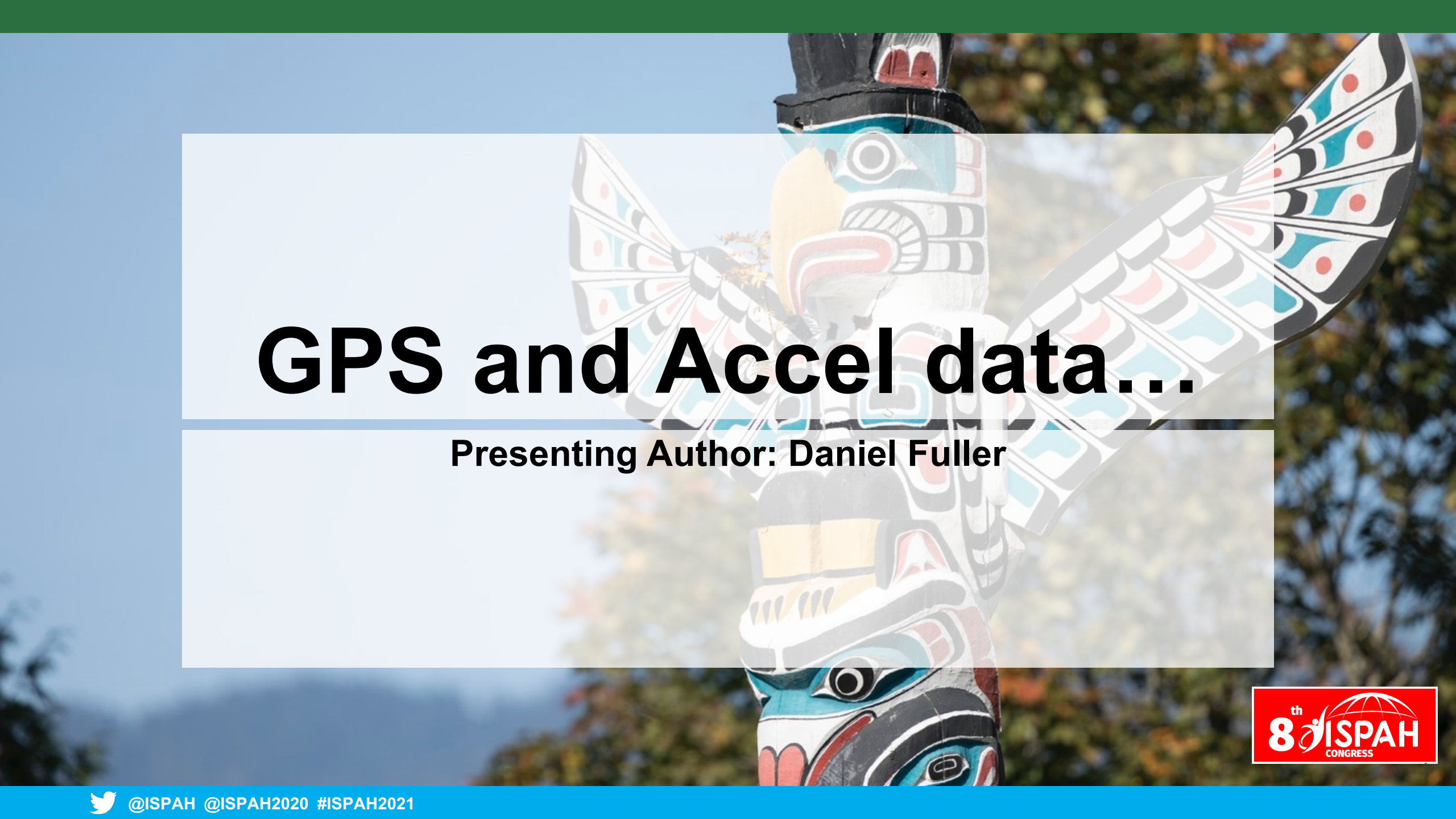


GPS and Accel data…

Presenting Author: Daniel Fuller

Territorial Acknowledgement

We respectfully acknowledge the territory in which we gather as the ancestral homelands of the Beothuk, and the island of Newfoundland as the ancestral homelands of the Mi'kmaq and Beothuk. We would also like to recognize the Inuit of Nunatsiavut and NunatuKavut and the Innu of Nitassinan, and their ancestors, as the original people of Labrador. We strive for respectful relationships with all the peoples of this province as we search for collective healing and true reconciliation and honour this beautiful land together.

