of nearby bus) create many outliers (Roberts et al. 2006). Assessing Nepal's 1996 Living Standards Measurement Survey, Jacoby (2000) notes that reported travel times within wards vary widely around the ward median value for these reasons. Similarly, when comparing reported travel times to a locally validated cost time model, Ahlström et al. (2011) found responses differed by up to 30% (+/-) of the mean modeled travel time for a given district. These findings imply that while such data are appropriate for household level analysis, aggregation and generalization from them is problematic. Reported times also do not work for assessing new or planned infrastructure and are prohibitively expensive to collect at scale.*

Banick et al, 2019

Cross-checking

Accessibility mapping in R
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

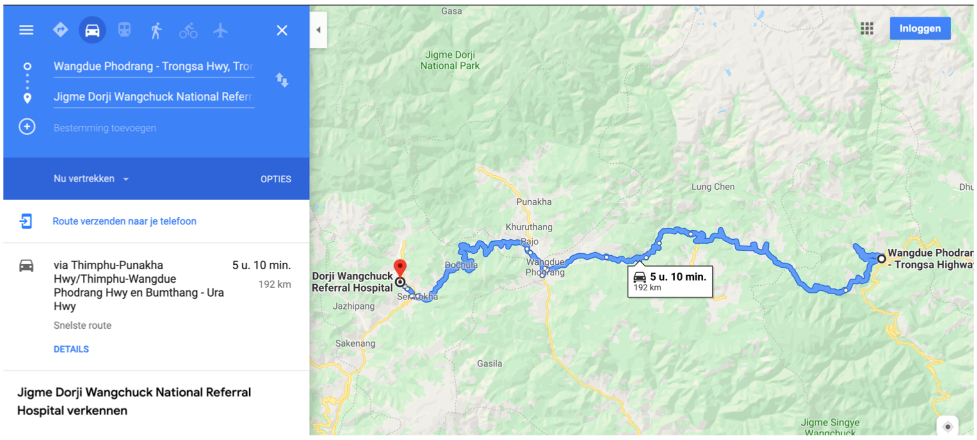
The accuracy and reliability of the result was tested and improved through consultations with professionals and organizations well-acquainted with travel patterns in diverse locations of Nepal. For additional validation we compared model results to reported service travel times from households surveyed in the Nepal Household Risk and Vulnerability Survey (HRVS).

Banick et al., 2019

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

One easily accessible trusted source

Google maps routing function



UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Accessibility mapping in R
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

That was a one-off

Can also script this for multiple routes:

```python
import simplejson, urllib3

orig_coord = 27.499024, 90.506626  # Trongsa main junction
dest_coord = 27.463743, 89.638934  # Thimpu Referral Hospital

http = urllib3.PoolManager()

url = "http://maps.googleapis.com/maps/api/distancematrix/json?\
       origins={0}&destinations={1}&mode=driving&language=en-EN&\
       sensor=false/".format(str(orig_coord),str(dest_coord))

r = http.request('GET', url)

print(r.data)

driving_time = result['rows'][0]['elements'][0]['duration']['value']
```

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Accessibility mapping in R
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

But such needs a Google account

Not necessarily money:

I've been making API calls without a key

I'm using APIs for free or without a key. Can I keep doing this?

You must have a valid API key and a billing account to access our APIs. When you enable billing, you will get $200 free usage every month for Maps, Routes, or Places. Based on the millions of users using our APIs today, most of them can continue to use Google Maps Platform for free with this credit. Having a billing account helps us understand our developers' needs better and allows you to scale seamlessly.

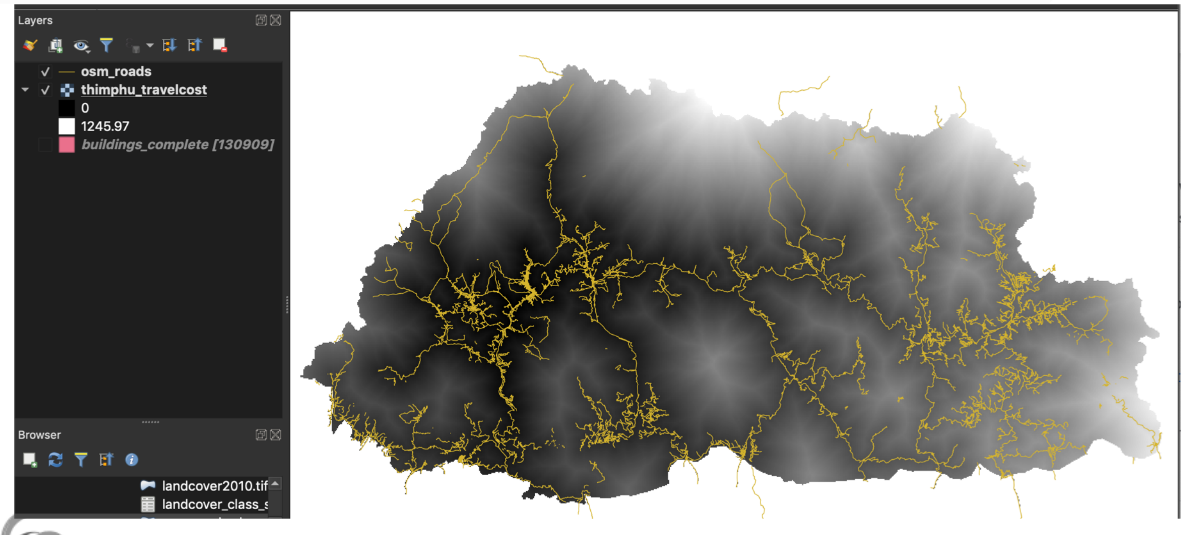
https://cloud.google.com/maps-platform/user-guide/account-changes/

The Thimphu Hospital travel time raster

```
################################################################
#  Validation;  Using travel times to the capital's referral hospital

thimphu_layer <- accCost_stack@layers[[22]]
writeRaster(thimphu_layer,
        filename=filepath("outputs/thimphu_travelcost.tif"),
        format="GTiff",
        overwrite=TRUE)
```

Our travel cost raster for Thimphu Hospital



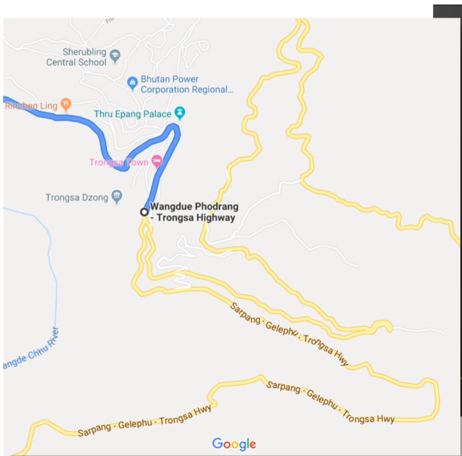
Accessibility mapping in R

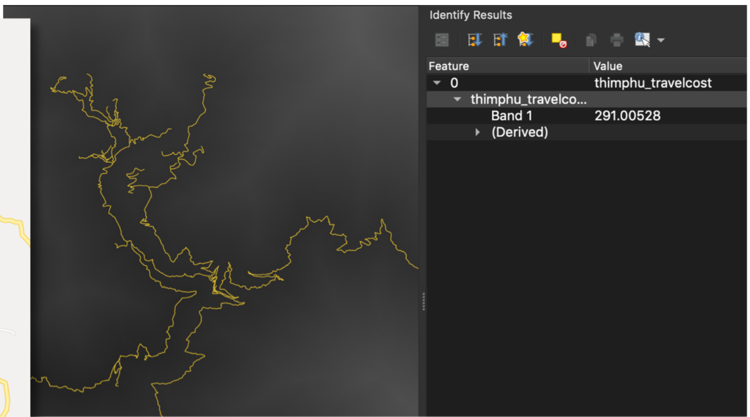
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

UNIVERSITY OF TWENTE.

ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Use the i-button and pick the junction's value





Accessibility mapping in R

Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

UNIVERSITY OF TWENTE.

ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Beware

Difference between
- *Actual travel times* (depending on traffic and time of day)
- *Nominal travel times* (ideal, no additional friction; government's ideal)

Understand Google's mechanism of travel time assessment:
- Believed to be based on people's actual driving time, averaged
- Depending on area may be time-sensitive (traffic intensity)

Calibration

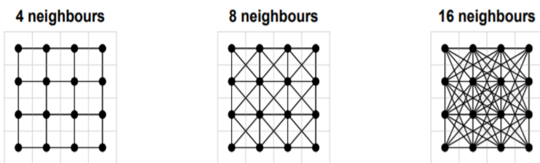
- Aims to improve model performance by tuning parameters and/or computational methods
- May target used input data, used constants
- Calibration often involves factoring out influence of other parameters and focusing on one single parameter
  - May mean identifying travel conditions where the one factor dominates travel time

Model critique

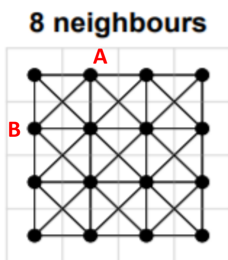
- Aims to predict nominal travel time
- Before calibration, generally seems to underestimate actual time (makes sense with bullet 1)
- Does not currently account for unidirectionality, which in mountainous areas is incorrect (technology in principle allows for it)
- Does not currently account for road curvature, which is known to affect speed (doable adjustment)
- Uses a coarse notion of slope (terrain slope, not road slope; affecting both driving directions equally)

Underlying technology

- Our model makes use of gdistance package, specifically its 8-neighbors incidence graph



- Observe cells (graph vertices) A and B on right. The incidence matrix has a weight for edge AB and one for BA. These need not be the same.



