Exploring the Service Delivery of First Responders in Perrysburg, Ohio

Ian Dunn, MPH, GISP
ESRI Health GIS Conference
September, 2015
Run Locations - All CFS Codes
April, 2007 - March, 2015

Run locations have been retrieved from the ProPhoenix database. The locations and times for data prior to April, 2007 are not recorded in a consistent manner and therefore are not included.

All Fire Division CFS codes (4100 - 4199) have been plotted on the map. Lat/long values for each run location are assigned in ProPhoenix before the data is exported.
✓ **ESRI Kernel Density tool**

✓ **Continuous surface**

✓ **Classification scheme**

Max Run Density Per Year

Run locations and times for data prior to April 2007 are not recorded in a consistent manner and therefore are not included.

All Fire Districts 113 codes (1109 - 1199) have been plotted on the map. Legend values for each run location are assigned into categories per year and mapped on the chart.
Run density is calculated using the Kernel Density tool. The more runs that occur in close proximity, the higher the run density and the darker the red. This data is classified using the April, 2013 – April, 2014 classification scheme.
Run Density - All CFS
April, 2008 - April, 2009

Run density is calculated using the Kernel Density tool. The more runs that occur in close proximity, the higher the run density and the darker the red. This data is classified using the April, 2013 – April, 2014 classification scheme.
Run Density - All CFS
April, 2009 - April, 2010

Run density is calculated using the Kernel Density tool. The more runs that occur in close proximity, the higher the run density and the darker the red. This data is classified using the April, 2013 – April, 2014 classification scheme.
Run Density - All CFS
April, 2010 - April, 2011

Run density is calculated using the Kernel Density tool. The more runs that occur in close proximity, the higher the run density and the darker the red. This data is classified using the April, 2013 – April, 2014 classification scheme.
Run Density - All CFS
April, 2011 - April, 2012

Run density is calculated using the Kernel Density tool. The more runs that occur in close proximity, the higher the run density and the darker the red. This data is classified using the April, 2013 – April, 2014 classification scheme.
Run Density - All CFS
April, 2012 - April, 2013

Run density is calculated using the Kernel Density tool. The more runs that occur in close proximity, the higher the run density and the darker the red. This data is classified using the April, 2013 – April, 2014 classification scheme.
Run Density - All CFS
April, 2013 - April, 2014

Run density is calculated using the Kernel Density tool. The more runs that occur in close proximity, the higher the run density and the darker the red. This data is classified using the April, 2013 – April, 2014 classification scheme.
Run density is calculated using the Kernel Density tool. The more runs that occur in close proximity, the higher the run density and the darker the red. This data is classified using the April, 2013 – April, 2014 classification scheme.
Run Density - All CFS
April, 2007 - March, 2015

Run density is calculated using the Kernel Density tool. The more runs that occur in close proximity, the higher the run density and the darker the red. This data is classified using the April, 2013 – April, 2014 classification scheme with two extra classes to accommodate the extremely high values.
Spatial Clustering
500 ft Hexagonal Grid

A hexagonal grid is used to aggregate CFS locations. The hexagons are 500 ft in diameter. Hexagonal grids provide better a better aggregation scheme over traditional rectangular grids: a) less distortion due to edge effects, b) even spacing between centroids, c) improved visual quality. A spatial clustering analysis was done on the hex grid to determine if the spatial pattern was random or not. The difference between the actual CFS per hexagon and the expected CFS per hexagon determine the Z-Score. A higher Z-Score indicates a higher than expected local sum and a lower Z-Score indicates a lower than expected local sum.

The further these Z-Scores are away from 0 (both + and -), the lower the probability (P-Values) it is that this spatial pattern occurs due to random chance. Hexagonal bins with a Z-score of $\geq 1.96$ indicate the probability is $< 5\%$ that this spatial pattern is due to random chance. Given the p-values shown to the right, the probability is actually much closer to 0 that this spatial pattern is due to random chance.
Mean Center & Directional Distribution

➤ Background Data

➤ Mean Center

► Middle location in a data set
► Similar to arithmetic mean but with locations not numbers

➤ Directional Distribution

► Ellipse that encompasses roughly 65% – 70% of the data
  > Large ellipse = dispersed data
  > Small ellipse = centralized data
► Direction locations are moving over time
All CFS Codes: 04/2007 - 03/2015
Mean Center & Directional Distribution

Calculating the mean center of a point distribution can help determine the general trend in the distribution of the points over time. The size and shape of the directional distribution ellipse can provide insight as to the general trend in the direction of the calls as well as if the calls are becoming more dispersed or not. The greater the area of the ellipse, the more dispersed the calls are.

One quarter of all calls originate from five locations. In order to achieve a more accurate representation of the calls from the general population, these five locations were removed from the analysis.
Current Vacant Land and Associated Zoning

While impossible to predict exactly what a parcel of land will be used for in the future, an educated guess can be made using currently vacant land and City zoning codes. The parcels outlined in gray and that are in full color represent areas that have not been developed.
Travel Time Estimation

➢ Where? Why? Which Way?

➢ Travel Time
  ➢ En route to the arrival on scene – a.k.a. drive time

➢ Network Analysis
  ➢ Estimates how far a vehicle can travel in a given time based on speed, turns, traffic data, etc.
Interpolation

- Creates a continuous surface from specific observations at unique points (i.e. temperature or rainfall map)
Interpolated Travel Time
Original Dataset

The interpolated travel time is calculated using the Inverse Distance Weighting method with a power of .05 and a variable search radius with 15 points.

The dataset consists of non In-Service and non Mutual Aid CFS classes ranging from April, 2007 – March, 2015. Points have not been modified in any way. The Travel Times used are the actual travel time to every location, not averages. Outliers have not been removed.

Interpolation estimates values between specific observations at unique points (travel times of actual runs) to generate a continuous contour map of travel times. This is same method used when creating temperature or rainfall maps and makes it possible to estimate values where there are none.

**Time Breaks**

- Purple: < 2 Min
- Light Blue: 2 - 3 Min
- Cyan: 3 - 4 Min
- Green: 4 - 5 Min
- Yellow: 5 - 6 Min
- Orange: 6 - 7 Min
- Red: > 7 Min
These maps highlight the variability in the geographic extent of calls that have travel times in the less than 1 and 1 – 2 minute range. Further study is needed to explain the variability. This does not include in service calls.
Dealing with Outliers

- **Integrate Run Locations**
  - Take points that occur within 100 ft
  - Shift them so they are spatially coincident

- **Average Run Times**
  - Helps negate but not eliminate the effect of travel time outliers

### Travel Times Prior to Integration and Averaging vs. Travel Times After Integration and Averaging

<table>
<thead>
<tr>
<th>Travel Time (min)</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.016667 to 10.620833</td>
<td></td>
</tr>
<tr>
<td>10.620833 to 21.225000</td>
<td></td>
</tr>
<tr>
<td>21.225000 to 31.829167</td>
<td></td>
</tr>
<tr>
<td>31.829167 to 42.4333333</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Travel Time (min)</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.033333 to 3.860417</td>
<td></td>
</tr>
<tr>
<td>3.860417 to 7.687500</td>
<td></td>
</tr>
<tr>
<td>7.687500 to 11.514583</td>
<td></td>
</tr>
<tr>
<td>11.514583 to 15.341667</td>
<td></td>
</tr>
</tbody>
</table>
Spatial Outliers After Integration & Averaging
Non In-Service & Mutual Aid CFS
April, 2007 - March, 2015

Integrating and Averaging run times and removing outliers helps create a more representative dataset of normal conditions. The effects of individual events (trains, traffic, construction, weather, red lights, etc.) have been removed from the data to create the basis for an interpolation model that can be compared across time and eventually space.

The red and blue points are spatial outliers and are labeled with the average travel time to that point.

Spatial Outliers
- Non Outlier Points
  - Red: High Travel Time surrounded by low Travel Times
  - Blue: Low Travel Time surrounded by high Travel Times
Interpolated Travel Time Revised Dataset

The interpolated travel time is calculated using the Inverse Distance Weighting method with a power of .05 and a variable search radius with 15 points.

The dataset consists of non In-Service and non Mutual Aid CFS classes ranging from April, 2007 – March, 2015. Outliers were removed from the dataset by integrating the runs within 100ft, averaging the run times of the integrated points, and removing spatial outliers using the Cluster and Outlier tool.

Interpolation estimates values between specific observations at unique points (travel times of actual runs) to generate a continuous contour map of travel times. This is same method used when creating temperature or rainfall maps and makes it possible to estimate values where there are none.

**Time Breaks**

- < 2 Min
- 2 - 3 Min
- 3 - 4 Min
- 4 - 5 Min
- 5 - 6 Min
- 6 - 7 Min
- > 7 Min
Interpolated Travel Time Revised Dataset

The interpolated travel time is calculated using a Weighting method with a power of 0.15 points.

The dataset consists of non InterCity CTS classes ranging from 1 to 4. Outliers were removed from the dataset.

Network Analysis

Interpolation
Current City Station Roadway Distance Bands

Mileage bands are calculated by summing the actual roadway miles from a point of origin as opposed to direct line of sight distance. There is no relation to speed limit or travel time.
Time as Distance Increases

Mileage bands are calculated by measuring the actual distance from the point of origin as opposed to direct line of sight distance. There is no relation to speed limit or travel time.
Time as Distance Increases with Regression Predicted Times

\[ y = 1.32x + 1.36 \]

\[ R^2 = .70 \]
\[ p < .001 \]
\[ n = 833 \]
Current City Station Regression Estimated Travel Times

