



Using GIS to Predict High Crash Segments in Highway

Lixin Huang, PhD, GISP

GIS Analyst III

Brevard County, Board of County Commissioners

2013 Esri International User Conference

July 10, 2013

Why Predicting High Crash Segments

- Help transportation agencies find the crash patterns and develop solutions for roadway segments with potential higher number of crashes
- Help emergency agencies allocate limited resources to the areas mostly needed
- Save lives

How Predicting High Crash Segments

- Collect crash data
- Group crash data by time of day
- Identify crash clusters in each group
- Interpret Results

Collect Crash Data

- Crash data format
 - Excel Spreadsheet: CSV, XLSX
 - Non-spatial database: DBF
 - Spatial database: Feature class
- Convert Excel Spreadsheet and non-spatial database to spatial database
 - CSV to Feature class
- Organize crash data
 - Yearly
 - Consolidated

Group Crash Data by Time of Day

Criteria of grouping crash data are based on the traffic pattern which varies during a week and a day.

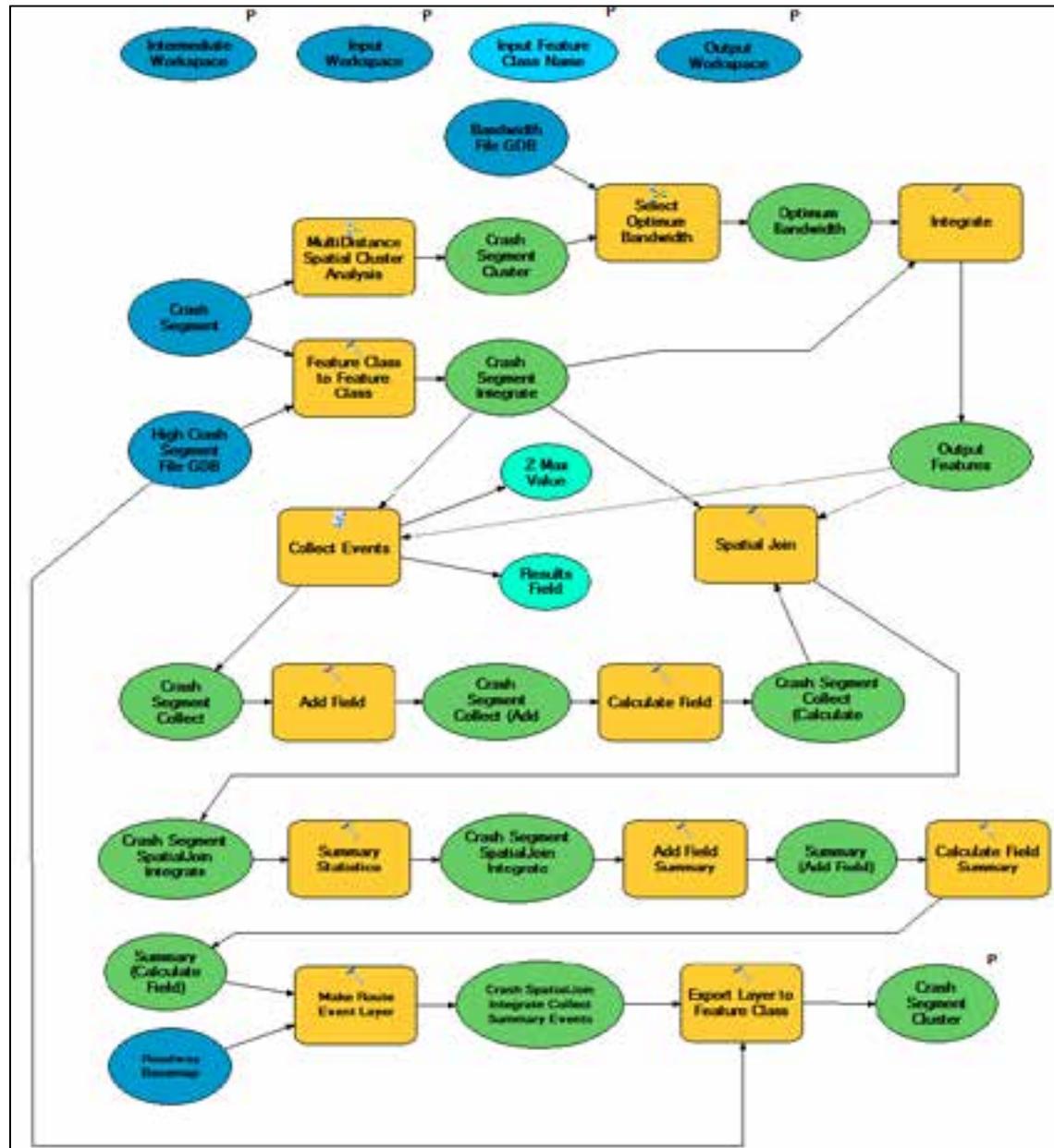
- Week

- Weekday: Monday ~ Friday
- Weekend: Saturday and Sunday

- Day

- One-hour interval: 12:00 am ~ 1:00 am, 1:00 ~ 2:00 am
- Three-hour interval: 12:00 am ~ 3:00 am, 3:00am ~ 6:00 am
- Six-hour interval: 12:00 am ~ 6:00 am, 6:00 am ~ 12:00 pm

