Best Practices in Data Reduction & Analysis of GPS Data

Jacqueline Kerr
Fred Raab
Ernesto Ramirez

University of California, San Diego
Audience Poll

• Who has collected GPS data?
• Who has analyzed or mapped the data?
Context of GPS work

- Older adults walking routes
- TEAN – adolescents travel
- GEI project: GEI/GPS/GIS workshop
- PALMS participants & interviews
- PALMS system
  - Calculations
Multiple data sources

• GPS (vs pedometer)
  – Walking/biking
  – Location
  – Time
  – Low cost
• Time stamped (survey) data
• Photos
• Accelerometer
• Heart rate
• Microscale
• GIS
GPS data process

- Data collection
  - Best practice for quality data
- Downloading
- Data checking/validation
  - Participant recall
- Cleaning & processing
- Visualization
- Variables
- Analyses & models
- Next Steps

• Additional information: handouts/roundtable
Sources of GPS Error
Terrestrial

Also:
• Tree coverage when outdoors
• Electronic interference

www.garmin.com
Sources of GPS Error

Atmospheric Delays

Position of the satellite in the sky effects accuracy; the higher the better.

Other sources of inference include sunspots and water vapor.

www.geoplane.com
GPS Issues
Jitter and Building Interference

Jitter while in building

Building interference while outside
GPS Issues
Urban Canyons
GPS Issues
Time to Acquire Fix
GPS Issues - Human

• Forgetfulness
  – Keeping device charged
  – Powering device on / off
  – Keeping in bag instead of on person

• Two device problem – GPS gets separated from accelerometer

• Leaving GPS in car

• Auto travel before acquiring fix
GPS Issues – Human
The Parking Lot Problem

Subject leaves GPS in car over multiple trips.
GPS Mitigations - Human

• Keeping device charged
  – GPS with battery life > 20 hours
  – Reminder messages via email or text

• Powering on / off
  – Instruction on meaning of LED indicators

• Keeping on person, leaving in car
  – Make devices as wearable as possible

• Two device problem
  – One holder for both devices if possible

• Time to acquire fix
  – Instruction on meaning of LED indicators
Configuration Decisions

- What data to save
- Sampling rate
- On device filtering
BT-335 Tradeoffs

• 1.4 MB Memory can store:
  – 100,000 position only
  – 62,000 position, date, time, speed
  – 43,000 position, date, time, speed, elevation

• Assuming 43,000 entries:
  – 15 seconds = 179 hrs (7.4 days)
  – 20 seconds = 238 hrs (9.9 days)
  – 60 seconds = 716 hrs (29.8 days)

• Problem with on-device filtering
  – Are gaps in time stamps caused by filtering or by loss of signal (or loss of power)? Was person really at the “filtered” location?
Downloading Data from GPS

Which file format(s) to use?
GPS File Formats

- Proprietary
  - Garmin, US GlobalSat, numerous others
- Standard
  - GPX
  - NMEA (National Marine Electronics Association)
- Other
  - CSV, TXT – columns are vendor-specific
  - KML – encoding is vendor-specific

Recommendation: Save in both vendor’s format and GPX
GPX Example

<trk>
  <name>174,2008-10-02:12:28:23</name>
  <trkseg>
    <trkpt lat="32.87053" lon="-117.23485">
      <ele>137.0</ele><time>2008-09-30T17:55:26Z</time>
    </trkpt>

    <trkpt lat="32.87093" lon="-117.23480">
      <ele>112.0</ele><time>2008-09-30T17:55:38Z</time>
    </trkpt>
  </trkseg>
</trk>

Types of GPS points

• Tracks and Trackpoints
  – Tracks consist of track segments which contain trackpoints
  – Typically used for a moving location
  – Lat/Lon, elevation, timestamp, speed, course
  – Most often displayed as a line segment connecting the trackpoints

• Waypoints
  – Typically used for a fixed location
  – Lat/Lon, elevation, creation / modification timestamp, name, type, icon
  – Displayed as a circle or specified icon
Types of GPS points

