Report No. K-TRAN: KU-17-2 ▪ FINAL REPORT ▪ January 2019

Improving the Accuracy and Applicability of Kansas Traffic Data

Alexandra Kondyli, Ph.D.
Steven D. Schrock, Ph.D., P.E., F.ITE
Georgios Chrysikopoulos

The University of Kansas

Kansas Department of Transportation

1 Report No. K-TRAN: KU-17-2	**2 Government Accession No.**	**3 Recipient Catalog No.**
4 Title and Subtitle Improving the Accuracy and Applicability of Kansas Traffic Data		**5 Report Date** January 2019 **6 Performing Organization Code**
7 Author(s) Alexandra Kondyli, Ph.D., Steven D. Schrock, Ph.D., P.E., F.ITE, Georgios Chrysikopoulos		**8 Performing Organization Report No.**
9 Performing Organization Name and Address The University of Kansas Department of Civil, Environmental & Architectural Engineering 1530 West 15th St Lawrence, Kansas 66045-7609		**10 Work Unit No. (TRAIS)** **11 Contract or Grant No.** C2099
12 Sponsoring Agency Name and Address Kansas Department of Transportation Bureau of Research 2300 SW Van Buren Topeka, Kansas 66611-1195		**13 Type of Report and Period Covered** Final Report May 2016–August 2017 **14 Sponsoring Agency Code** RE-0715-01
15 Supplementary Notes For more information write to address in block 9. Data quality example spreadsheet is available separately upon request to KDOT#Research.Library@ks.gov.		

16 Abstract

This research investigated the accuracy and reliability of KC Scout traffic sensors in the Kansas City metropolitan area. The locations and sensors under investigation are of particular importance to the Kansas Department of Transportation (KDOT) because they regularly report accurate traffic counts to the Federal Highway Administration (FHWA). The research team used dual-directional video recorders at six locations along I-35, I-70, I-435, and I-635 to collect field counts during two phases for a total of 9 hours. Comparison of the manual field counts and KC Scout records showed acceptable accuracy (less than 5% difference compared to field counts) if KC Scout verified that the subject sensors were recently checked and calibrated. Reliability assessment of the subject sensors revealed excellent year-long quality of data for two pairs of detectors (about 98% good data), while the other four pairs provided approximately 85%, 79%, 75%, and 68% good-quality data, respectively. Sources of error that resulted in bad-quality data were duplicate records, zero traffic counts when vehicle speed was non-zero, extremely high traffic counts, or missing intervals. Investigating the temporal distribution of each error type provided insight on their probable causes. Finally, an assessment of the automated quality checks process of the KC Scout data portal identified errors and inconsistencies, which hinder the use of aggregate data by KDOT. Strategies that can mitigate the above issues were suggested.

17 Key Words Traffic Volumes, Sensor Accuracy, Count Reliability		**18 Distribution Statement** No restrictions. This document is available to the public through the National Technical Information Service www.ntis.gov.	
19 Security Classification (of this report) Unclassified	**20 Security Classification (of this page)** Unclassified	**21 No. of pages** 58	**22 Price**

Form DOT F 1700.7 (8-72)

This page intentionally left blank.

Improving the Accuracy and Applicability of Kansas Traffic Data

Final Report

Prepared by

Alexandra Kondyli, Ph.D.
Steven D. Schrock, Ph.D., P.E., F.ITE
Georgios Chrysikopoulos

The University of Kansas

A Report on Research Sponsored by

THE KANSAS DEPARTMENT OF TRANSPORTATION
TOPEKA, KANSAS

and

THE UNIVERSITY OF KANSAS
LAWRENCE, KANSAS

January 2019

PREFACE

The Kansas Department of Transportation's (KDOT) Kansas Transportation Research and New-Developments (K-TRAN) Research Program funded this research project. It is an ongoing, cooperative and comprehensive research program addressing transportation needs of the state of Kansas utilizing academic and research resources from KDOT, Kansas State University and the University of Kansas. Transportation professionals in KDOT and the universities jointly develop the projects included in the research program.

NOTICE

The authors and the state of Kansas do not endorse products or manufacturers. Trade and manufacturers names appear herein solely because they are considered essential to the object of this report.

This information is available in alternative accessible formats. To obtain an alternative format, contact the Office of Public Affairs, Kansas Department of Transportation, 700 SW Harrison, 2nd Floor – West Wing, Topeka, Kansas 66603-3745 or phone (785) 296-3585 (Voice) (TDD).

DISCLAIMER

The contents of this report reflect the views of the authors who are responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the views or the policies of the state of Kansas. This report does not constitute a standard, specification or regulation.

