Review of remotely sensed data products for disease mapping and epidemiology

Copernicus Global Land User Conference, Toulouse, France
23–25 October, 2018

Sabelo Nick Dlamini (MSc, PhD)
Department of Geography, Environmental Science and Planning,
University of Eswatini
Outline

- Background: why are we interested?
  - Applications in disease mapping
  - Geostatistical modeling
  - Remotely sensed data convenience
  - Advances in technology

- Examples of applications in epidemiology
  - Disease vector distributions
  - Modeling examples

- Remotely sensed data products review
  - Review findings
  - Problems
  - Conclusions way forward
Background
Disease applications: remotely sensed data, geostatistics and GIS

- Geostatistical methods combining remote sensing (RS) techniques and geographic information systems (GIS) remain effective in estimating spatial and temporal effects of environmental determinants on disease outcomes.

- Geostatistics allow production of model-based maps which are used to predict outcomes at unsampled geographic locations.

- In case of disease control programmes this aids and support targeted control.

Email: sabelodlamini2004@yahoo.com
Geostatistical modelling in disease mapping

- Often, epidemics occur with quasi-regular cycles and are predictable with a reasonable degree of precision
- Meaning they could be preventable, if control programmes are adequately informed and prepared
- Geostatistical modeling is applied to improve our understanding and predictability of such epidemics

Email: sabelodlamini2004@yahoo.com
Advantages of using RS data products is its

- Near real-time availability for rapid assessment of at risk areas
- Allows prediction of disease distributions, especially in inaccessible areas with poor ground measurement meteorological station networks that may also lack baseline data
- Directly feeds into algorithms and models that attempt to create smooth/continuos climatic surfaces where ground measurement coverage is poor
The increase in the launch of higher resolution satellites and advances in processing techniques and Computational power have enabled wider adoption of RS data.

Large datasets that were previously difficult to handle and to process can now be efficiently processed and results obtained within reasonable time frames.
Examples of applications in epidemiology
Disease vector distributions

- Vector borne diseases such as malaria, are highly sensitive to **changing temperatures, precipitation and altitude** (RS products)
- Which affect vector breeding habits
  - For instance, stagnant pools of water make the physical environment more suitable for vector to oviposit and hatch
- These conditions determine epidemic situations

Email: sabelodlamini2004@yahoo.com
Examples: RS products for Epidemiology

EO in Epidemiology

Vegetation Maps

Medium Resolution
Land Cover/Use Maps

Email: sabelodlamini2004@yahoo.com
Examples: RS products for Epidemiology

- EO in Epidemiology

Digital Elevation Maps

Water Body Maps

Wind Blown Dust Maps

Land Surface Temperature Maps

Email: sabelodlamini2004@yahoo.com
Example Land Cover Classification

RapidEye 09/10/2011

Land Cover Classification

Legend
- Forest/Woodland
- Bushland
- Grassland
- Agriculture
- Bare Soil/Rock
- Water
- Settlement/Infrastructure
- Roads
- Cloud/Cloudshadow

(c) RapidEye (2011), provided under EC/ESA GSC-DA

Email: sabelodlamini2004@yahoo.com
Applications in modeling: Negative binomial model fitted in WinBUGS

Let $Y_i$ be the average number of malaria cases at a given location $s = i, \ldots, n$

with likelihood $Y_i \sim \text{dnegbin}(P_i, r)$

where $P_i$ is the proportion of malaria cases in a defined location and $r$ is the dispersion parameter and $\mu_i = r \frac{1-p}{p}$ while $\sigma_i^2 = r(1-p)p^{-2}$. The model was written as: $\text{Logit}(\mu_i) = \text{logit}(Popu_i) + \beta_0 + \beta_1 X_i \ldots + \beta_{12} X_i + \varepsilon_i$, where $\mu_i$ is the number of malaria cases in each location and $\beta$ are the regression coefficients, $X$ are the model covariates and $\varepsilon$ are spatial/temporal random effects. The distributed lag model was restricted to a polynomial function of power 2 which was formulated as follows:

Email: sabelodlamini2004@yahoo.com
Distributed lag model restricted to a polynomial function of power 2 could be formulated as follows:

$$\beta_i = \sum_{k=0}^{2} a_k i^k$$

Where $k$ is the categorical variable for the covariate corresponding to $\beta_i$ coefficient and $a$ is the intercept for locations $i$ .... $n$. The model describes the relationship between an independent value of $X_i$ and the corresponding dependent mean $Y_i$. This is summarized as $E(y|x)$. The model gives the expected $\mu_i$ of malaria cases given the corresponding value of each categorical variable at location s.
Bayesian geostatistical modeling to assess spatio–temporal variations and elapsing time for malaria incidence risk in Eswatini
Sabelo Nick Dlamini*¹, and Sizwe Mabaso¹
¹University of Eswatini, Kwaluseni Camous, Matsapha, Eswatini
Example: some model outputs

- Water Body Mapping
- Distance to water assessment

(c) RapidEye (2011), provided under EC/ESA GSC-DA

Email: sabelodlamini2004@yahoo.com
Potential vector breeding sites classification of the malaria-endemic area covered by 5-m resolution RapidEye data (northern part of the Lowveld)

Assessing the relationship between environmental factors and malaria vector breeding sites in Swaziland using multi-scale remotely sensed data

Sabelo Nick Dlamini,1,2 Jonas Franke,3 Penelope Vounatsou1,2
1Swiss Tropical and Public Health Institute, Basel; 2University of Basel, Switzerland; 3Remote Sensing Solutions GmbH, Baierbrunn, Germany
So what?

