

SPATIAL DATA IN R

RACHEL KUZMICH
(not a developer)



Intro

About me

Before

- BSc in geography and biology at Trent University
- thesis on time series analysis of forest disturbances in the Alberta oil sands
- published papers about on too, and uav work for tree species discrimination

Now

- pursuing a MSc in geography at Queen's University
- research project (and eventually thesis) on cerulean warbler habitat preferences using lidar and passive acoustic monitoring

About today's talk

- what is gis and remote sensing
- spatial data types:
 - raster, vector, las, tabular
- where spatial data comes from and how to find it
- time series example:
 - analysis and visualization basics in R

GIS and remote sensing

GIS

- geographic information system/science
- antecedents: cartography, navigation, epidemiology
- grandfather of modern GIS is Roger Tomlinson
- storage, analysis, editing, display, sharing of geographic information
- makes location the primary variable
- can use a gis to make maps, but gis is not the same as cartography

Remote sensing

science of acquiring information about the surface of, or objects on, the earth (or another planet) without coming in direct contact with them

- antecedents: surveying, reconnaissance
- earthrise image (Apollo 8, 1968), blue marble image (Apollo 17, 1972)
- sensing and recording of reflected or emitted energy
- reliant on the electromagnetic spectrum (visible and non visible wavelengths)
- involves a sensor (i.e., consumer grade camera, single band, multispectral...) mounted on a platform (i.e., ground-based, uav, aircraft, satellite...)

Why use R?

- language and environment (1993)
- based on S (1976)
- open source
- free to run, share, modify
- versatile
- customization
- statistics
- integration with other languages (c, c++)
- extended capacity via a whole bunch of packages
- and a whole bunch of forums and communities
- download R: <http://cran.utstat.utoronto.ca/>
- download R Studio: <https://www.rstudio.com/products/rstudio/download/>

Spatial data

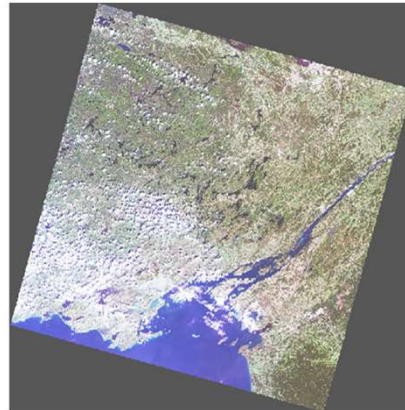
Spatial data standards

(geo)spatial data: data about geographic locations

- instrumentation calibration
- data collection
- naming
- reporting
- language
- sharing
- assessment
- metadata

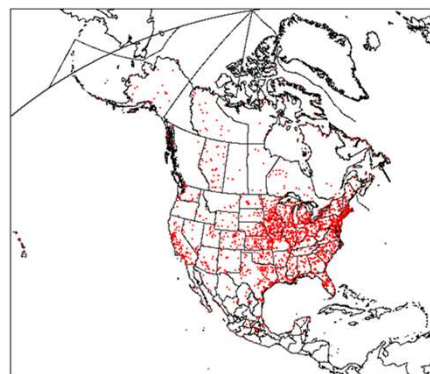
Raster

- matrix of cells
- fundamental units are pixels
- integer (discrete/categorical), floating point (continuous)
- pixel resolution related to real-world space
- in addition to location, each pixel has a single value of some attribute class



Vector

- vertices and paths
- points: one set of coordinates (x, y, and sometimes z too)
- lines: connected ordered sets of coordinates
- polygons: closed line, first and final coordinates are the same
- organized as shapefile representing a single feature class with different attributes



LAS

- collected using lidar (aka als)
- involves a platform, laser scanner system, GPS, INS
- measures time it takes for pulse return after interacting with surface/object
- output is dense three dimensional spatially explicit point cloud
- each point has x, y, z coordinates and intensity values