COI : No

The 8th International Society for Physical Activity and Health Congress

Disclosure of Conflict of Interest

Name of first author: Daniel Fuller
Affiliation: Memorial University of Newfoundland

I have no Conflict of Interest with regard to our presentation.

Overview

- Promise and challenges of GPS and accelerometer data
  - 20 Minutes
- Coding and examples (Using R)
  - 40 minutes
- Machine learning – The future?
  - 20 minutes

Background questions

1. Rate your familiarity with processing GPS data (post in the chat)
   0 = No idea ; 10 = Expert
2. Rate your familiarity with processing accel data (post in the chat)
   0 = No idea ; 10 = Expert
3. Rate your familiarity with processing gps+accel data (post in the chat)
   0 = No idea ; 10 = Expert

Promise and challenges

- Many conceptual papers about GPS and accel data
  - James P, Jankowska M, Marx C, et al. “Spatial Energetics.” *Am J Prev Med*. 2016;51(5):792-800. doi:10.1016/j.amepre.2016.06.006
  - Duncan MJ, Badland HM, Mummery WK. Applying GPS to enhance understanding of transport-related physical activity. *J Sci Med Sport*. 2009;12(5):549-556. doi:10.1016/j.jsams.2008.10.010
- Attempts to simplify and automate the processes
  - PALMS
  - HABITUS (https://www.habitus.eu/)

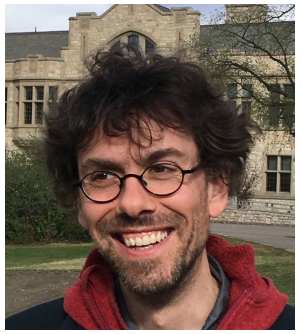
Promise and challenges

- Fewer consistent and well developed papers combining and applying GPS and accel data
- No standard methods: Everything is bespoke decisions, assumptions, and code
- Need to work to standardize or make more explicit our goals

Our team








- INTERACT is a pan-Canadian collaboration of scientists, urban planners, and citizens uncovering the impact of urban changes on health and equity.
- 4 cities, 2500 participants
- Over 8TB of GPS and Accel data
- Dedicated GPS and Accel + Smartphone app

Accel data

- Measures acceleration in X,Y,Z axis
  - Hip, wrist, or other locations
- Used to calculate activity intensity, activity type, or both
  - Intensity (Sedentary, Light, Moderate, Vigorous)
  - Type (Siting, Standing, Walking, Running)
  - Intensity and Type (Siting, Walking 3METS, Walking 7METS)
- Many, many, many approaches… probably too many

Accel data

- Processing
  - Raw Accelerometer processing
    - Filter and/or smooth the high frequency accelerometer signal (10-100Hz)
    - Combine the filtered x,y,z data in some way
    - Summarize the combined filtered x,y,z data
  - ActiGraph Activity Counts (AAC)
  - Euclidean Norm Minus One (ENMO)
  - Monitor Independent Movement Summary (MIMS)

Accel data

- Wearing
  - Device wear detection
  - Choi algorithm is the standard
- Classification
  - Classify activity into groups based on combined x,y,z data
    - Rule based classification
      - Cut-points
    - Machine learning based classification
      - More on that later

Accel data

- End result
  - Minute by minute classification for activity intensity, type, or both depending on the method you use
  - Can sum up to to get minutes of MVPA or sedentary minutes per day, week, etc.
  - Can also sum up by spatial concepts like activity space, trips along bike lanes, or home location.

GPS data

- Collects spatial location data based on trilateration of location from Global Positioning System Satellites (https://en.wikipedia.org/wiki/Satellite_navigation)
  - GPS (USA)
  - GLONASS (Russia)
  - BeiDU (China)
- Location can also come from cell phone towers through cell phone.
- Sensor fusion can create one location that combines GPS and cell phone location to get a more accurate location estimate

GPS data

- GPS Strengths and Limitations
  - Strengths
    - Global
    - Accurate (a best +/- 5 meters)
    - Very common
  - Limitations
    - Not good inside
    - Urban canyons
    - Slow to connect when signal is lost

Why combine data?

- Better measures of exposures that are spatially distributed
- Can create new metrics
  - Trips and dwells
  - Transportation mode detection

Questions?