• Routes
  – Set of waypoints to help you navigate from one point to another.
  – Navigation systems compute routes dynamically

• Track logs
  – Some handheld GPS devices offer the option of saving track logs on the device. BEWARE: timestamps may be removed to save space.
Lat/Lon Formats

• Degree / Minutes / Seconds (DMS)
  38° 24’ 57.63” N  122° 50’ 27.34” W
  (one second = 100 feet)

• Degree / Decimal Minutes
  38° 24.96’  -122° 50.45’

• Decimal Degrees
  38.4160  -122.8381
Time Formats

• GPS time typically reported in UTC 2008-09-30T17:55:26Z
• Does not automatically adjust for timezone or daylight savings
• GPS hardware reports:
  – Date in year, month, day
  – Time in hours, minutes, seconds, milliseconds
  – Milliseconds often truncated by software
• NOTE: Other devices (accelerometers, heart rate monitors) report in subject’s local time
Other GPS Software

EasyGPS

Works directly with Garmin & Magellan hardware units

www.easygps.com
GPS Software Tools

GPSBabel

Translates between GPS file formats

Can also filter data

www.gpsbabel.org
Geotagging

• Adds GPS data to digital photos (.jpg files)
• Data stored in Exif header
• Matches timestamps
• Available as desktop program and web applications

Locr GPS Photo software included with BT-334
www.locr.com
Displaying Geotagged Photos

Flicker.com  Picasaweb.google.com
GPS Data Processing

• Algorithms
  – Derived values calculations
    • Time (duration)
    • Distance
    • Speed
    • Elevation delta
    • Grade
  – Filtering
  – Clustering
  – Trip detection
  – Location detection
  – Mode of transportation
  – Indoors / outdoors
Using interviews to check data & develop algorithms

- Reviewed raw track logs (displayed in Google Earth) and raw accelerometer / heart rate graphs with study participant
- Created a timeline of the participant’s day
- Identified location clusters, trips, and modes of transportation
- Probed periods of peak activity and/or heart rate
- Specifically queried problem areas: gaps in data, GPS in parking lot, odd GPS patterns, indoor activities
GPS Algorithms
Derived Values

Given two fixes F1, F2 with time, lat, lon, ele

Distance = great_circle(F1,F2)
Duration = F2.time – F1.time
Elevation Delta = F2.ele – F1.ele
Speed = Distance / Duration
Grade = Elevation Delta / Distance
 (change of 6m over 100m = 6% grade)
GPS Algorithms
Filtering

Remove extraneous data points caused by jitter when stationary

Color coded by speed
GPS Algorithms
Filtering

Goals:

• Identify points 1-5 as stationary fixes at one location and optionally map to one point

• Identify point 6 as start of trip and points 7-10 as points on the trip
If $D(n,n+1) < \text{minimum}\_\text{distance}$, then remove $n+1$

If $D(n,n+2) < \text{minimum}\_\text{distance}$, then remove $n+1$

Other GPS errors typically seen:

- Improbable change in elevation
- Improbable speed
GPS Algorithms
Clustering

Detect clusters of activity, time spent, most visited, etc.
GPS Algorithms
Location Detection

First cluster, then map all points within a radius to the cluster centroid
GPS Algorithms
Trip Detection

- Detects starting and stopping locations of trips, as well as short pauses
GPS Algorithms
Mode of Transportation

- Classify trips as walking, running, bicycle, vehicle
- Speed used as initial classifier
- Average trip speed is better
- Assisted by use of accelerometer and/or heart rate data
GPS Algorithms
Geodata Anonymization (points)

- Techniques used to hide subject’s home (office, etc) location

- Creates n points at random within x distance of subject’s actual location

- Show the mean location of the points as the subject’s location
GPS Algorithms

Geodata Anonymization (paths)

Difficult with path data

Other approaches:

- If physical location is not important, generate a random offset to apply to lat/lon values.