Identify Crash Clusters in Each Group

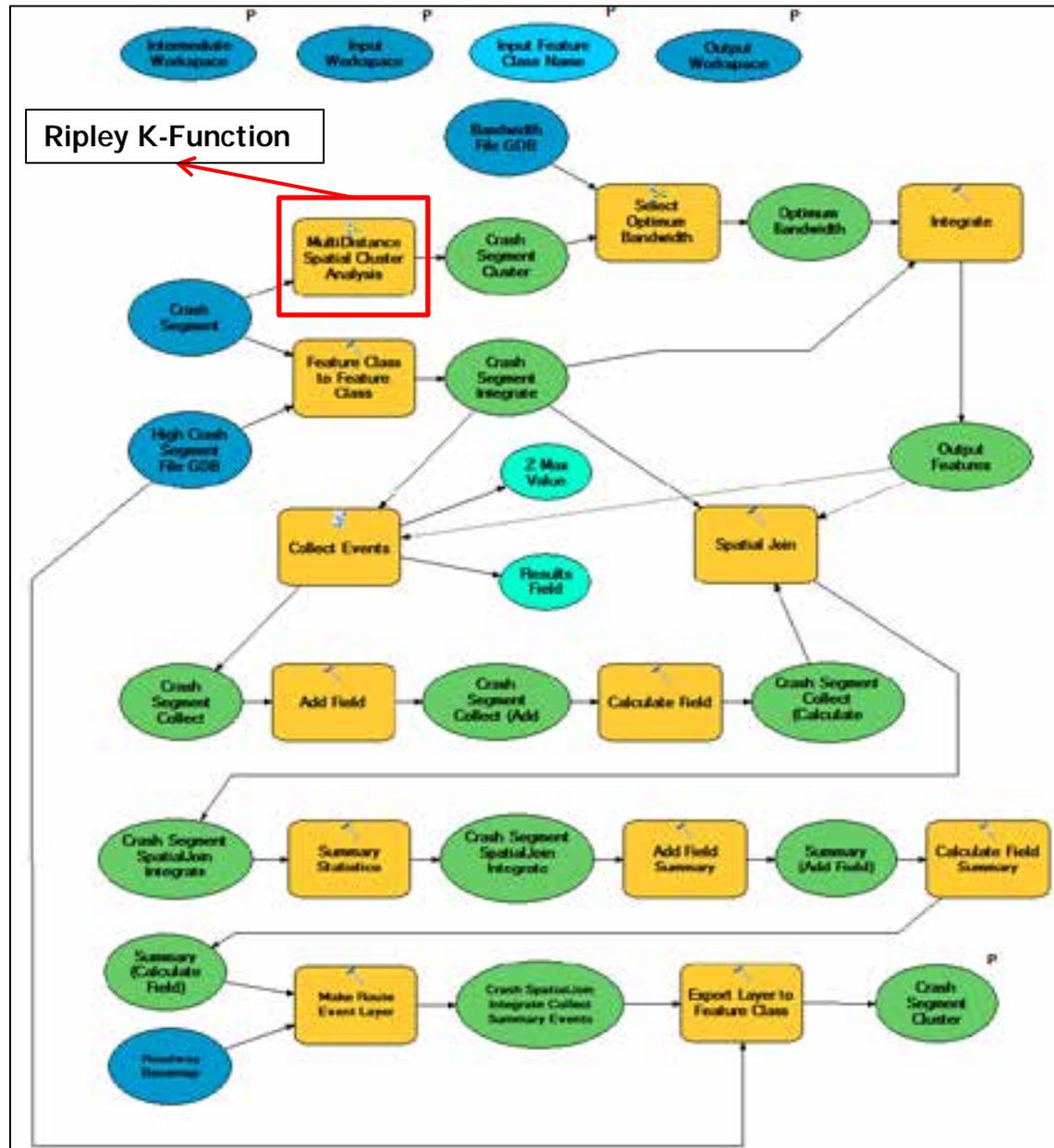


Identify Crash Clusters in Each Group (Cont'd)

Main Model Components

- Ripley's K-Function: Find the clustered crash segments
- Bandwidth Selection: Select the optimum bandwidth at which the crash is mostly clustered

Ripley's K-Function



Ripley's K-Function (Cont'd)

Equation

$$L(d) = \sqrt{\frac{A \sum_{i=1}^n \sum_{j=1, j \neq i}^n k(i, j)}{\rho n (n - 1)}}$$

$L(d)$: Common transformation of K-Function;

A : Total area of the features;

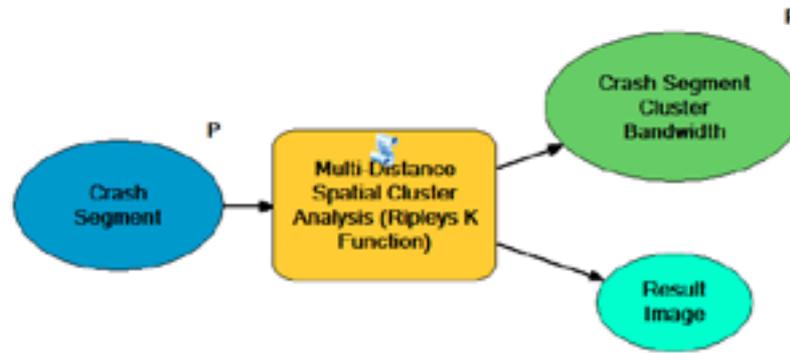
d : Distance;

n : Total number of features

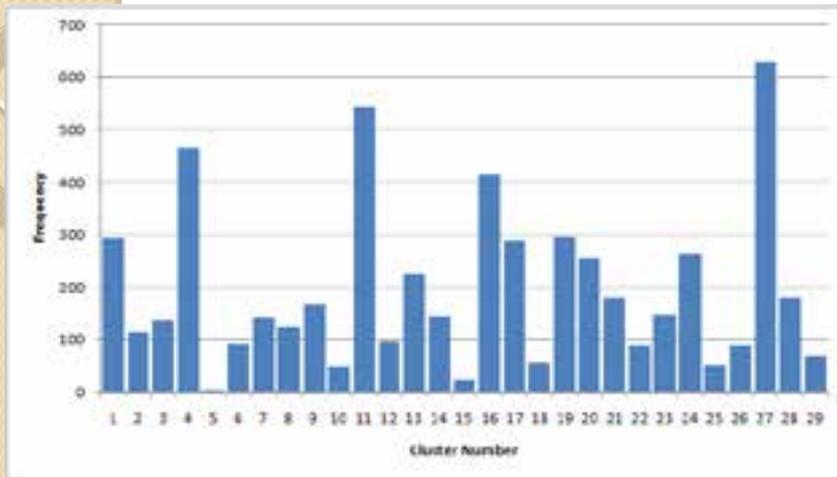
$k(i, j)$: Weight.