Abstract

This research investigated the accuracy and reliability of KC Scout traffic sensors in the Kansas City metropolitan area. The locations and sensors under investigation are of particular importance to the Kansas Department of Transportation (KDOT) because they regularly report accurate traffic counts to the Federal Highway Administration (FHWA). The research team used dual-directional video recorders at six locations along I-35, I-70, I-435, and I-635 to collect field counts during two phases for a total of 9 hours. Comparison of the manual field counts and KC Scout records showed acceptable accuracy (less than 5% difference compared to field counts) if KC Scout verified that the subject sensors were recently checked and calibrated. Reliability assessment of the subject sensors revealed excellent year-long quality of data for two pairs of detectors (about 98% good data), while the other four pairs provided approximately 85%, 79%, 75%, and 68% good-quality data, respectively. Sources of error that resulted in bad-quality data were duplicate records, zero traffic counts when vehicle speed was non-zero, extremely high traffic counts, or missing intervals. Investigating the temporal distribution of each error type provided insight on their probable causes. Finally, an assessment of the automated quality checks process of the KC Scout data portal identified errors and inconsistencies, which hinder the use of aggregate data by KDOT. Strategies that can mitigate the above issues were suggested.

Acknowledgments

The authors of this report thank project monitors Alan Spicer, Steven Buckley, and William Hughes for their guidance and suggestions. The authors also thank KC Scout personnel Karsten Sommerhauser, Gary Covey, Lindsay Harris, and Nancy Powell for assisting with the sensors and the traffic count data.

Table of Contents

Abstract ... v

Acknowledgments ... vi

Table of Contents ... vii

List of Tables ... ix

List of Figures ... x

Chapter 1: Introduction ... 1

1.1 Background ... 1

1.2 Objectives ... 2

Chapter 2: Review of Current Practice ... 3

2.1 KC Scout Data Aggregation and Quality Checks ... 3

2.2 Accuracy of Wavetronix Radar Sensors ... 5

2.3 FHWA Highway Performance Monitoring System (HPMS) Data Requirements ... 7

2.4 FHWA Traffic Volume Trends (TVT) Data Requirements ... 8

Chapter 3: Data Collection ... 9

3.1 Location Selection ... 9

3.2 Ground Truth Data ... 9

3.3 Sensor Data ... 14

3.4 System Reliability Data ... 14

Chapter 4: Analysis of Data Accuracy ... 15

4.1 Comparison Methodology ... 15

4.2 Summarized Results ... 16

4.2.1 First Data Collection Phase ... 16

4.2.2 Second Data Collection Phase ... 16

4.3 Detailed Results ... 17

4.3.1 Eastbound I-70 at 72nd Street ... 17

4.3.2 Westbound I-70 at 72nd Street ... 18

4.3.3 Northbound I-35 at Cambridge Circle ... 18

4.3.4 Southbound I-35 at Cambridge Circle ... 19

4.3.5 Southbound I-35 at Prairie Street ... 20

4.3.6 Northbound I-35 at Prairie Street 20
4.3.7 Northbound I-35 at 95th Street 21
4.3.8 Southbound I-35 at 95th Street 22
4.3.9 Southbound I-435, North of Kansas Avenue 22
4.3.10 Northbound I-435, North of Kansas Avenue 23
4.3.11 Northbound I-635 at Shawnee Drive 24
4.3.12 Southbound I-635 at Shawnee Drive 24
Chapter 5: Assessment of System Reliability 26
5.1 One-Year Reliability Assessment 26
5.2 Reliability Assessment During Field Data Collection 31
Chapter 6: Conclusions and Recommendations 33
References 37
Appendix A: Using the KC Scout Data Portal 39
A.1 User Authentication and Navigation Interface 39
A.2 Detector Report Queries 40
A.3 Data Quality Assessment 43

List of Tables

Table 2.1: KC Scout Data Quality Checks and Actions 4
Table 2.2: Traffic Data Items Reported for HPMS 8
Table 3.1: Study Sites 9
Table 3.2: Data Collection Periods 10
Table 4.1: Comparison of Field Data and KC Scout Data for Phase I 16
Table 4.2: Comparison of Field Data and KC Scout Data for Phase II 17
Table 4.3: Comparison of Field Data and KC Scout Data at I-70 E at 72nd Street 17
Table 4.4: Comparison of Field Data and KC Scout Data at I-70 W at 72nd Street 18
Table 4.5: Comparison of Field Data and KC Scout Data at I-35 N at Cambridge Circle 19
Table 4.6: Comparison of Field Data and KC Scout Data at I-35 S at Cambridge Circle 19
Table 4.7: Comparison of Field Data and KC Scout Data at I-35 S at Prairie Street 20
Table 4.8: Comparison of Field Data and KC Scout Data at I-35 N at Prairie Street 21
Table 4.9: Comparison of Field Data and KC Scout Data at I-35 N at 95th Street 21
Table 4.10: Comparison of Field Data and KC Scout Data at I-35 S at 95th Street 22
Table 4.11: Comparison of Field Data and KC Scout Data at I-435 S, North of Kansas Avenue 23
Table 4.12: Comparison of Field Data and KC Scout Data at I-435 N, North of Kansas Avenue 23
Table 4.13: Comparison of Field Data and KC Scout Data at I-635 N at Shawnee Drive 24
Table 4.14: Comparison of Field Data and KC Scout Data at I-635 S at Shawnee Drive 25
Table 5.1: One-Year Data Quality Analysis (November 2016 to October 2017) 27
Table 5.2: Data Quality Analysis Throughout the Year 28
Table 5.3: KC Scout Data Portal Quality Report Throughout the Year 29
Table 5.4: Field Data Quality Analysis 32
Table 5.5: KC Scout Field Data Quality Report 32