Source: Inference Attacks on Location Tracks, John Krumm, Microsoft Research, Pervasive 2007

Figure 5: This demonstrates the effect of our three methods of corrupting GPS data to enhance privacy on a set of data from one of our subjects. The upper left image shows the raw, uncorrupted GPS data. The white circle in each image shows the location of the subject’s house.
Time & space

- Physical activity
- Travel
- Environmental context
- Location

- Points
- Lines
- Polygons

TIME

Microscale attributes
Case studies and research questions

• Participatory photo mapping (Dennis & Gaulocher)
• Activity levels in key locations, school, work, home, neighborhoods (Rodriguez & Cohen, TAAG study)
• Active commuting behavior (Sallis et al. TEAN study)
• Route taken to school (Roemmich)
• Use of park facilities (Troped)
• Avoidance behavior (Patla & Frank)
• Behavior before and after hip surgery (Chang)
• Time spent being active near freeway (Avol)
• SES influences on distance traveled (Matthews)
Types of data from GPS

- Time spent generally
- Time spent in specific locations
  - With attributes
- Time spent in traveling
  - With attributes
- Travel mode
- Activity locations
Data Merging

• Need to merge GPS data with sensor data (typically by matching timestamps)
  – Add location, elevation to sensor record
  – Add sensor data to GPS tracklog

• Choice depends on your objectives
  – Identify events (i.e. bouts of activity), then add GPS fixes to events
  – Process GPS fixes into locations and trips, then match events to GPS fixes
Issues with Data Fusion
How to match epochs and adjust for gaps in data?

<table>
<thead>
<tr>
<th>Inputs</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>GPS Samples</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accelerometer</td>
<td>y</td>
<td>y</td>
<td>y</td>
<td>y</td>
<td>y</td>
<td>y</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outputs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPS to acc</td>
<td>yx</td>
<td>yx</td>
<td>yx</td>
<td>y</td>
<td>yk</td>
<td>yx</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acc to GPS</td>
<td>x</td>
<td>xy</td>
<td>x</td>
<td>xy</td>
<td>x</td>
<td>xy</td>
</tr>
<tr>
<td></td>
<td>x</td>
<td>xy</td>
<td>x</td>
<td>xy</td>
<td>x</td>
<td>xy</td>
</tr>
<tr>
<td></td>
<td>xy</td>
<td>x</td>
<td>xy</td>
<td>x</td>
<td>x</td>
<td>xy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avg Acc to GPS</td>
<td>xy/2</td>
<td>xy/2</td>
<td>xy/2</td>
<td>xy/2</td>
<td>xy/2</td>
<td>xy/2</td>
</tr>
<tr>
<td></td>
<td>x2y</td>
<td>xy</td>
<td>xy/2</td>
<td>xy/2</td>
<td>xy/2</td>
<td>xy/2</td>
</tr>
<tr>
<td>Timeline</td>
<td>xy/2</td>
<td>xy/2</td>
<td>xy/2</td>
<td>xy/2</td>
<td>xy/2</td>
<td>y/2</td>
</tr>
<tr>
<td></td>
<td>x</td>
<td>xy/2</td>
<td>xy/2</td>
<td>xy/2</td>
<td>xy/2</td>
<td>y/2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Case 1: Match GPS location to accelerometer reading
- Case 2: Match accelerometer reading to GPS location
- Case 3: Match average accelerometer value to GPS location
- Case 4: Establish a time line and match GPS locations and average accelerometer values to timeline
- In all cases, need to establish rules and cutoffs for missing data
Visualizing GPS Data

- **Vendor Software**
  - Good for verifying data as received from subject
  - Typically designed to display track of one subject at a time

- **Google Earth**
  - Will directly import and display GPX file (but doesn’t provide much more functionality than vendor software)
  - Real power is displaying KML files, but you need an application to create the KML file

- **Google Map (out of the box)**
  - Needs a KML file to display
  - KML file resides on Google server – potential privacy concerns
  - Google Map can be used as the mapping engine in a custom web application without the need to send data to Google
Visualizing GPS Data

Websites
www.geovisualizer.com

- Provides some filtering and analysis functions
- Can process data from multiple subjects
- Typically has a data size limitation
- Potential privacy concerns
- Good for prototyping with non-sensitive data
Visualizing GPS Data

GIS Systems

ESRI ArcGIS
Quantum GIS

- Desktop systems – data never leaves your control
- Can process large data sets from multiple subjects
- Many analytical functions available
- Many datasets (layers) publicly available detailing the natural and built environment
GPS Processing

openGTS

Open GPS Tracking System (OpenGTS)

This is a free demonstration of the features of the OpenGTS open-source project available on SourceForge. For more information regarding the OpenGTS project (including download links), visit the project web-site at http://www.opengts.org/

Account: Demo Account (OpenGTS User)

Event Detail

<table>
<thead>
<tr>
<th>#</th>
<th>Date</th>
<th>Time</th>
<th>Status</th>
<th>Lat</th>
<th>Lon</th>
<th>Speed mph</th>
<th>Altitude ft</th>
<th>Odometer Miles</th>
<th>Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2007/03/16</td>
<td>11:08:57</td>
<td>Start</td>
<td>38.6487</td>
<td>-121.8028</td>
<td>11.7 m</td>
<td>23</td>
<td></td>
<td>480 North Highlands, CA</td>
</tr>
<tr>
<td>2</td>
<td>2007/03/16</td>
<td>11:11:44</td>
<td>InMotion</td>
<td>38.6384</td>
<td>-121.9312</td>
<td>41.4 m</td>
<td>14</td>
<td></td>
<td>100, Sacramento, CA</td>
</tr>
<tr>
<td>3</td>
<td>2007/03/16</td>
<td>11:18:50</td>
<td>InMotion</td>
<td>38.6755</td>
<td>-121.8702</td>
<td>64.5 m</td>
<td>14</td>
<td></td>
<td>480 West Sacramento, CA</td>
</tr>
<tr>
<td>4</td>
<td>2007/03/16</td>
<td>11:20:58</td>
<td>InMotion</td>
<td>38.6868</td>
<td>-121.6781</td>
<td>66.6 m</td>
<td>16</td>
<td></td>
<td>45217 E Chiles Rd, University of California Davis Campus, CA 95616</td>
</tr>
<tr>
<td>5</td>
<td>2007/03/16</td>
<td>11:26:05</td>
<td>InMotion</td>
<td>38.5152</td>
<td>-121.7559</td>
<td>65.2 m</td>
<td>33</td>
<td></td>
<td>Dixon, CA</td>
</tr>
<tr>
<td>6</td>
<td>2007/03/16</td>
<td>11:31:18</td>
<td>InMotion</td>
<td>38.4265</td>
<td>-121.8850</td>
<td>66.9 m</td>
<td>32</td>
<td></td>
<td>Dixon, CA</td>
</tr>
<tr>
<td>7</td>
<td>2007/03/16</td>
<td>11:36:26</td>
<td>InMotion</td>
<td>38.3820</td>
<td>-121.9426</td>
<td>64.0 m</td>
<td>92</td>
<td></td>
<td>Dixon, CA</td>
</tr>
<tr>
<td>8</td>
<td>2007/03/16</td>
<td>11:41:34</td>
<td>InMotion</td>
<td>38.3225</td>
<td>-122.0259</td>
<td>65.1 m</td>
<td>42</td>
<td></td>
<td>Dixon, CA</td>
</tr>
<tr>
<td>9</td>
<td>2007/03/16</td>
<td>11:46:40</td>
<td>InMotion</td>
<td>38.2446</td>
<td>-122.0820</td>
<td>65.1 m</td>
<td>92</td>
<td></td>
<td>Dixon, CA</td>
</tr>
<tr>
<td>10</td>
<td>2007/03/16</td>
<td>11:51:44</td>
<td>InMotion</td>
<td>38.1189</td>
<td>-122.1309</td>
<td>66.6 m</td>
<td>16</td>
<td></td>
<td>Dixon, CA</td>
</tr>
<tr>
<td>11</td>
<td>2007/03/16</td>
<td>11:56:02</td>
<td>InMotion</td>
<td>38.1189</td>
<td>-122.1309</td>
<td>66.6 m</td>
<td>16</td>
<td></td>
<td>Dixon, CA</td>
</tr>
<tr>
<td>12</td>
<td>2007/03/16</td>
<td>12:01:59</td>
<td>InMotion</td>
<td>38.0987</td>
<td>-122.2478</td>
<td>66.6 m</td>
<td>16</td>
<td></td>
<td>Dixon, CA</td>
</tr>
<tr>
<td>13</td>
<td>2007/03/16</td>
<td>12:07:08</td>
<td>InMotion</td>
<td>37.9756</td>
<td>-122.1817</td>
<td>66.6 m</td>
<td>16</td>
<td></td>
<td>Dixon, CA</td>
</tr>
<tr>
<td>14</td>
<td>2007/03/16</td>
<td>12:12:14</td>
<td>InMotion</td>
<td>37.9973</td>
<td>-122.3091</td>
<td>66.6 m</td>
<td>16</td>
<td></td>
<td>Dixon, CA</td>
</tr>
<tr>
<td>15</td>
<td>2007/03/16</td>
<td>12:17:20</td>
<td>InMotion</td>
<td>37.8223</td>
<td>-122.3229</td>
<td>66.6 m</td>
<td>16</td>
<td></td>
<td>Dixon, CA</td>
</tr>
<tr>
<td>16</td>
<td>2007/03/16</td>
<td>12:22:24</td>
<td>InMotion</td>
<td>37.8077</td>
<td>-122.3677</td>
<td>66.6 m</td>
<td>16</td>
<td></td>
<td>Dixon, CA</td>
</tr>
<tr>
<td>17</td>
<td>2007/03/16</td>
<td>12:27:32</td>
<td>InMotion</td>
<td>37.7889</td>
<td>-122.3879</td>
<td>66.6 m</td>
<td>16</td>
<td></td>
<td>Dixon, CA</td>
</tr>
<tr>
<td>18</td>
<td>2007/03/16</td>
<td>12:32:32</td>
<td>InMotion</td>
<td>37.7816</td>
<td>-122.3994</td>
<td>66.6 m</td>
<td>16</td>
<td></td>
<td>Dixon, CA</td>
</tr>
<tr>
<td>19</td>
<td>2007/03/16</td>
<td>12:37:54</td>
<td>InMotion</td>
<td>37.7855</td>
<td>-122.4002</td>
<td>66.6 m</td>
<td>16</td>
<td></td>
<td>Dixon, CA</td>
</tr>
<tr>
<td>20</td>
<td>2007/03/16</td>
<td>12:42:40</td>
<td>InMotion</td>
<td>37.7855</td>
<td>-122.4002</td>
<td>66.6 m</td>
<td>16</td>
<td></td>
<td>Dixon, CA</td>
</tr>
<tr>
<td>21</td>
<td>2007/03/16</td>
<td>12:47:41</td>
<td>InMotion</td>
<td>37.7834</td>
<td>-122.4026</td>
<td>66.6 m</td>
<td>16</td>
<td></td>
<td>Dixon, CA</td>
</tr>
<tr>
<td>22</td>
<td>2007/03/16</td>
<td>12:52:46</td>
<td>Stop</td>
<td>37.7847</td>
<td>-122.3981</td>
<td>66.6 m</td>
<td>16</td>
<td></td>
<td>Dixon, CA</td>
</tr>
</tbody>
</table>
Visualization Examples - Population Heat Maps
Visualization Examples - Population
3D Overlay
Figure 2
Example of space-time path in GIS.
Figure 3
Time density map of Marion County by percent of black or Latino population.
Specific location: park

- Park location
  - Where in park
- Frequency of visit to park
- Time spent in park
  - Total/min/max/average
  - % of total time
- Time of day, day of week
- Time/location since last visit (latencies)
- Attributes of park
- Intensity of activity in park
  - When, where, how long
  - Bouts/ guidelines
Travel to park

- Start location, time
- Journey time
- Route taken
- Distance traveled
- Speed traveled
- Elevation
- Mode of transport
- Energy expenditure
Context of park visit

• Why this park?
  – Was it the closest?
  – Within a neighborhood buffer?