Mileage bands are calculated by summing the actual roadway miles from a point of origin as opposed to direct line of sight distance. Average distances were calculated for each band (for example a band with a distance range 1 - 1.25 would have an average distance of 1.125) and run locations were spatially joined to the individual distance bands they were contained in. This gave each run location a travel time and a distance from the station. Using Microsoft Excel, linear regression was performed to calculate predicted travel times for each distance band. The distance bands were then symbolized with these predicted values to achieve the time breakdowns.
Mileage bands are calculated by summing the distance from the point of origin as opposed to direct line distances. The distance bands were calculated for each station and run locations were spatially distributed within the distance bands they were assigned. The run location a travel time band was calculated based on the number of runs per day for each station. The Microsoft® Excel® tool was used to calculate the bands.
Hypothetical Station Locations

Hypothetical station locations were chosen based on current or possible future availability of land as well as a perceived need for stations in that area. Ten locations in addition to the current City Station were chosen. Of the 11 total stations, 4 (A, G, H, I) were not included in the following Regression Estimated Travel Time analysis because the travel distance from the hypothetical station to the furthest possible address exceeds that of the current City station.
Percent coverage is calculated for two different sets of demand points. The first is every commercial and residential address point in the City (n=10,596). The second is all previous run locations with a valid travel time (n=12,366). The rankings are determined by the number of demand points covered in a given time frame (4 min, 5 min).
Hypothetical Station J
Regression Estimated Travel Times

Mileage bands are calculated by summing the actual roadway miles from a point of origin as opposed to direct line of sight distance. Average distances were calculated for each band (for example a band with a distance range 1 – 1.25 would have an average distance of 1.125) and run locations were spatially joined to the individual distance bands they were contained in. This gave each run location a travel time and a distance from the station. Using Microsoft Excel, linear regression was performed to calculate predicted travel times for each distance band. The distance bands were then symbolized with these predicted values to achieve the time breakdowns.

Time Breaks

< 2 Min
2 - 3 Min
3 - 4 Min
4 - 5 Min
5 - 6 Min
6 - 7 Min
> 7 Min

Percent coverage is calculated for two different sets of demand points. The first is every commercial and residential address point in the City (n=10,596). The second is all previous run locations with a valid travel time (n=12,366). The rankings are determined by the number of demand points covered in a given time frame (4 min, 5 min).

<table>
<thead>
<tr>
<th>Rankings</th>
<th>All City Address Points Covered by A Single Station</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Station</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>2 2 2.0</td>
<td>J</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rankings</th>
<th>All Previous Runs Covered by A Single Station</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Station</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>2 1 1.5</td>
<td>J</td>
</tr>
</tbody>
</table>
Two Station Rankings

<table>
<thead>
<tr>
<th>Rankings</th>
<th>All Previous Runs Covered by A Two Stations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>April, 2007 - March, 2015 n=12,366</td>
</tr>
<tr>
<td>4 min</td>
<td>5 min</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
</tr>
</tbody>
</table>

Percent coverage is calculated for two different sets of demand points. The first is every commercial and residential address point in the City (n=10,596). The second is all previous run locations with a valid travel time (n=12,366). The rankings are determined by the number of demand points covered in a given time frame (4 min, 5 min).
Current City Station & Hypothetical Station D Regression Estimated Travel Times

Mileage bands are calculated by summing the actual roadway miles from a point of origin as opposed to direct line of sight distance. Average distances were calculated for each band (for example a band with a distance range 1 – 1.25 would have an average distance of 1.125) and run locations were spatially joined to the individual distance bands they were contained in. This gave each run location a travel time and a distance from the station. Using Microsoft Excel, linear regression was performed to calculate predicted travel times for each distance band. The distance bands were then symbolized with these predicted values to achieve the time breakdowns.

Time Breaks

<table>
<thead>
<tr>
<th>Time Range</th>
<th>Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 2 Min</td>
<td>Purple</td>
</tr>
<tr>
<td>2 - 3 Min</td>
<td>Blue</td>
</tr>
<tr>
<td>3 - 4 Min</td>
<td>Green</td>
</tr>
<tr>
<td>4 - 5 Min</td>
<td>Yellow</td>
</tr>
<tr>
<td>5 - 6 Min</td>
<td>Orange</td>
</tr>
<tr>
<td>6 - 7 Min</td>
<td>Red</td>
</tr>
<tr>
<td>&gt; 7 Min</td>
<td>Brown</td>
</tr>
</tbody>
</table>

Percent coverage is calculated for two different sets of demand points. The first is every commercial and residential address point in the City (n=10,596). The second is all previous run locations with a valid travel time (n=12,366). The rankings are determined by the number of demand points covered in a given time frame (4 min, 5 min).

<table>
<thead>
<tr>
<th>Rankings</th>
<th>All City Address Points Covered by Two Stations n=10,596</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rankings</th>
<th>All Previous Runs Covered by Two Stations April, 2007 – March, 2015 n=12,366</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Engine 38 GPS Data
6/26/2015 - 8/26/2015

- Average Speeds
  - Specific Road and Road Type
  - Time of Day
  - Weather
  - Traffic
- Intersection/Turn Delays
- Rail Crossings
- Optimize Routing
Spatial Clustering
Directional Trending
Network Analysis
Spatial Outliers
Station Optimization
Thank You

Data Driven

GIS Centric

Evidence Based