Ripley's K-Function (Cont'd)

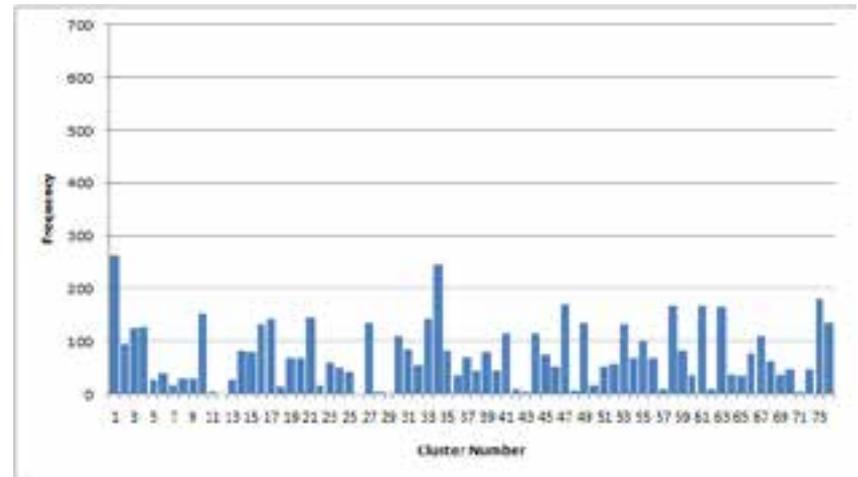
Model Diagram



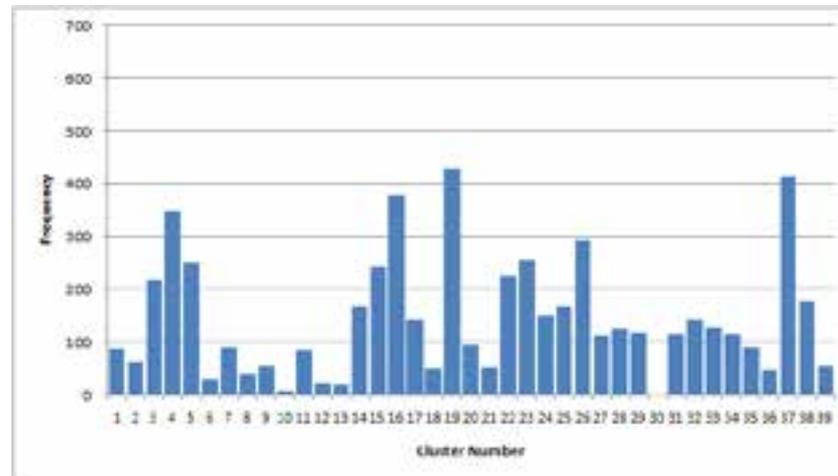
Bandwidth Selection (Cont'd)



Larger Bandwidth (Less Segments)



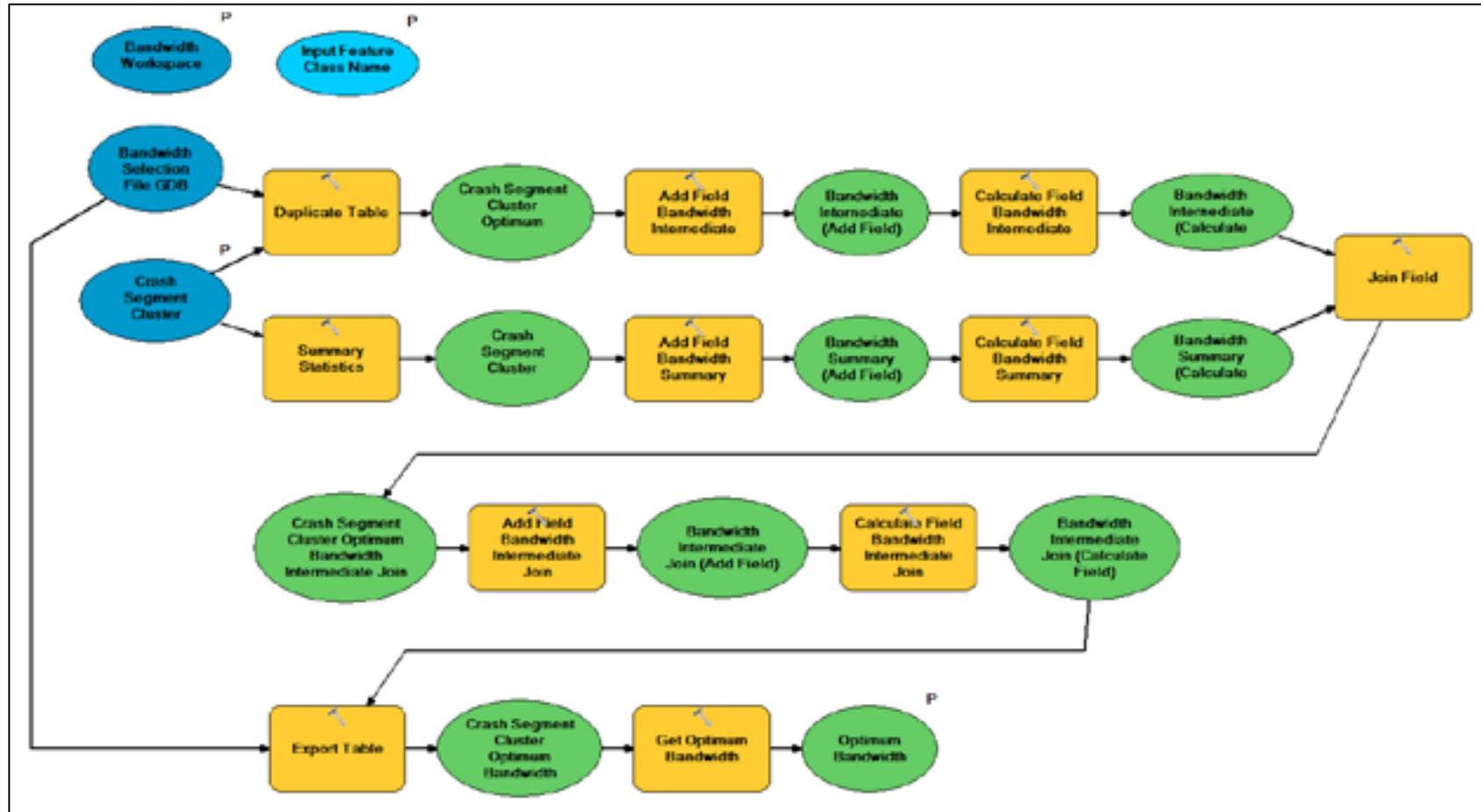
Smaller Bandwidth (More Segment)



Optimum Bandwidth

Bandwidth Selection (Cont'd)

Model Diagram



Interpret Results

- Input
 - Six Years Crash Data
 - Three-hour interval on weekday and weekend
- Output
 - Crash Clusters
 - Crash Frequency
 - Cluster Begin and End Locations
 - Cluster Length
 - Time
 - GIS Web Application with Time Enabled

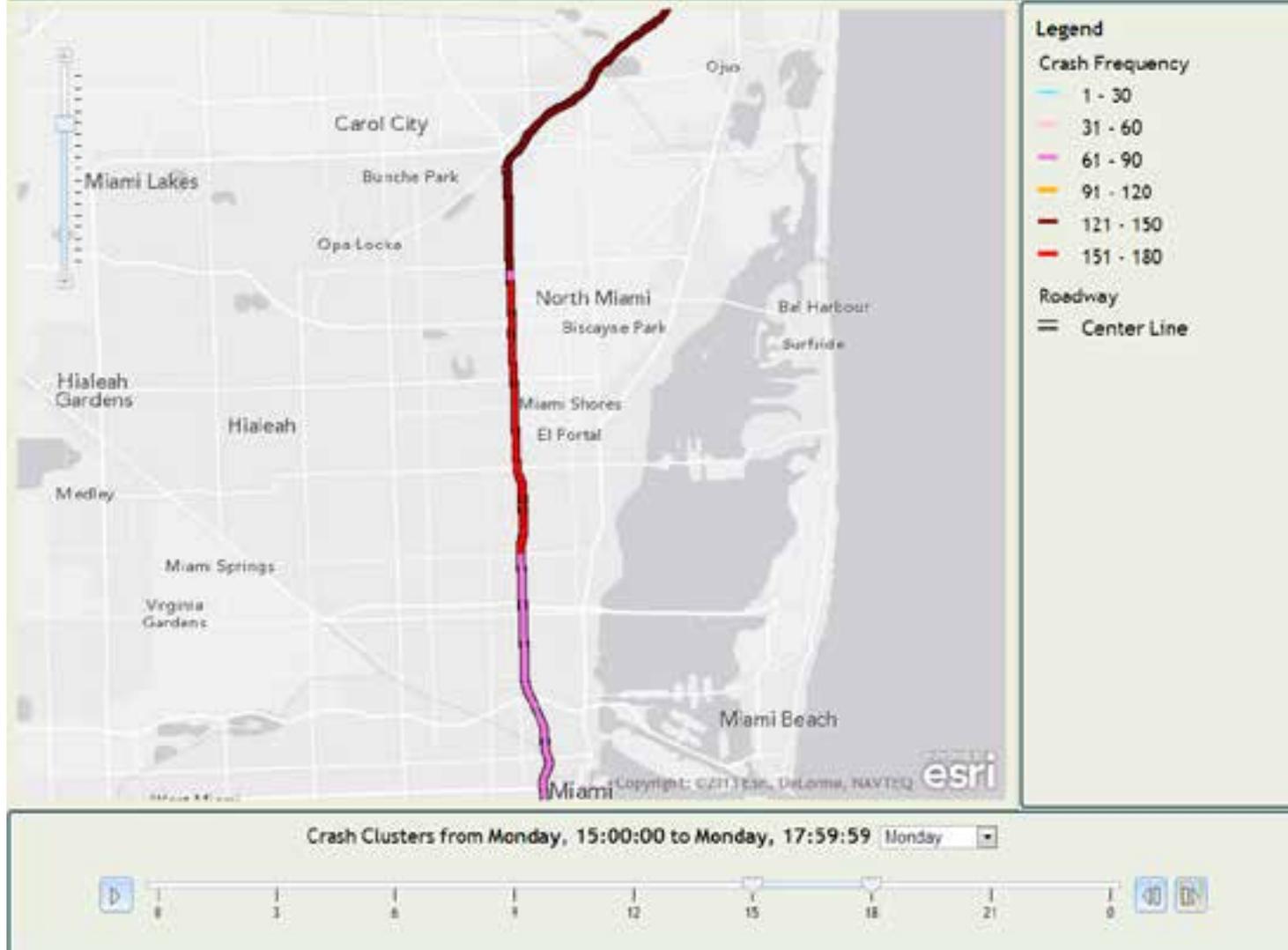
Results - Crash Clusters

FREQUENCY	MIN_LOCMILEPT	MAX_LOCMILEPT	SEG_LEN	TIME_START	TIME_END
5	12.483	13.356	0.873	2010/12/20 00:00:00	2010/12/20 02:59:59
10	9.286	10.158	0.872	2010/12/20 00:00:00	2010/12/20 02:59:59
7	0.575	0.994	0.419	2010/12/20 00:00:00	2010/12/20 02:59:59
9	6.708	7.689	0.981	2010/12/20 00:00:00	2010/12/20 02:59:59
3	14.714	15.56	0.846	2010/12/20 00:00:00	2010/12/20 02:59:59
11	4.23	5.104	0.874	2010/12/20 00:00:00	2010/12/20 02:59:59
8	5.732	6.217	0.485	2010/12/20 00:00:00	2010/12/20 02:59:59
6	16.598	17.01	0.414	2010/12/20 00:00:00	2010/12/20 02:59:59
4	2.923	3.681	0.758	2010/12/20 00:00:00	2010/12/20 02:59:59
14	10.768	12.095	1.329	2010/12/20 00:00:00	2010/12/20 02:59:59
10	8.005	8.84	0.835	2010/12/20 00:00:00	2010/12/20 02:59:59
5	0.707	3.515	2.808	2010/12/20 03:00:00	2010/12/20 05:59:59
39	10.29	17.01	6.72	2010/12/20 03:00:00	2010/12/20 05:59:59
42	5.139	9.758	4.619	2010/12/20 03:00:00	2010/12/20 05:59:59
32	10.445	11.223	0.778	2010/12/20 08:00:00	2010/12/20 08:59:59
2	0.994	1.418	0.424	2010/12/20 08:00:00	2010/12/20 08:59:59
26	5.642	6.417	0.775	2010/12/20 08:00:00	2010/12/20 08:59:59
19	11.559	12.345	0.786	2010/12/20 08:00:00	2010/12/20 08:59:59
19	2.836	3.922	1.086	2010/12/20 08:00:00	2010/12/20 08:59:59
2	0.519	0.546	0.027	2010/12/20 08:00:00	2010/12/20 08:59:59
26	9.342	10.319	0.977	2010/12/20 08:00:00	2010/12/20 08:59:59
25	14.156	15.464	1.308	2010/12/20 08:00:00	2010/12/20 08:59:59
9	13.356	13.892	0.536	2010/12/20 08:00:00	2010/12/20 08:59:59
37	8.071	9.062	0.991	2010/12/20 08:00:00	2010/12/20 08:59:59
24	7	7.771	0.771	2010/12/20 08:00:00	2010/12/20 08:59:59
33	4.318	5.354	1.036	2010/12/20 08:00:00	2010/12/20 08:59:59
11	12.672	13.152	0.48	2010/12/20 08:00:00	2010/12/20 08:59:59
17	16.06	17.004	0.944	2010/12/20 08:00:00	2010/12/20 08:59:59
17	11.559	12.345	0.786	2010/12/20 09:00:00	2010/12/20 11:59:59
10	12.539	12.994	0.455	2010/12/20 09:00:00	2010/12/20 11:59:59

GIS Web Application

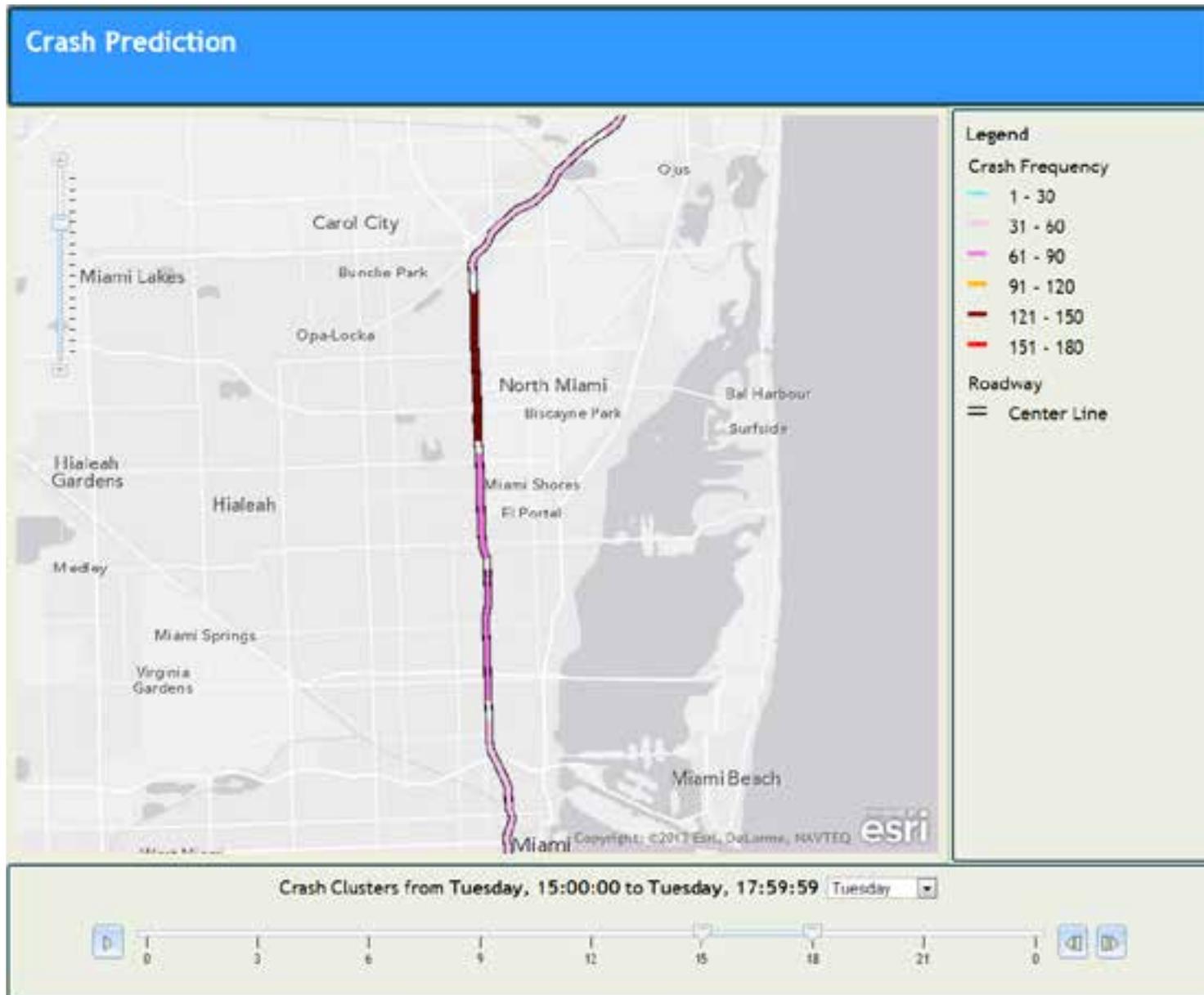
Crash Cluster @ 15:00 ~ 18:00, Monday

Crash Prediction



GIS Web Application

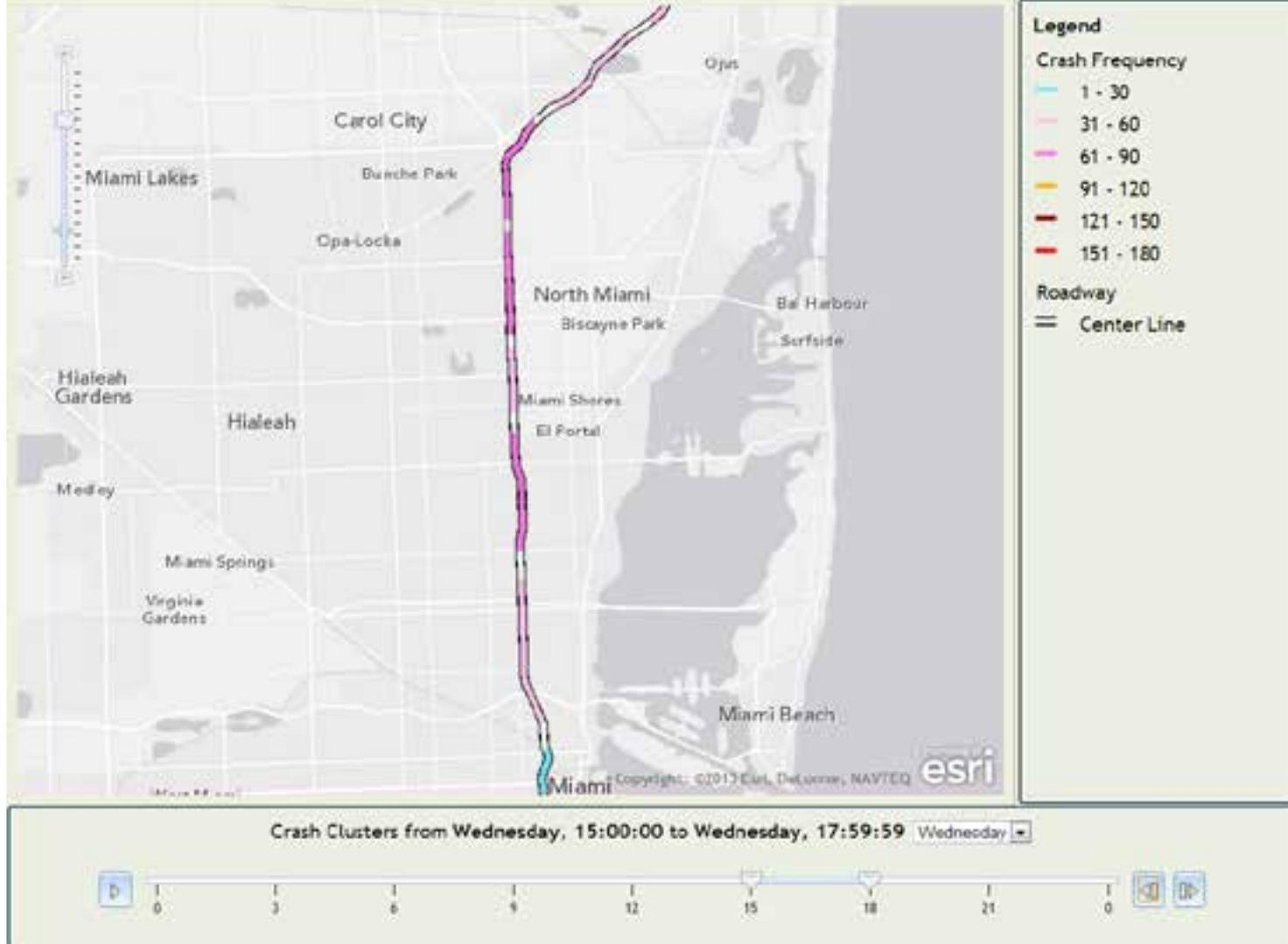