• Why this route?
  – Was it shortest?
  – How do you attribute GIS characteristics to the route?

• Where stop on way?

• What exposures?
Day in a life

- Percent time in home neighborhood
- Distance traveled
- Time spent in car
- Time spent outside/indoors
- Least/most active locations
- Exposure history
Research methods questions arising from GPS data

• Sampling
  – How many people, how long?
• Missing data
• Distance & time spent outside of home neighborhood
  – What buffers do we use?
  – How sample locations for quality data?
  – Definitions of neighborhood

• Points, polygons, lines and TIME
  – Point data
    • Point pattern analysis used for incidents
    • Clusters
  – Moving point data
    • Agent based modeling
  – Multiple data points
    • Time points nested within people
    • Intensive longitudinal data analysis: multiple time series analysis
What models & frameworks can we use?

- Spatial analysis
  - Spatial interaction or "gravity models" estimate the flow of people between locations in geo-space
  - Kwan M: geovisualization; time-geographic approach
  - Geographically weighted regression

- Artificial Neural Networks/ Supervised Learning Algorithms /Complex Adaptive Systems

- Agent Based Modeling

- Pattern Analysis/Recognition

- Markov Model (to predict movement (in time))

- State-Space Approach

- Ecological Momentary Assessment

3Gs Workshop & report

- Current state of field
- Challenges
- Best practice
- What is holding field back
- Next steps, research questions & technical expectations

- Location based prompts
What questions do you have?

• We will try to raise unanswered questions at the 3Gs workshop with the “uber” experts.

• ALR
  – Active Learning Research
  – New transdisciplinary partners

• Roundtable
  – Thurs 7.30am (!)

• What research questions have you attempted to answer?
• How have you analyzed or mapped the data?
Additional Slides
Some references

GPS Systems: Data Logger

GlobalSat DG-100

- Cost: $70
- Battery Life: 25 hours
- Memory: 60,000 points
- Data Capture Settings
  - Epochs
  - Radius
  - Speed
- Data Export:
  - KML/Google Maps
  - Text
  - Excel file
  - RMC
  - GPX format
GPS Systems: Data Logger

 Globalsat BT-335

• Cost: $70
• Battery Life: 25 hours
• Memory: 60,000 points
• Data Capture Settings
  o Epochs
  o Radius
  o Speed
• Data Export:
  o KML/Google Maps
  o Text
  o Excel file
  o RMC
  o GPX format
• Bluetooth compatible
GPS Systems: Wearable

Garmin Forerunner 405
- Battery Life: 8 hrs of data capture
- Cost: $300-350
- Displays continuous user data
- Data export
  - Garmin Software
  - Garmin web-based application
  - Excel
  - KML
- Easily wearable: watch style
- Wireless Data Download (ANT)
GPS : Vendor Websites

GPS Vendors
GlobalSat GPS and Dataloggers
  • www.globalstat.com.tw/eng/index.htm
GPS Dataloggers
  • www.semsons.com/datalogger.html
Garmin
  • www.garmin.com
How GPS Works

GPS receiver uses triangulation to determine location.
How GPS Works

1. All satellites have clock set to exactly the same time
2. All satellites know their exact position from data sent to them from the system controllers
3. Each satellite transmits its position and a time signal
4. The signals travel to the receiver delayed by distance traveled
5. The differences in distance traveled make each satellite appear to have a different time
6. The receiver calculates the distance to each satellite and can then calculate its own position

www.aero.org
GPS Accuracy
Best Case

SA – Selective Availability – error injected by DOD, turned off 5/1/2000

WAAS – Wide Area Augmentation System, turned on 2003

garmin
GPS Issues – Accuracy Testing at Fixed Location

- Six Garmin Foretrex 201 GPS
- Each recorded 121 observations during 1 hour period (n = 726)
- Average distance from geodetic point: 3.02 m (SD 2.51)
- 81% within 5 meters
- 99.4% within 15 meters

Daniel Rodriguez & Elizabeth Shay, 2008 ALR Conference Presentation
GPS Issues – Accuracy
In the field

• Wore units along route in three built environment scenarios
  – Open space, clustered development and urban
  – Route within each scenario had 25 locations
  – Each scenario was tested three times

• Inter-unit reliability
  – Comparison of the distance between each unit’s recorded location and the average location recorded by the other 5 units.