Our model: key questions

A test for your understanding of what our model computes

- In constructing the transition matrix, we base its weights on a raster; where in the code does this happen, what is the raster, and what does it express?
- Which formula assigns weights to edges like AB and BA in the incidence matrix? Is our matrix direction sensitive?
- How do we express the effect of slope on driving speed in the model? Is this a sensible method? Why? Can you think of alternatives?

Done

Accessibility mapping in R
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society
UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Accessibility modelling in R
CGIAR-CSI Annual meeting
Amsterdam, March 2020

Population discrepancy analysis

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

BHUTAN THIMPHU OKTOBER 28 - NOVEMBER 8 2019

The problem

We are getting different total population numbers from census and population raster. Why and what caused it?

601,301: **select sum(ntotal) --- this is SQL**
from vw_population2017_v4

571,951: **cellStats(population,stat='sum') # this is R**

5% difference is too much!

Accessibility mapping in R
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Strategy 1

We think *the census data are correct*, and we have used these to produce the population raster. So, somewhere we made an error producing that raster.

A first thought is that there may still be *census tracts* (aka popwogs) that have population, but we have *no buildings* there. So, let us find popwogs without buildings, and see how many people are in these.

All the relevant data for this strategy is in the database, so we will use SQL.

Strategy 1 code and results

```
select p.id, p.ntotal
from vw_population2017_v4 as p
where not exists
  (select *
   from building_v4 as b
   where st_intersects(p.geom,b.geom))
```

Data Output | Explain | Me

	id integer	ntotal integer
1	631	124
2	764	30
3	767	34
4	2913	83
5	3116	41
6	3122	29
7	3128	28
8	3179	22

We have found 391 people missing.

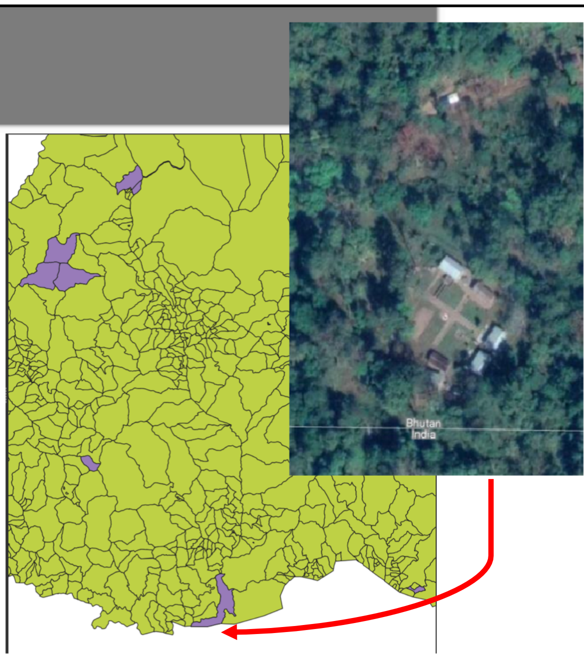
Discuss what to do about this case.
Invent buildings?

Strategy 1 results spatially

This is where the census tracts are.

The largest number (124) is found in Dagana—Nichula--Damchhunang.

We found at least some houses in that popwog via Google imagery.



Accessibility mapping in R
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Strategy 2

Harder work: recount the population from the raster by popwog and compare these numbers to popwog census numbers. We do not know what to expect, but differences should show up.

Method: use R to operate on the raster, much like we worked on single hospital service areas, but now popwogs as polygons.

Accessibility mapping in R
Quantification and Mapping of SDG Indicators that identify the most Vulnerable in today's Society

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Resolution

- After some data analysis, we identified the true problem:
  - Part of our workflow accounts for buildings and aggregates *by popwog vector geometry*
  - Then an average #people per building (appb) is determined for the popwog
  - We found that often popwog boundaries run along mountain ridges and that these are also often chosen road trajectories.
  - This means many buildings along the road, and thus on either side of the shared popwog boundary
  - For travel time analysis we work with population per raster cell, and use the appb of the dominant popwog. Buildings within a raster cell that belong to the other popwog do not count there and do count for the wrong popwog …