- Evidently RS products greatly support disease mapping distributions
- **Very high to high resolution** remotely sensed (RS) data products remains of interest to disease mapping studies
- Although high costs had been a barrier for most end users – there are changes already taking place in RS landscape – DRONES
  - May have implications on future pricing

Email: sabelodlamini2004@yahoo.com
Remotely sensed data products review
We conducted an online review to find out what RS data products were accessible for disease mapping and epidemiology.

Purpose was:
- To document RS data products for disease mapping
- To propose new products that could be incorporated in disease mapping studies
- Need for information availability on the type of data products obtainable from satellite imagery in order to facilitate access and utilization

Email: sabelodlamini2004@yahoo.com
RS review findings
Need for Information Bulletin

- National Aeronautics and Space Administration (NASA) monthly bulletin called Spacewarn was launched in 1991 (http://nssdc.gsfc.nasa.gov/spacewarn/) to raise awareness about newly launched satellites and their missions to the general public.
- Unfortunately, this bulletin was discontinued with the last issue available until July 2011—how about products?

Email: sabelodlamini2004@yahoo.com
Review findings

- Little effort made to inventorize existing and potential RS data products that could be used in disease mapping and epidemiology
- Space activity (satellite launches, landuse/cover observations) has increased
- Information on the type of RS data products obtainable from these missions is somewhat elusive.

Email: sabelodlamini2004@yahoo.com
Review findings: distribution of spatial resolutions of remote sensing sensors and products

Review of remotely sensed data products for disease mapping and epidemiology
Sabelo Nick Dlamini*1, and Anton Beloconi2,3,
1University of Eswatini, Kwaluseni Campus, Matsapha, Eswatini
2Swiss Tropical and Public Health Institute, Basel, Switzerland
*Correspondence: sabelodlamini2004@yahoo.com

In Review: Remote Sensing Journal: Society and Environment
Present documentation of RS data products from satellites imagery is ad-hoc, incomplete and characterized by duplication and redundancy between access websites. E.g. MODIS products were found across access sources such as Copernicus (Sentinel Scientific Data Hub), USGS Earth Explorer, NASA Earth Data among others.

- (Host, supply agency or middleman?)
- Good descriptions but direct download not found?
- Magnified resolutions!!!
Review findings cont.

- ESA comprised of 58 already launched missions and about 18 planned launches from year 2018 to 2028 while for NASA missions we found about 200 launches are faithfully recorded but there is little info on the resultant products for end users.

- We found various inexhaustive data download ftp sites managed by both remote sensing agencies and private web bloggers who enthusiastically follow remote sensing issues.

Email: sabelodlamini2004@yahoo.com
For instance, such websites included NASA Goddard Space Flight Center found on https://www.nasa.gov/goddard, Copernicus Open Access Hub) previously known as Sentinels Scientific Data Hub (https://scihub.copernicus.eu/), Observing Systems Capability Analysis and Review Tool (OSCAR), including NORAD Catalog (http://satellitedebris.net/Database/), and Global Visualization Viewer (GLOVIS) among others.

(although they may look few but when looking for some products and you are not sure who has it as it is made mention of across agencies but you can not really trace it to a supply agency e.g NDWI, soil moisture)

When you get products via a middleman its authenticity may be in doubt.

Email:
sabelodlamini2004@yahoo.com
Modeling problems

- Alignment in time – e.g. daily LST against dekadal rainfall or even 8/16 days NDVI
- Indices like LST do not directly measure ambient air temperature, critical consideration in vector dispersal
- Not easy to know continuous or sustained acquisition programme from once-off project specific products
  - End of RS data products supply not known

Email: sabelodlamini2004@yahoo.com
We prepared a catalogue of indices from ecological studies that could be used as covariates/predictors in disease mapping and epidemiology studies.

Remotely sensed data products related to climate, meteorology, land use/cover, cartography and urban mapping were explored as potential indices for disease mapping and epidemiology.
Conclusions

- There remains a substantial amount of work to be conducted on the evaluation and validation of some of the indices presented in this study.
- Synergies between remote sensing experts and epidemiologists could be useful in the uptake and testing of new RS products as well as some of the RS data products presented in this catalogue.
Conclusions cont.

- Development of indices that are sensitive to vector conditions or take into account the climatic conducive environments of disease causing vectors could be useful for epidemiologists.

- Need to identify and define RS data products in terms of their characteristics and formats as this relates to the decision to select data products for any analysis.

Why?

- The characteristics and formats of RS data products as well as availability of handling software for interoperability still determine their accessibility and utilization by end users.
Siyabonga
Merci
Danke schön
Thank you: