Mobile Data Collection:
Lessons from the Escambia County Bus Stop Inventory

Martin Catalá
GIS Manager
Center for Urban Transportation Research
4202 East Fowler Avenue (CUT100)
Tampa, Florida 33620
http://www.cutr.usf.edu
catala@cutr.usf.edu
ECAT BSI

• Project Goal
  – Create a bus stop inventory (BSI)
  – Survey riders using the collected BSI

• Purpose
  – Establish accurate baseline ridership data
  – Identify opportunities for operations savings
  – Realignment of routes
  – Prioritize stop improvements
Project Tasks

- Determine data collection methodology
- Determine hardware and GPS
- Determine data collection needs
  - What data to collect in addition to location
- Create ArcPAD forms
- Establish a methodology for systematically collecting the route data
  - Avoid duplication of effort
- Collect the data
Project Tasks

• Collect the data
• Analyze the data (QC)
• Create route and trip specific ridecheck sheets from the Bus Stop inventory for surveyors
• Enter ridecheck data into database
• Join ridecheck data with GIS BSI
• Perform analysis
ECAT Transit System

- 33 Fixed Routes
- UWF and Beach Trolley Services
- 1,516 Stops
- 1.6 million passenger trips annually
Selecting a Methodology

- Three evolutions of Collection efforts at CUTR
  - “Old School”
    - Hand write data collection
  - “Not So Old School”
    - Hand write GPS coordinates into paper form
  - “New School”
    - In-Field GIS data collection
“Old School” Methodology

• Old School
  • Recorded on paper only with cross street and distance from cross street information recorded
    - Location and amenities written down
  • Heads up geocoding performed from paper survey
  • Data manually entered
  • Data and stop information joined and mapped
“Not so old School” Methodology

• Not so old school
  • All amenities are recorded on paper
  • Location data acquired by GPS units
    – Coordinates written down on forms
  • Data manually entered
    – Stops created from GPS data
  • Data and stop information joined and mapped
Old School Benefits - Drawbacks

• Benefits
  – Simple
    • Little training needed
    • Minimally reliable on technology
  – Hard copy paper trail

• Drawbacks
  – Time consuming
  – Multiple opportunities for errors
    • Collection, data entry, transferring data into GIS
“New School” Methodology

• New school
  • All amenities are recorded on PocketPC device
  • Location data acquired by GPS units
    – Coordinates calculated and imported into GIS database
• Attribute data collected on mobile device
• Data directly imported into GIS
New School Benefits - Drawbacks

• Benefits
  – Collect data once
  – Accurate and uniform data collection
    • Forms
  – Directly importable into ArcGIS
  – More Timely
New School Benefits - Drawbacks

• Drawbacks
  – Not simple
  • Required programming time with ArcPAD forms
  • Hardware field testing
    – Groundtruthing
    – Communication glitches with Bluetooth technology
  • Not friendly to the technology / GIS challenged
  • More time for training
    – Software
    – Data descriptions (not different than old method)
New School Benefits - Drawbacks

• **Drawbacks**
  - The promise of technology
  - Changing data “needs”
  - Never-ending list of “new” fields and other data
    • Lost some efficiencies
    • (i.e. property names, roadway classification, speed limit, number of lanes, lane width)
Methodology - New

- New Method
  - In-field data collection
  - GPS with Bluetooth technology
  - PocketPC with ArcPad 6.0
  - ArcPad forms
Selecting Software and Hardware

• University Site license
  – ArcPAD 6.0
  – ArcGIS

• Using specifications from ESRI
  – Recommended hardware bundle

• HP PocketPC and Bluetooth enabled GPS
  – Wireless GPS with Pocket PC connection
Determining BSI Data Needs

• Unique ID (MUST HAVE)

• Stop Amenities
  – Shelter
    • Type of shelter
  – Bench
    • Type of bench
  – Sign
    • Condition
    • Type
BSI Data Needs

• Stop Conditions
  – Sidewalk
    • Width
    • Condition
  – Street Light
  – Curb cut
  – ADA (Americans with Disability Act) Accessible
    • Description of ADA features
    • Description of ADA impediments
BSI Data Needs

• X-Coordinate
• Y-Coordinate
• Adjacent land-use
• Property description
  – Business name
    • Used for ridecheck
• Comments
  – Text field for manual comments
Creating ArcPAD Forms

• ArcPAD Application Developer
  – Packaged with ArcPAD 6.0

• GIS Analyst
  – On-line resources
  – Documentation

• Significant learning curve
  – 40-60 hours creating basic forms
  – 40 hours for testing forms
Data Collection Methodology

• Multiple teams or Single team
  – Multiple requires coordination to avoid duplication
  – Single team requires less coordination but effort should be made for optimize collection effort

• Multiple teams
  – Logistics for route corridors
Collecting the Data

- Two Teams of two
  - One driver, one recorder and navigator
- Each had a printed map
  - Directions for streets not to collect beyond
    - i.e. do not go north of Fairfield Ave. on Palafax St.
- Two days designated to collect data
  - Team One collected 158 stops on day one
  - Team Two collected 93 stops on day one
Collecting Data

- End of Two days of collecting data
  - Nearly 500 stops collected
    - Only one-third

- Three more days allocated to collect data
  - At end of 5 total days
    - Over 1500 stops collected with 43 columns of data
Problems encountered

- Hurricane Ivan
  Sept. 16, 2004
Problems

• Political, administrative and fiscal environment
  - Bus Stop Inventory needed to be complete in order to conduct ridership survey
• Ridership survey used for Transit Development Plan (TDP) and Comprehensive Operations Analysis (COA)
• Deadline did not change for TDP and COA
Problems

- Problems encountered
  - Many stops missing, destroyed or damaged
  - Too much time to collect data at each stop
  - Too many variables (for time allocated)
  - GPS drops in urbanized area
  - Better base map for when GPS signal dropped
Problems

- Coding of the forms
  - Using radio buttons require a binary field type
  - Fast for in field coding
  - Problem with ArcGIS
    - Binary format not supported
    - Transferring to ArcGIS fine
    - Bringing data (Shp) back to ArcPAD problem for updating
- Use more pull downs for bench and sign types
Problems

• Frequent GPS drops in more urban areas
  – Also around communication towers
    • (coincidence?)
  – Base street database was TIGER
    • Appearance that the GPS was inaccurate
  – Did not include Aerials or parcel data for base map
    • Would help with heads-up geocoding when GPS dropped
Creating Ride Check Sheets

- Each route needed a separate ride check sheet for each trip

<table>
<thead>
<tr>
<th>Time</th>
<th>Route 1</th>
<th>Route 2</th>
<th>Route 3</th>
<th>Route 4</th>
<th>Route 5</th>
<th>Route 6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ECAT</td>
<td>Davis Hwy</td>
<td>Brent Lane</td>
<td>West Florida Hospital</td>
<td>Davis Hwy</td>
<td>Brent Lane</td>
</tr>
<tr>
<td>5:45</td>
<td>5:49</td>
<td>5:54</td>
<td>6:05</td>
<td>6:12</td>
<td>6:20</td>
<td>6:25</td>
</tr>
<tr>
<td>6:45</td>
<td>6:49</td>
<td>6:54</td>
<td>7:05</td>
<td>7:12</td>
<td>7:20</td>
<td>7:25</td>
</tr>
<tr>
<td>7:45</td>
<td>7:49</td>
<td>7:54</td>
<td>8:05</td>
<td>8:12</td>
<td>8:20</td>
<td>8:25</td>
</tr>
<tr>
<td>8:45</td>
<td>8:49</td>
<td>9:05</td>
<td>9:12</td>
<td>9:20</td>
<td>9:25</td>
<td>9:40</td>
</tr>
<tr>
<td>10:45</td>
<td>10:49</td>
<td>11:05</td>
<td>11:12</td>
<td>11:20</td>
<td>11:25</td>
<td>11:40</td>
</tr>
<tr>
<td>11:45</td>
<td>11:49</td>
<td>12:05</td>
<td>12:12</td>
<td>12:20</td>
<td>12:25</td>
<td>12:40</td>
</tr>
<tr>
<td>12:45</td>
<td>12:49</td>
<td>1:05</td>
<td>1:12</td>
<td>1:20</td>
<td>1:25</td>
<td>1:40</td>
</tr>
<tr>
<td>1:45</td>
<td>1:49</td>
<td>1:54</td>
<td>2:05</td>
<td>2:12</td>
<td>2:20</td>
<td>2:25</td>
</tr>
<tr>
<td>2:45</td>
<td>2:49</td>
<td>2:54</td>
<td>3:05</td>
<td>3:12</td>
<td>3:20</td>
<td>3:25</td>
</tr>
<tr>
<td>3:45</td>
<td>3:49</td>
<td>4:05</td>
<td>4:12</td>
<td>4:20</td>
<td>4:25</td>
<td>4:40</td>
</tr>
<tr>
<td>4:45</td>
<td>4:49</td>
<td>5:05</td>
<td>5:12</td>
<td>5:20</td>
<td>5:25</td>
<td>5:40</td>
</tr>
</tbody>
</table>
Creating Ride Check Sheets

- Query of the data must be able to produce ridership by:
  - Route
  - Stop
  - Time of Day
  - Trip time
  - Day of the week
  - Time point
<table>
<thead>
<tr>
<th>ID</th>
<th>On Street</th>
<th>Cross Street</th>
<th>On</th>
<th>Off</th>
<th>Total On</th>
<th>Off</th>
<th>ArrvTime</th>
<th>Time Point</th>
<th>Adj. Land Use</th>
<th>Rt_ID</th>
<th>Trip ID</th>
<th>Start Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>251</td>
<td>University Office Blvd</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>D</td>
<td></td>
<td>University Mall (north)</td>
<td>RT_19Wkdy</td>
<td></td>
<td>8:45 AM</td>
</tr>
<tr>
<td>252</td>
<td>University Office Blvd</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>University Mall (south)</td>
<td>RT_19Wkdy</td>
<td></td>
<td>8:45 AM</td>
</tr>
<tr>
<td>253</td>
<td>N Davis Hwy</td>
<td>Creighton Rd</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sears</td>
<td>RT_19Wkdy</td>
<td></td>
<td>8:45 AM</td>
</tr>
<tr>
<td>254</td>
<td>N Davis Hwy</td>
<td>E Burgess Rd</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Compass Bank</td>
<td>RT_19Wkdy</td>
<td></td>
<td>8:45 AM</td>
</tr>
<tr>
<td>259</td>
<td>N Davis Hwy</td>
<td>E Burgess Rd</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Regions Bank</td>
<td>RT_19Wkdy</td>
<td></td>
<td>8:45 AM</td>
</tr>
<tr>
<td>265</td>
<td>N Davis Hwy</td>
<td>6645</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Rooms &amp; More Furniture</td>
<td>RT_19Wkdy</td>
<td></td>
<td>8:45 AM</td>
</tr>
<tr>
<td>266</td>
<td>N Davis Hwy</td>
<td>6601</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Trade Winds Shopping Center</td>
<td>RT_19Wkdy</td>
<td></td>
<td>8:45 AM</td>
</tr>
<tr>
<td>256</td>
<td>N Davis Hwy</td>
<td>Schubart Dr</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RT_19Wkdy</td>
<td></td>
<td>8:45 AM</td>
</tr>
<tr>
<td>257</td>
<td>N Davis Hwy</td>
<td>6235</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hobby Lobby</td>
<td>RT_19Wkdy</td>
<td></td>
<td>8:45 AM</td>
</tr>
<tr>
<td>258</td>
<td>N Davis Hwy</td>
<td>6115</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pensacola Heritage Apartments</td>
<td>RT_19Wkdy</td>
<td></td>
<td>8:45 AM</td>
</tr>
<tr>
<td>458</td>
<td>N Davis Hwy</td>
<td>Fontaine St</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C</td>
<td>Home Depot</td>
<td>RT_19Wkdy</td>
<td></td>
<td>8:45 AM</td>
</tr>
</tbody>
</table>
Ride Check Data

- Imported into Microsoft Access
- Reports for ridership
- Query Statements:

SQL for Weekday Ridership by Route

```sql
SELECT RideCheckData2.Rt_ID,
       SUM(RideCheckData2.[On]) AS SumOfOn,
       SUM(RideCheckData2.[Off]) AS SumOfOff
FROM RideCheckData2
GROUP BY RideCheckData2.Rt_ID
HAVING (((RideCheckData2.Rt_ID) Like "Wkdy"))
ORDER BY RideCheckData2.Rt_ID;
```

<table>
<thead>
<tr>
<th>Rt_ID</th>
<th>SumOfOn</th>
<th>SumOfOff</th>
</tr>
</thead>
<tbody>
<tr>
<td>RT_01Wkdy</td>
<td>68</td>
<td>53</td>
</tr>
<tr>
<td>RT_02Wkdy</td>
<td>132</td>
<td>102</td>
</tr>
<tr>
<td>RT_03Wkdy</td>
<td>73</td>
<td>64</td>
</tr>
<tr>
<td>RT_04Wkdy</td>
<td>150</td>
<td>139</td>
</tr>
<tr>
<td>RT_05Wkdy</td>
<td>222</td>
<td>199</td>
</tr>
<tr>
<td>RT_06Wkdy</td>
<td>139</td>
<td>107</td>
</tr>
<tr>
<td>RT_09Wkdy</td>
<td>164</td>
<td>136</td>
</tr>
<tr>
<td>RT_12Wkdy</td>
<td>179</td>
<td>177</td>
</tr>
<tr>
<td>RT_13Wkdy</td>
<td>92</td>
<td>65</td>
</tr>
<tr>
<td>RT_14Wkdy</td>
<td>236</td>
<td>236</td>
</tr>
<tr>
<td>RT_15Wkdy</td>
<td>304</td>
<td>297</td>
</tr>
<tr>
<td>RT_18Wkdy</td>
<td>87</td>
<td>71</td>
</tr>
<tr>
<td>RT_19Wkdy</td>
<td>159</td>
<td>155</td>
</tr>
<tr>
<td>RT_20Wkdy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RT_21Wkdy</td>
<td>23</td>
<td>22</td>
</tr>
<tr>
<td>RT_22Wkdy</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>RT_23Wkdy</td>
<td>21</td>
<td>20</td>
</tr>
<tr>
<td>RT_31Wkdy</td>
<td>42</td>
<td>42</td>
</tr>
<tr>
<td>RT_32Wkdy</td>
<td>15</td>
<td>13</td>
</tr>
<tr>
<td>RT_33Wkdy</td>
<td>41</td>
<td>34</td>
</tr>
<tr>
<td>RT_34Wkdy</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>RT_35Wkdy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RT07Wkdy</td>
<td>156</td>
<td>116</td>
</tr>
<tr>
<td>RT07BWkdy</td>
<td>51</td>
<td>35</td>
</tr>
<tr>
<td>RT10BWkdy</td>
<td>200</td>
<td>183</td>
</tr>
<tr>
<td>RT10BWkdy</td>
<td>86</td>
<td>73</td>
</tr>
<tr>
<td>RT11BWkdy</td>
<td>157</td>
<td>156</td>
</tr>
<tr>
<td>RT11BWkdy</td>
<td>89</td>
<td>89</td>
</tr>
<tr>
<td>RT16BWkdy</td>
<td>60</td>
<td>67</td>
</tr>
<tr>
<td>RTU11Wkdy</td>
<td>82</td>
<td>69</td>
</tr>
<tr>
<td>RTU12Wkdy</td>
<td>60</td>
<td>37</td>
</tr>
<tr>
<td>RTU21Wkdy</td>
<td>61</td>
<td>43</td>
</tr>
<tr>
<td>RTU22Wkdy</td>
<td>37</td>
<td>22</td>
</tr>
</tbody>
</table>
Ridership Analysis

• Route data combined to create corridor analysis
• Ridership between timepoints
• Help identify opportunities to consolidate routes.
Ridership Route Analysis

- Ridership by Stop and Route
Segment Level Analysis

- Used with corridor analysis
- Justify route consolidation
If we could do it all over

• Spend more time before project securing local parcel and street data
• Include parcels and aerial data for times which the GPS unit did not work
• Create a pull down menu for cross street names
  – Two fields using street file from County
If we could do it all over

- Focus on data that is necessary for project
  - If future needs require more data, schedule time and money for additional data
If we could do it all over

• Establish techniques for maintaining the system
  – Hardware and data all delivered to the client
  – No guidebook created to help with maintenance and use of the GPS and PocketPC