Jump to R

- Caveats
  - There are no right answers
  - This is not a definitive guide
  - Lots of hard thinking needs to happen
  - Stop me and ask questions

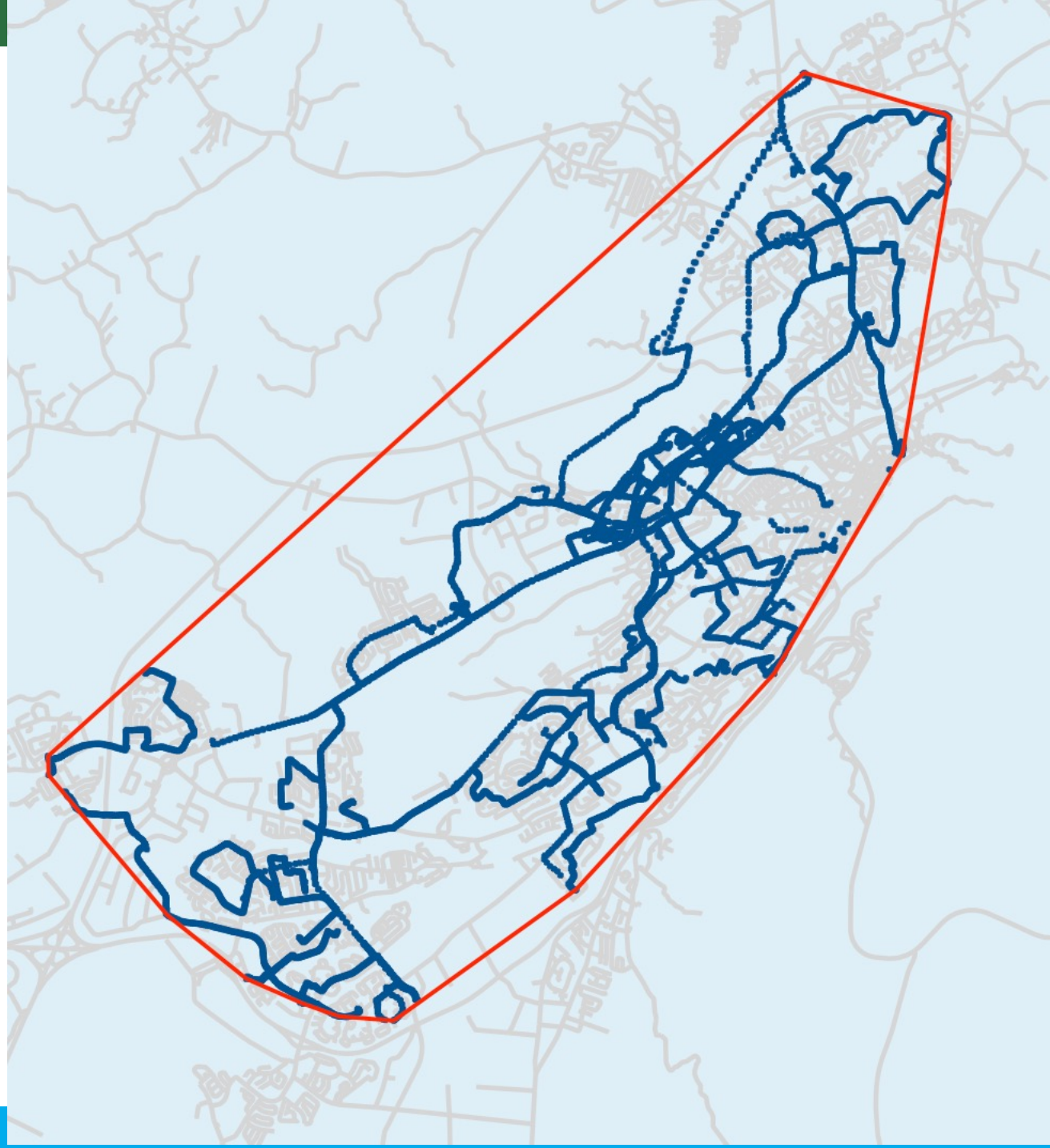
Back from R

- Time is terrible
- You will drop a lot of data (typically accel data)