Daniel Rodriguez & Elizabeth Shay, 2008 ALR Conference Presentation
GPS Issues – Accuracy
In the field

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Average distance (m)</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open space</td>
<td>10.7</td>
<td>11.9</td>
<td>450</td>
</tr>
<tr>
<td>Clustered development</td>
<td>20.1</td>
<td>21.8</td>
<td>450</td>
</tr>
<tr>
<td>Urban</td>
<td>18.5</td>
<td>18.4</td>
<td>450</td>
</tr>
</tbody>
</table>

Daniel Rodriguez & Elizabeth Shay, 2008 ALR Conference Presentation
UCSD Physical Activity Measurement System
PALMS

Physical Activity Measurement Sensor

Visualization/Data Analysis

System Architecture

Data analysis

Wearable environmental sensors

Portable GPS devices

User Interface

GIS

Visualization

Cell phones and Ecological momentary assessments

Accelerometers
PALMS GPS processing

• Trap within participant-provided locations (home, work, market, etc.)
  If A is within X meters of B, then A = B

• Filter GPS errors
  If distance between A & B > X meters
  AND distance between A & C < Y meters,
  then mark B as invalid
PALMS GPS Processing

- Detect location clusters and trap within centroid
- Detect trips: start, end, pause
- Classify trips based on speed walking, jogging, running, vehicle (heart rate aids in classification)
- Attempt to fill-in gaps (typically when user is in a building.)
PALMS accelerometer processing

- Detect non-wearing conditions defined as a # of zeros in a row
- Detect sedentary periods
- Detect bouts of activity counts > X for at least Y minutes allowing for Z minutes of counts < X
PALMS heartrate processing

- Detect non-wearing conditions defined as a # of zeros in a row
- Detect and mark noise
- Replace noise with average HR
- Compute Energy Expenditure (ee)
- Detect bouts of activity
  HR > X for at least Y minutes allowing for Z minutes of counts < X
PALMS Data Fusion

• Create timeframe based on:
  – Starting date
  – Ending date
  – Interval

• Align GPS, accelerometer, heart rate data timestamps to timeframe

• Detect and mark gaps in data
PALMS Data Summary

• Summarize data by
  – Day
  – Hour
  – Location
  – Mode of transportation
  – Activity levels
  – Duration
  – Etc
PALMS Data Analysis (under development)

• Issues:
  – How much statistical / spatial analysis to do within PALMS vs GIS / SPSS?
  – How to merge data from multiple participants in a meaningful way?
PALMS Export

• Export of raw data, processed data or merged data as KML (Google Earth) or CSV (ArcGIS)
GPS Algorithms
Distance using Great Circle Formula

- Radius_Feet = 20889108
- Radius_KM = 6367
- $a = \text{toradians}(90 - F1.\text{lat})$
- $b = \text{toradians}(90 - F2.\text{lat})$
- $t = \text{toradians}(F2.\text{lon} - F1.\text{lon})$
- $c = \text{acos}(\cos(a) \times \cos(b) + \sin(a) \times \sin(b) \times \cos(t))$
- Distance = $c \times \text{Radius}_{xx}$

www.mappinghacks.com/project/distance
GPS Algorithms
Problem with Great Circle Formula

Assumes Earth is a perfect sphere with radius of 6367km
Actual equatorial radius = 6378
Actual polar radius = 6356
Difference of 22km
Error using 6367 = 0.00175

Mapping Hacks, pg 119-120