Tabular

- often as attributes joined to other data
- sometimes standalone

```

20800394      URN:catalog:BSC-EOC:OBBA2BE_SUMM:15UP89-ABDU      2008-05-14 17:18:40.157000000 Observation      BSC-EOC OBBA2BE_SUMM      15UP89-ABDU      Anas rubripes      AnimaliaChorda
est Coordinates of the 10x10km square      OBBA2      Breeding Bird Atlas      BBATLAS Breeding birds      NORAC, 2006      http://www.bsc-eoc.org/norac/atlascont
ANGO      2008-05-14 17:18:40.157000000      Observation      BSC-EOC OBBA2BE_SUMM      15UP89-ANGO      Carduelis tristis      AnimaliaChordataAvesPasseriformesFringillidae      Animalia
re      OBBA2      Breeding Bird Atlas      BBATLAS Breeding birds      NORAC, 2006      http://www.bsc-eoc.org/norac/atlascont.htm      15UP89      10000
0      Observation      BSC-EOC OBBA2BE_SUMM      15UP89-AMRO      Turdus migratorius      AnimaliaChordataAvesPasseriformesTurdidae      Animalia
rds      NORAC, 2006      http://www.bsc-eoc.org/norac/atlascont.htm      15UP89      10000      SQUARE      10000      10000      0      15UP89      15UP89      Chordata      Aves      Passer:
ula      AnimaliaChordataAvesPasseriformesIcteridae      Animalia      Chordata      Aves      Passeriformes      Icteridae      Icterus galbula      49.883      hours      North America
scont.htm      15UP89      10000      SQUARE      10000      10000      0      15UP89      15UP89      Coccyzus      erythrophthalmus      49.883      hours      North America
culidae      Animalia      Chordata      Aves      Cuculiformes      Cuculidae      15UP89      10000      15UP89      Coccyzus      erythrophthalmus      49.883      hours      North America
0      15UP89      15UP89      iformes      Alcedinidae      Ceryle      alcyon      North America      Canada      Ontario      1      No
America      Canada      Ontario      20800416      URN:catalog:BSC-EOC:OBBA2BE_SUMM:15UP89-BOBO      2008-05-14 17:18:40.157000000 Observation      BSC-EOC OBBA2BE_SUMM      15UP89-BOBO      Toxostoma rufum      AnimaliaChordataAvesPa
NAD83      10000      AOU      7.048      2007      BLTE      H      20800416      URN:catalog:BSC-EOC:OBBA2BE_SUMM:15UP89-BOBO      2008-05-14 17:18:40.157000000 Observation      BSC-EOC OBBA2BE_SUMM      15UP89-BOBO      Toxostoma rufum      AnimaliaChordataAvesPa
9      URN:catalog:BSC-EOC:OBBA2BE_SUMM:15UP89-BRTH      2008-05-14 17:18:40.157000000 Observation      BSC-EOC OBBA2BE_SUMM      15UP89-BRTH      Toxostoma rufum      AnimaliaChordataAvesPa
Coordinates of the 10x10km square      OBBA2      Breeding Bird Atlas      BBATLAS Breeding birds      NORAC, 2006      http://www.bsc-eoc.org/norac/atlascont
8:40.157000000 Observation      BSC-EOC OBBA2BE_SUMM      15UP89-CAGO      Branta canadensis      AnimaliaChordataAvesAnseriformesAnatidae      Animalia      Chordata      Aves
BBATLAS Breeding birds      NORAC, 2006      http://www.bsc-eoc.org/norac/atlascont.htm      15UP89      10000      SQUARE      10000      10000      0      15UP89      15UP89      Passeriformes      Emberizidae      Spizel
BA2BE_SUMM      15UP89-CHSP      Spizella passerina      AnimaliaChordataAvesPasseriformesEmberizidae      Animalia      Chordata      Aves      Passeriformes      Emberizidae      Spizel
006      http://www.bsc-eoc.org/norac/atlascont.htm      15UP89      10000      SQUARE      10000      10000      0      15UP89      15UP89      Anseriformes      Anatidae      Bucephala      clangula      49.883      hours
la clangula      AnimaliaChordataAvesAnseriformesAnatidae      Animalia      Chordata      Aves      Anseriformes      Anatidae      Bucephala      clangula      49.883      hours
ac/atlascont.htm      15UP89      10000      SQUARE      10000      10000      0      15UP89      15UP89      Chordeiles      minor      North America
lgidae      Animalia      Chordata      Aves      Caprimulgiformes      Caprimulgidae      Chordeiles      minor      North America
10000      0      15UP89      15UP89      olopacidae      Gallinago      gallinago      North America      Canada      Ontario      1
  
```

What should you use?

Depends on what you are trying to do, what's important, relevant and meaningful to your analysis, where you got your data from and its nature, and how hard you want to make your life.