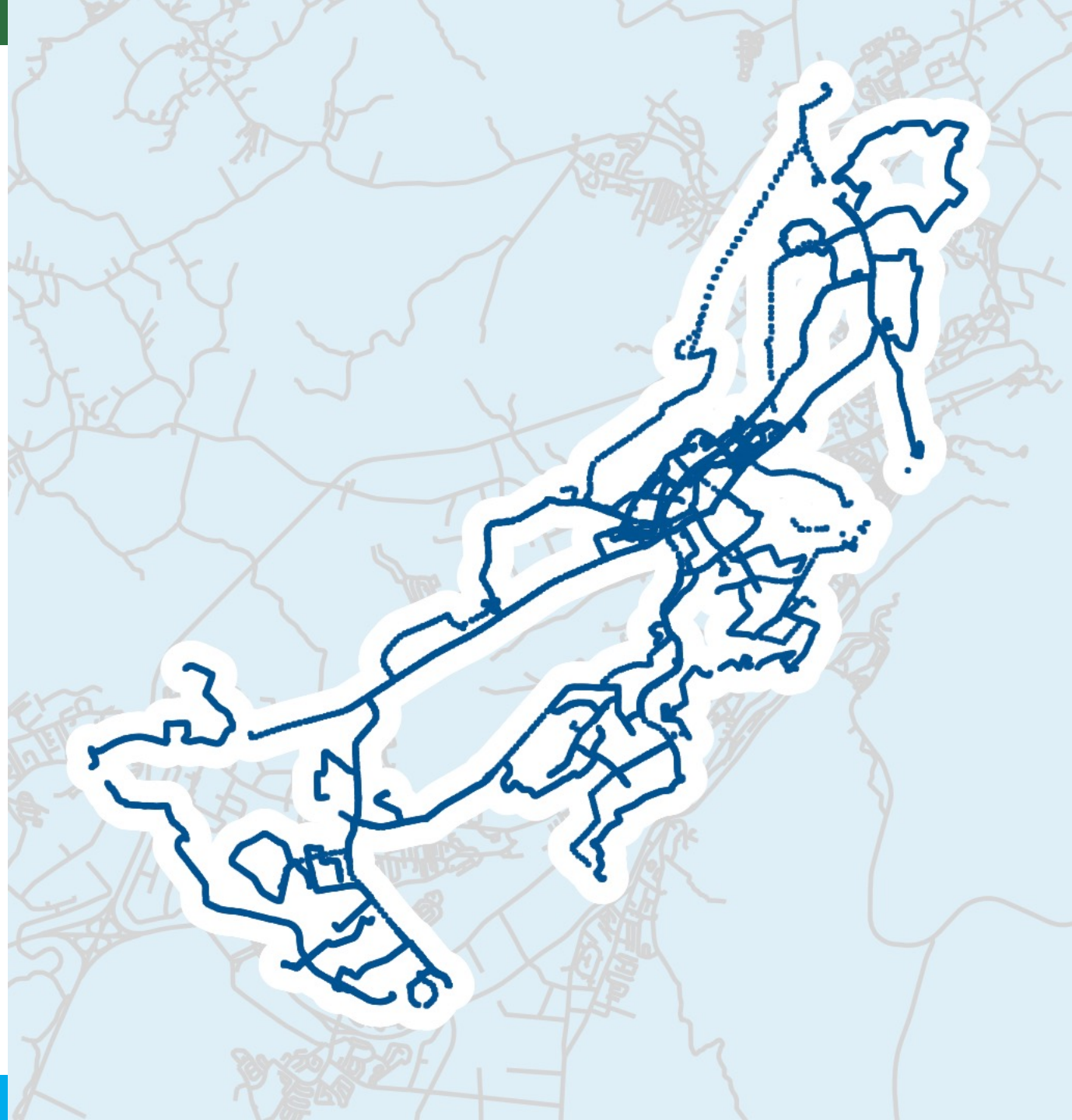
Machine learning and fancy stuff

- Activity spaces
  - Convex hull
  - Standard deviation ellipse
  - Buffer area
  - Entropy
  - Fractal Dimension
- First 3 can easily be calculated for the GPS data.
  - https://geocompr.robinlovelace.net/ will get you started