Crash Cluster @ 15:00 ~ 18:00, Tuesday



GIS Web Application

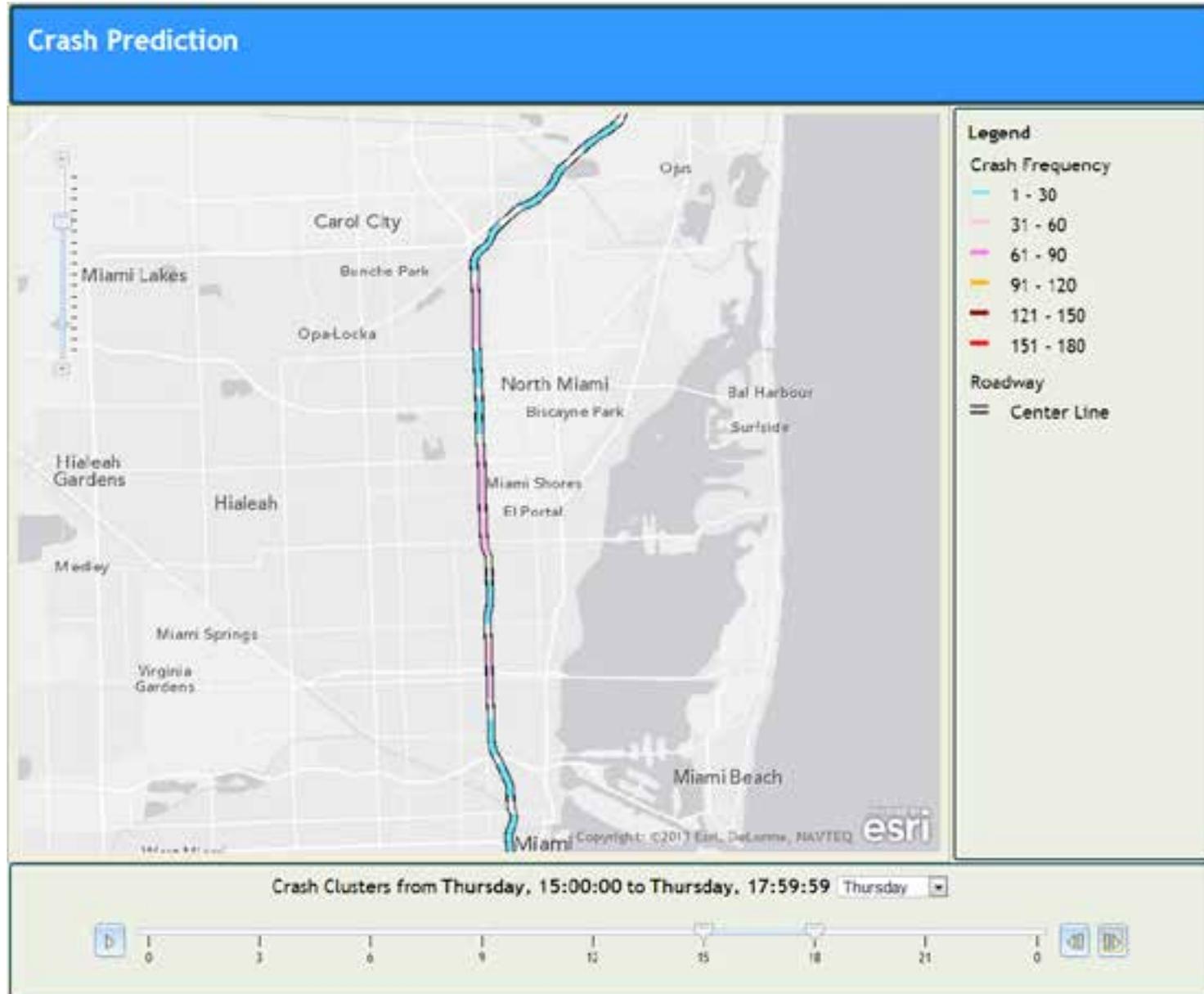
Crash Cluster @ 15:00 ~ 18:00, Wednesday

