Technical Writing

Unit Availability vs. Response Time (An explanation of how cities calculate their future needs for fire stations, intended for nonspecialist officials.)

Municipalities need to know whether their emergency-response system is meeting their current needs, and whether the system will continue to meet needs in the future. As urban areas experience growth and development, emergency calls increase and there is a gradual decrease in the ability of fire stations to provide a high level of emergency response. If no corrective action is taken, at some point response times will fall below acceptable standards. For the purpose of planning it is important to be able to predict the point at which this will occur—the point at which the number of calls will exceed a station's capacity—and to take corrective action ahead of time. The starting point of this assessment is to have a clear picture of an area's current performance.

The Center for Public Safety Excellence (CPSE) has devised a means of measuring a fire department's current performance and predicting a drop in the department's ability to respond to calls—in other words, to measure and predict the capacity of stations and areas. This method compares two factors:

1. *Unit availability:* How often the closest station (the intended closest unit) is available to handle a call

2. *Performance:* How often the travel time is within the desired response time

By comparing those two factors we can see what happens to response time when the intended closest unit is not available to respond to a call: other units have to respond from farther away, response time increases, and at some point it falls below the established performance standard. For the example below we will use a standard of arriving within four minutes 80% of the time.

Figure 1 shows how unit availability and response time can be plotted graphically to study their relationship. The horizontal **Unit Availability axis** shows how often the intended closest unit was available for service out of a period of 24 hours. If a station is available for 18 hours out of 24, it would have an availability of 75%. Various factors reduce any unit's availability: workload, inspections, training, maintaining equipment, and so on. Those factors reduce a station's performance.

The vertical **Performance Reliability axis** shows how often the unit arrives within the established time goal. The dotted line across the middle of the graph represents a performance standard of 80%, meaning in our example that a unit arrives within four minutes 80% of the time.

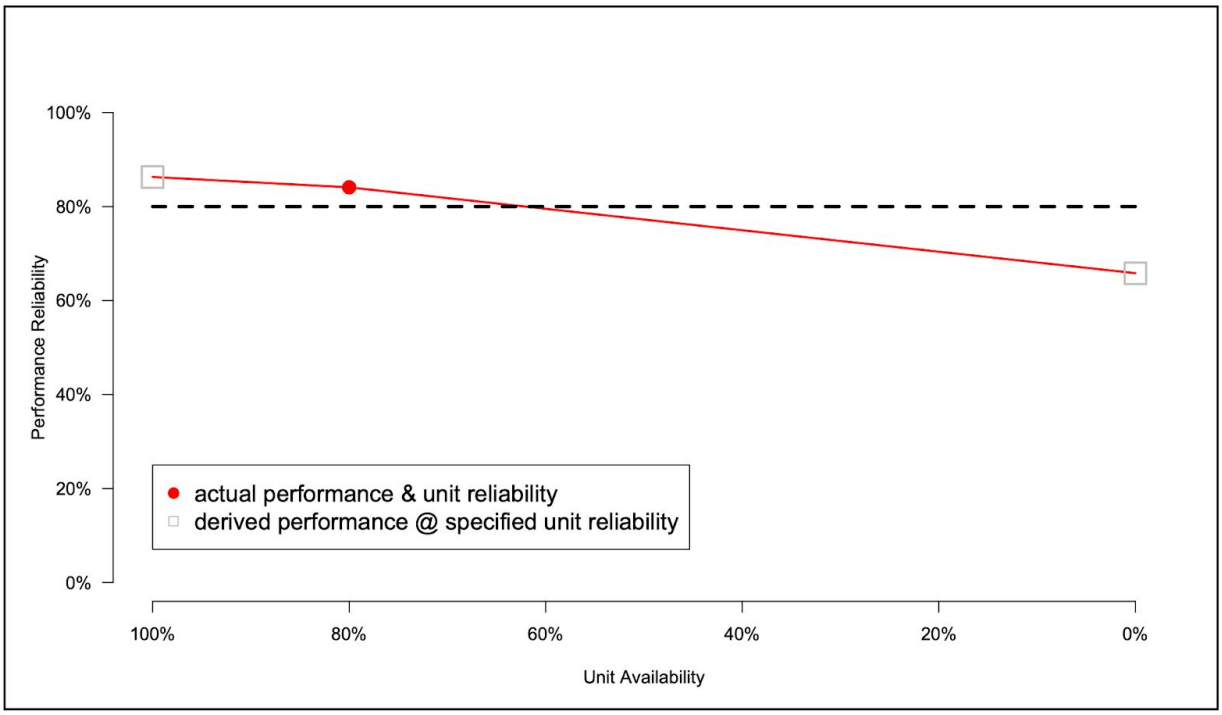
The graph below describes the performance of Station 8, the first-due station in the area. The square on the left side of the graph represents a theoretical condition of 100% availability for Station 8. It shows that if Station 8 were available 100% of the time, 24 hours a day, and able to respond to every call, it would achieve a compliance level of about 85%, meaning it would arrive within four minutes about 85% of the time. This is above (better than) the area's targeted performance standard.

But in fact Station 8 is only available about 80% of the time. The red dot represents the station's actual performance, which is a bit lower than it would be if it were available 100% of the time, but still above the performance standard. If the red dot were below the dotted line that would indicate that the station was not currently meeting the performance standard.

What would happen if Station 8 were never available, meaning always busy with other calls and unable to respond to a new call? In that case other units from outside the area would need to respond to calls, and the result would be lower performance (increased response time). The square on the right side of the graph shows a theoretical zero availability for Station 8 and a performance of only about 67% for units outside the area that take up the slack. Units are arriving within four minutes only 67% of the time, which is below the performance goal of 80% because the emergency-response system is now stretched beyond its capacity.

The red line connecting these three data points shows how performance changes as unit availability declines. We can see that when the unit availability declines to 60%, the performance standard is barely met. Below 60% availability, the standard is no longer being met. Corrective action needs to be taken *before* the 60% availability level is reached, so it's important to know when this is going to occur.

Figure 1: Example of Response Time Performance



The slope of the line reflects *workload sensitivity*, or how an area is impacted by an increase or decrease in calls. Areas that have multiple units available or several nearby fire stations that can cover calls are less impacted by an increase in calls. They have a low workload sensitivity. This would be shown on the graph by a nearly horizontal line—as availability declines, response time remains about the same.

In an area that has fewer units available and other stations are located farther away, an increase in calls will exceed the area's capacity to respond and response times by units outside the area will be excessive. This area is workload sensitive—heavily dependent on unit availability. This would appear on the graph as a line that slopes strongly downward.

If the square on the left side of the graph—representing 100% availability—falls below the 80% performance goal, this indicates a problem with fire station location. The Performance

Reliability station is not meeting its time-response goal even though it is always available, because it is too far away from calls. The unit is arriving late 20% of the time or more, indicating that it is not well located to reach all parts of its first-due area. This problem can be corrected by relocating the current station, building a new station, or—perhaps the best solution—reevaluating first-due boundaries to make sure that the most appropriate station is responding.

This system of graphing can be used to plan for future station locations by monitoring response trends over time and noting when response time, due to declining unit availability, is heading toward a point when action must be taken to avoid falling below the performance goal. If a station's current capacity is known, the graph can be used to calculate remaining capacity that is available while continuing to meet the performance goal.