Convex Hull

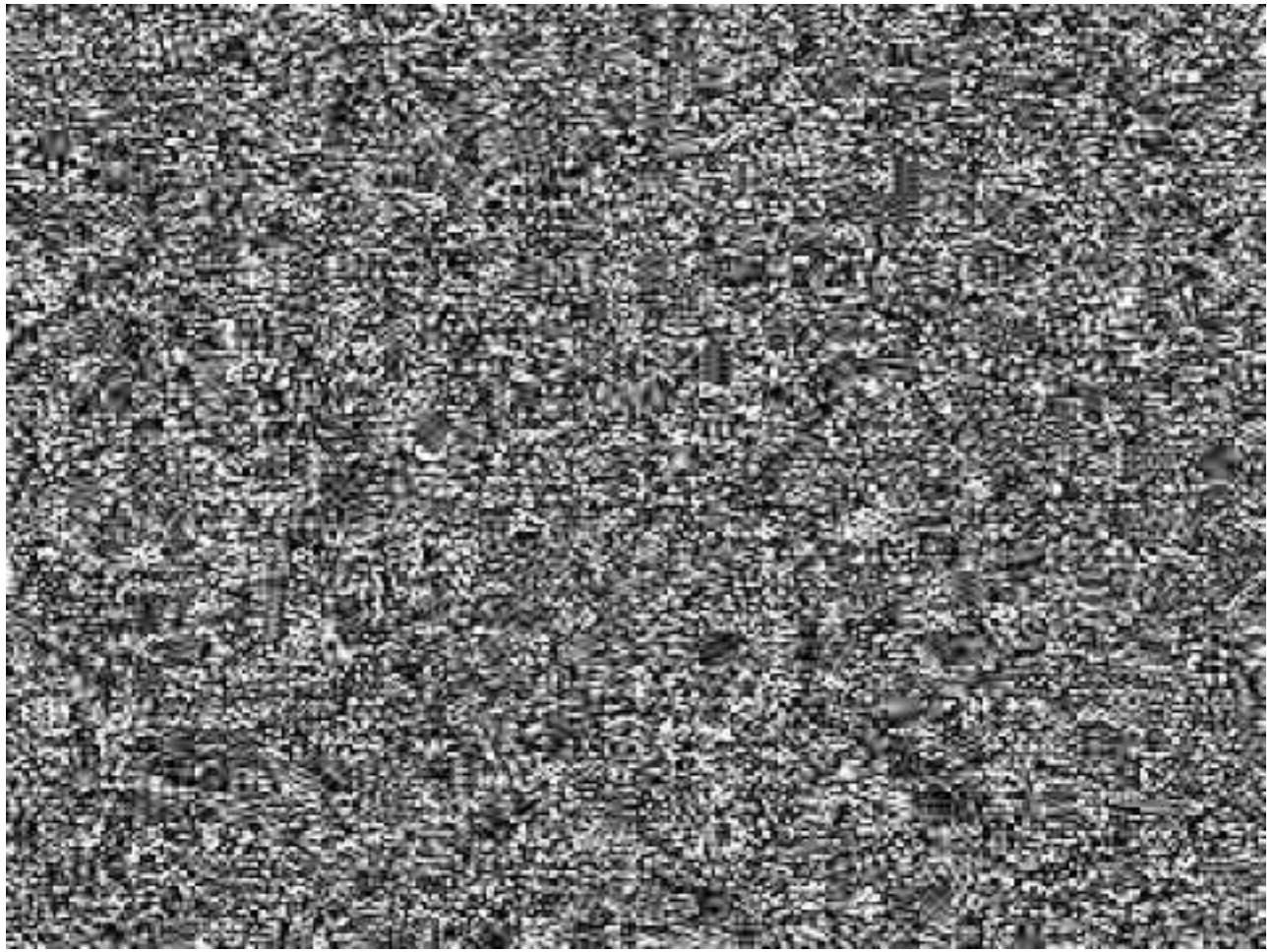


Buffer Area



Entropy

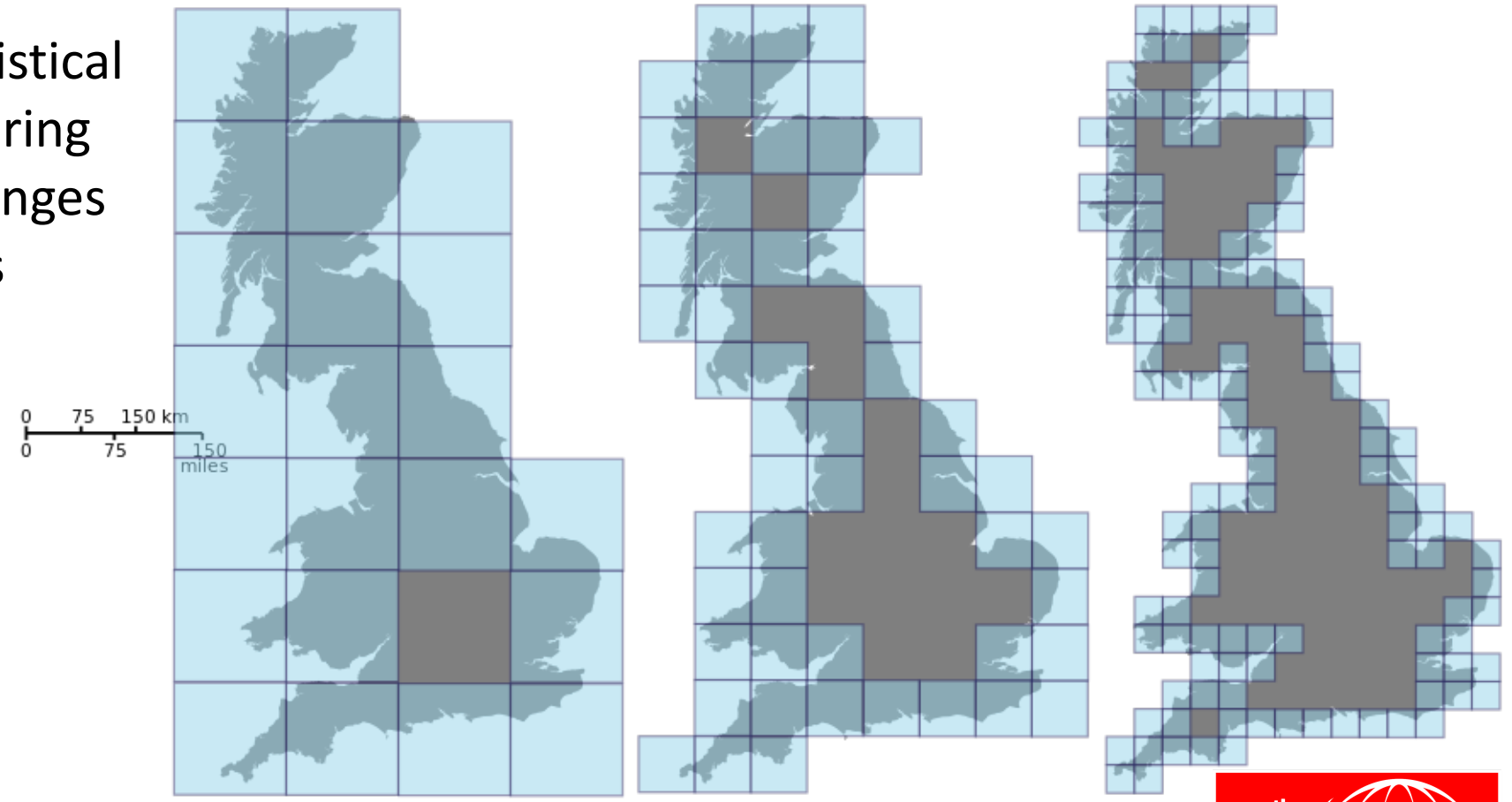
Entropy is a scientific concept, as well as a measurable physical property that is most commonly associated with a state of disorder, randomness, or uncertainty.



Paul, et al. 2018. Multiscale entropy rate analysis of complex mobile agents. Royal Society open science 5, 10 (2018), 180488.

Fractal Dimension

Fractal Dimension is a statistical index of complexity comparing how detail in a pattern changes with the scale at which it is measured.



Zhang R, Stanley K, **Fuller D,** Bell S. Differentiating Population Spatial Behavior using Representative Features of Geospatial Mobility (ReFGeM). Transactions on Spatial Algorithms and Systems. 2020:6(1). https://dl.acm.org/doi/10.1145/3362063

Machine learning and fancy stuff

- Trip and dwell detection
  - Thierry et al. 2013
  - Broach 2019
- Feasible to do with most if not all GPS data. Tools exist to help with this. Gets you cleaner data that you can them calculate for example how active in terms of minutes of MVPA each trips in persons day ways.

Machine learning and fancy stuff

- Transportation Mode Detection
  - Many approaches to this
    - Roy et al 2020
    - Feng et al 2013
  - Hard to implement because
    - Your data will not be the same
    - You will have different transportation modes in your city
    - How people move in your city is different
  - Fast moving research area. Maybe generalizable methods in 1-2 years.

Why can’t we just automate this?

- The data from different devices is not the same
- Data type matters a LOT!
  - Accel
    - Different frequencies of data collection
    - Different wear locations
    - Clock drift
  - GPS
    - Varying frequencies of data collection with the study
    - Different locations detection methods

  - Objectives are different so not everyone wants the same thing
  - Cleaning is decision making. Joining is decision making.

Conclusions

- It seems easy but it’s not.
- Many, many decisions and trade offs
- We need to do better publishing detail methods papers in this area

Acknowledgments

- https://teaminteract.ca/
- http://www.beaplab.com/home/
- @walkabilly
- dfuller@mun.ca