Crash Prediction



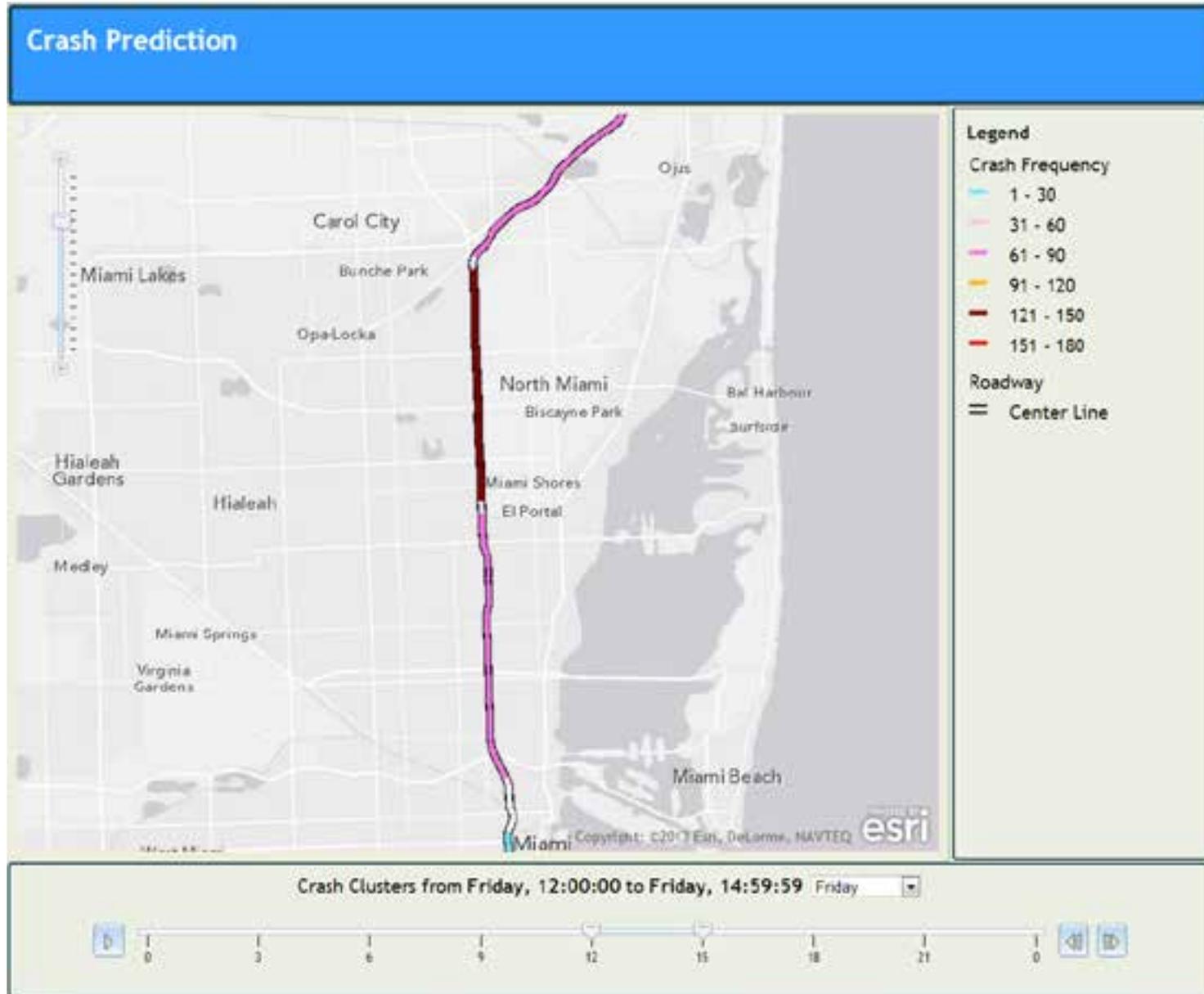
GIS Web Application

Crash Cluster @ 15:00 ~ 18:00, Thursday



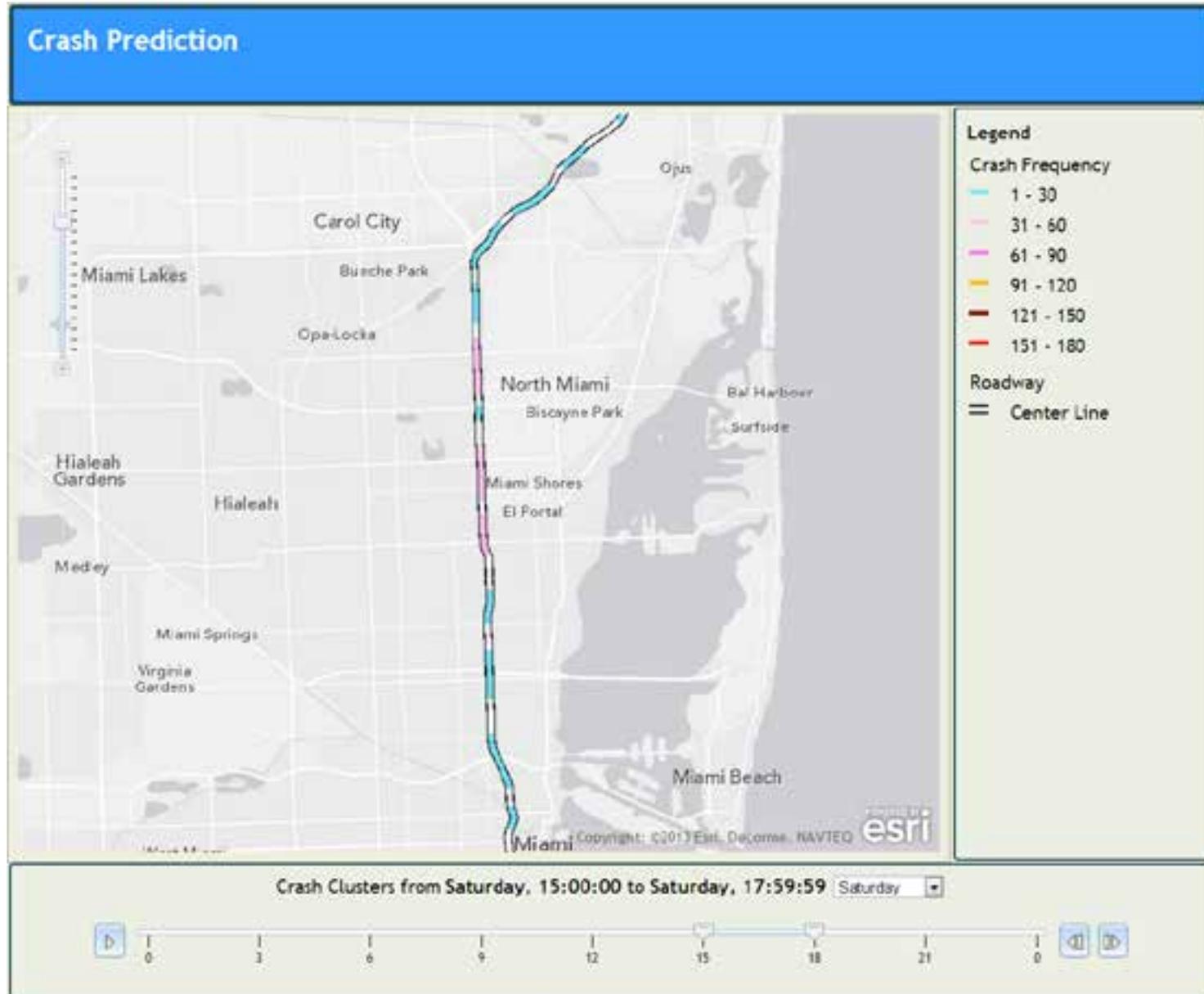
GIS Web Application

Crash Cluster @ 15:00 ~ 18:00, Friday



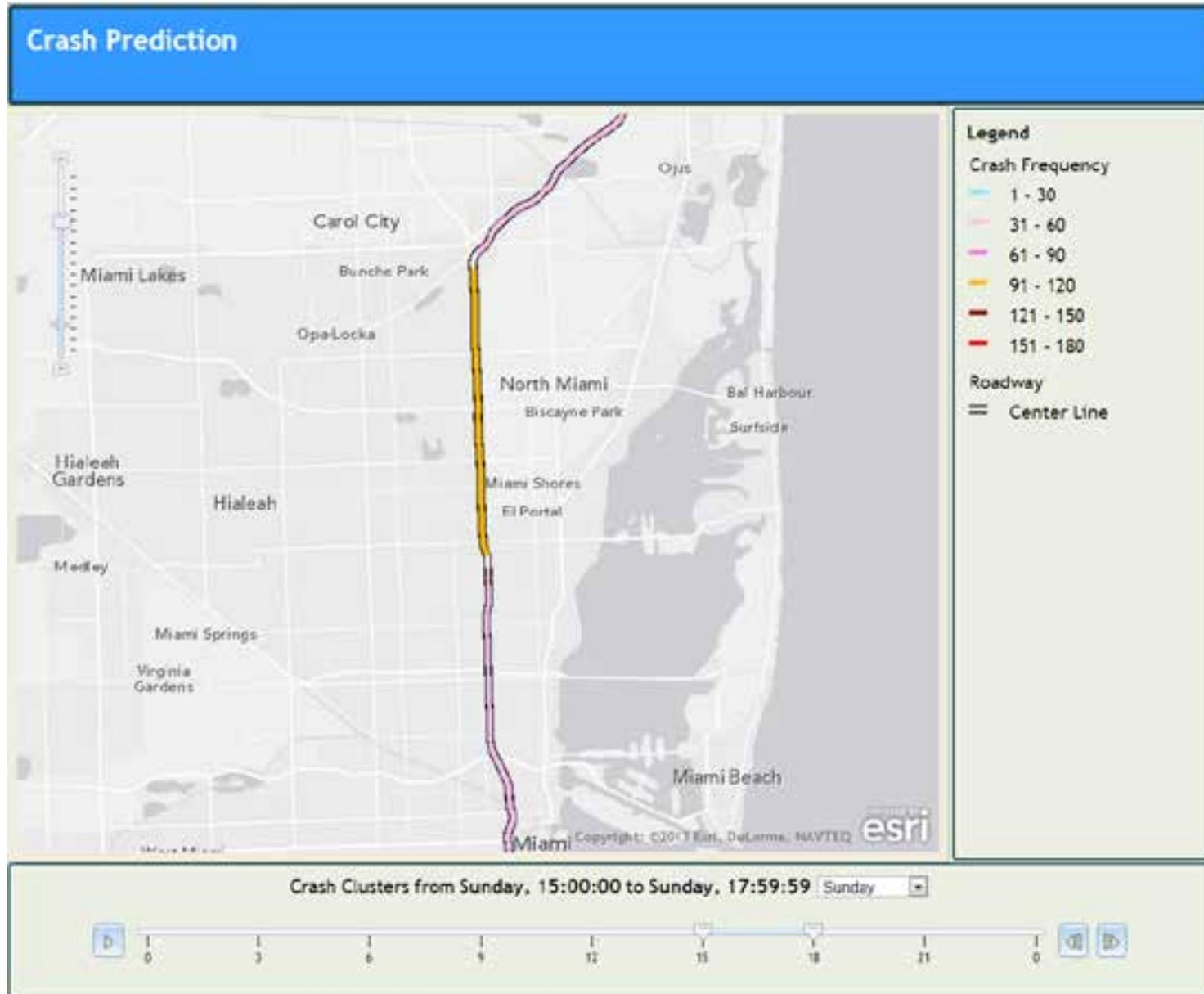
GIS Web Application

Crash Cluster @ 15:00 ~ 18:00, Saturday



GIS Web Application

Crash Cluster @ 15:00 ~ 18:00, Sunday



Summary

- Know the locations where crashes could most probably happen.
- Allocate resources, e.g., fire rescue, efficiently.

Questions