Raster

- continuous phenomena
- categorical data

Vector

- discrete locations
- extents of spatial objects

LAS

- 3d environments
- surfaces

Assumptions, pros and cons

Raster

- data collected by remote sensing
- assumes the world is full, up to you to partition it
- display sensitive to resolution

Vector

- data collected by surveying
- assumes the world is empty, up to you to add elements
- simplification and data reduction

LAS

- data collected either by remote sensing or surveying
- assumes the world is 3d
- more point data than can be used

Spatial data sources

Spatial data sources

Conventional data collection

- typically done by (or with support from) government agencies

Crowdsourcing

- common in emergency situations, study of species occupying a wide area

“Volunteered geographic information”

- data collected through apps when you agree to the terms and conditions

Create your own

Accessing spatial data

Landsat satellite archive: earthexplorer.usgs.gov

Government

- Canada: open.canada.ca/data/en/dataset
- Ontario: ontario.ca/search/data-catalogue
- Kingston: cityofkingston.ca/explore/data-catalogue

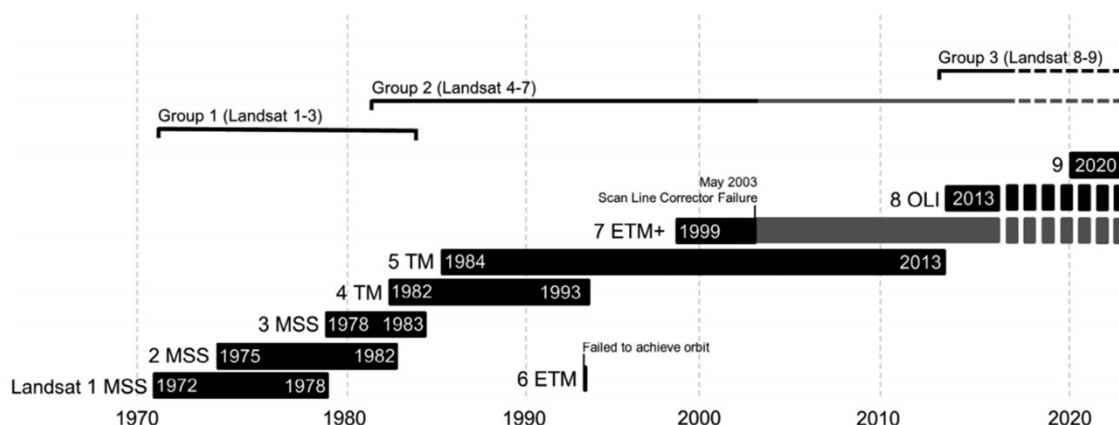
Other places

- agencies, libraries and repositories: fao, un, cec, usgs, statscan, conservation authorities, land inventory ontario, scholars geoportal, esri
- private (paid): digital globe



**Time series
example**

Background and context



Band	Name	Wavelength	Applications
1	Ultra blue	0.435 - 0.451	Coastal and aerosol studies
2	Blue	0.452 - 0.512	Bathymetric mapping, distinguishing soil from vegetation, and deciduous from coniferous vegetation
3	Green	0.533 - 0.590	Emphasizes peak vegetation, which is useful for assessing plant vigor
4	Red	0.636 - 0.673	Discriminates vegetation slopes
5	Near infrared	0.851 - 0.879	Emphasizes biomass content and shorelines
6	Shortwave infrared I	1.566 - 1.651	Discriminates moisture content of soil and vegetation; penetrates thin clouds
7	Shortwave infrared II	2.107 - 2.294	Improved moisture content of soil and vegetation and thin cloud penetration
8	Panchromatic	0.503 - 0.676	15 metre resolution, sharper image definition
9	Cirrus	1.363 - 1.384	Improved detection of cirrus cloud contamination
10	Thermal infrared I	10.60 - 11.19	100 metre resolution, thermal mapping and estimated soil moisture
11	Thermal infrared II	11.50 - 12.51	100 metre resolution, Improved thermal mapping and estimated soil moisture

Scene ID	Landsat Product Identifier
LXSPPPRRRRYYYYDDGSIVV	LXSS_LLLL_PPPRRR_YYYYMMDD_yyyymmdd_CC_TX
<p>L = Landsat X = Sensor S = Satellite PPP = WRS path RRR = WRS row YYYY = Year DDD = Julian day of year GSI = Ground station identifier VV = Archive version number</p> <p>Examples: LC80290302015343LGN00 LE70160392004262EDC02 LT40170361982320XX08 LM10170391976031AAA01</p>	<p>L = Landsat X = Sensor ("C" = OLI/TIRS Combined, "O" = OLI-only, "T" = TIRS-only, "E" = ETM+, "T" = TM, "M" = MSS) SS = Satellite ("07" = Landsat 7, "08" = Landsat 8) LLLL = Processing correction level ("L1TP": Precision Terrain, "L1GT": Systematic Terrain, "L1GS": Systematic) PPP = WRS path RRR = WRS row YYYYMMDD = Acquisition year (YYYY) / Month (MM) / Day (DD) yyymmdd = Processing year (yyyy) / Month (mm) / Day (dd) CC = Collection number ("01", "02") TX = Collection category: ("RT" for Real-Time, "T1" for Tier 1, or "T2" for Tier 2)</p> <p>Examples: LC08_L1GT_029030_20151209_20160131_01_RT LE07_L1TP_016039_20040918_20160211_01_T1 LT04_L1GS_017036_19821115_20160315_01_T2 LM01_L1GS_017039_19760131_20160225_01_T2</p>

ISPRS Journal of Photogrammetry and Remote Sensing 130 (2017) 370–384

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Change detection using landsat time series: A review of frequencies, preprocessing, algorithms, and applications

Zhe Zhu^{a,*}

^a Department of Geosciences, Texas Tech University, Lubbock, TX 79409, United States
Center for Geospatial Technology, Texas Tech University, Lubbock, TX 79409, United States
Climate Science Center, Texas Tech University, Lubbock, TX 79409, United States

INTERNATIONAL JOURNAL OF DIGITAL EARTH, 2016
<http://dx.doi.org/10.1080/17538947.2016.1187673>

Mass data processing of time series Landsat imagery: pixels to data products for forest monitoring

Txomin Hermosilla^a, Michael A. Wulder^b, Joanne C. White^b, Nicholas C. Coops^a, Geordie W. Hobart^c and Lorraine B. Campbell^d

^a Canadian Forest Service (Pacific Forestry Centre), Natural Resources Canada, 506 West Burrard Road, Victoria, BC V8T 1X5, Canada
^b Department of Forest Resource Management, Forest Sciences Centre, 2424 Main Mall, University of British Columbia, Vancouver, BC V6T 1Z6, Canada

REMOTE SENSING LETTERS, 2018
VOL. 9, NO. 2, 160–169
<https://doi.org/10.1080/2150704X.2017.1410293>

Detection accuracy of new well sites using Landsat time series data: a case study in the Alberta Oil Sands Region

Rachel J. Wasson and Steven E. Franklin

School of Environment, Trent University, Peterborough, Ontario, Canada

How Similar Are Forest Disturbance Maps Derived from Different Landsat Time Series Algorithms?

Warren B. Cohen^{a,*}, Sean P. Healey^b, Zhiqiang Yang^c, Stephen V. Stehman^d, C. Kenneth Brewer^e, Evan B. Brooks^f, Niel Gorelick^g, Chongqiang Huang^h, M. Joseph Hughesⁱ, Robert E. Kennedy^j, Thomas R. Loveland^k, Gretchen G. Mladenov^l, Todd A. Schoneveld^{m,n}, James E. Vogelmann^o, Curtis E. Woodcock^o, Limin Yang^o and Zhe Zhu^o

^a USDA Forest Service, Pacific Northwest Research Station, 3200 SW Jefferson Way, Corvallis, OR 97331, USA
^b USDA Forest Service, Rocky Mountain Research Station, 9700 NE 28th Avenue, Ogden, UT 84401, USA
^c USDA Forest Service, Pacific Northwest Research Station, 3200 SW Jefferson Way, Corvallis, OR 97331, USA
^d USDA Forest Service, Pacific Northwest Research Station, 3200 SW Jefferson Way, Corvallis, OR 97331, USA
^e USDA Forest Service, Pacific Northwest Research Station, 3200 SW Jefferson Way, Corvallis, OR 97331, USA
^f USDA Forest Service, Pacific Northwest Research Station, 3200 SW Jefferson Way, Corvallis, OR 97331, USA
^g USDA Forest Service, Pacific Northwest Research Station, 3200 SW Jefferson Way, Corvallis, OR 97331, USA
^h USDA Forest Service, Pacific Northwest Research Station, 3200 SW Jefferson Way, Corvallis, OR 97331, USA
ⁱ USDA Forest Service, Pacific Northwest Research Station, 3200 SW Jefferson Way, Corvallis, OR 97331, USA
^j USDA Forest Service, Pacific Northwest Research Station, 3200 SW Jefferson Way, Corvallis, OR 97331, USA
^k USDA Forest Service, Pacific Northwest Research Station, 3200 SW Jefferson Way, Corvallis, OR 97331, USA
^l USDA Forest Service, Pacific Northwest Research Station, 3200 SW Jefferson Way, Corvallis, OR 97331, USA
^m USDA Forest Service, Pacific Northwest Research Station, 3200 SW Jefferson Way, Corvallis, OR 97331, USA
ⁿ USDA Forest Service, Pacific Northwest Research Station, 3200 SW Jefferson Way, Corvallis, OR 97331, USA
^o USDA Forest Service, Pacific Northwest Research Station, 3200 SW Jefferson Way, Corvallis, OR 97331, USA



Over to R....



**Thanks
dudes**