List of Figures

Figure 2.1: KDOT October 2017 Fact and Trends Report, Kansas City Area Permanent Count Stations 8
Figure 3.1: I-70 West of North 72nd Street, Detector and Recording Locations 11
Figure 3.2: I-435 North of Kansas Avenue, Detector and Recording Locations 11
Figure 3.3: I-635 North of Shawnee Drive, Detector and Recording Locations 12
Figure 3.4: I-35 West of Cambridge Circle, Detector and Recording Locations 12
Figure 3.5: I-35 at West 95th Street, Detector and Recording Locations 13
Figure 3.6: I-35 at Prairie Street, Detector and Recording Locations 13
Figure A.1: Data Portal Home Screen 39
Figure A.2: Edit Detector Report Screen 40
Figure A.3: Report File Example 43
Figure A.4: *Quality Checks* Tab in the Data Quality Example Spreadsheet, Including Summary Statistics Lines 44
Figure A.5: *Summary* Tab in the Data Quality Example Spreadsheet 45

Chapter 1: Introduction

1.1 Background

Highway agencies must collect accurate, reliable traffic data to efficiently operate roadway networks and plan for future growth. These agencies must also submit annual traffic reports to the Federal Highway Administration (FHWA, 2016a) for performance measurements and resource allocation as part of the Highway Performance Monitoring System (HPMS). In addition, monthly traffic counts, in the form of hourly traffic by lane, must be submitted for the monthly Traffic Volume Trends (TVT) report (FHWA, 2016b). The Kansas Department of Transportation (KDOT) collects traffic data at selected permanent count locations across the state and at four locations in the Kansas City metropolitan area.

According to FHWA (2016b), permanent count location selection should consider:

- Quality of traffic data: sensor downtimes resulting in missing data and data imputations may negatively impact the quality of collected data;
- Quality of continuous count equipment: outdated or malfunctioning equipment should be avoided;
- Existing locations: available locations from other programs may be beneficially utilized;
- Locations used by HPMS;
- Distribution over geographical areas of the state and by functional class system; and
- Bias reduction: new locations should be randomly selected.

Although KDOT has access to traffic data from Kansas City's traffic management system, KC Scout, KDOT has questioned the accuracy of raw traffic data obtained from KC Scout sensors (Wavetronix) as well as the accuracy and quality of aggregate data provided by the KC Scout data portal. As such, KDOT employs portable traffic counters for data collection, resulting in excess work and missed opportunities to advantageously utilize current automatic data collection.

1.2 Objectives

The main objectives of this project included evaluation of the reliability of traffic data KC Scout uses to report performance measures and determination of more efficient use of available traffic data for monitoring. This research provides procedures for KDOT to effectively utilize KC Scout data for planning and operations. Research results will benefit the state of Kansas because data evaluation will identify strengths and opportunities for improvement in data accuracy and reliability, potentially reducing the time and effort required by KDOT personnel to analyze KC Scout data.

Chapter 2: Review of Current Practice

This chapter reviews KC Scout's current data cleaning, quality assessment, and aggregation practices and summarizes previous research findings that investigated the accuracy of nonintrusive radar sensors, such as Wavetronix sensors used within the Kansas City area freeway network. Because these data are essential for reporting purposes to the FHWA, this chapter also reviews the requirements of the HPMS.

2.1 KC Scout Data Aggregation and Quality Checks

KC Scout's detectors produce 30-second records (raw data) that are automatically aggregated into 5-minute records (TransCore ITS, 2016). Data quality checks are performed during the 30-second to 5-minute aggregation process and are included in the 5-minute records. 15-minute, 30-minute, hourly, and daily records can then be derived from the 5-minute records according to user-defined parameters when a report is run in the TransSuite Data Portal. This 5-minute to higher-level aggregation is performed by adding or averaging the corresponding values, such as vehicle count, average speed, average occupancy, and average data quality.

Table 2.1 shows the data quality checks performed on each 30-second record as it is aggregated into the 5-minute records. Minimum and maximum thresholds for volume, speed, occupancy, and density are configurable by TransCore and KC Scout system administrators; however, threshold values are not publicly available. If a quality check fails, equivalent action presented in Table 2.1 is taken. The worst (lowest) status value prevails for each lane and station in which multiple quality checks fail in the same interval. Quality statistics are then gathered for all 10 of the 30-second records and summarized in the 5-minute record. The percentages of valid volume, speed, and occupancy values for lanes and stations are each represented by a completeness factor. If all records are good (100% percentage of valid values), the completeness factor is 100, while a completeness factor between 0 and 100 indicates bad or missing records. For example, if 20% of the volume values, 10% of the speed values, and 40% of the occupancy values are invalid, the volume quality completeness factor will be 80, the speed quality completeness factor will be 90, and the occupancy quality completeness factor will be 60.

Table 2.1: KC Scout Data Quality Checks and Actions

Quality Control Test	Action
System error (special codes in the raw data indicate that the controller has detected an error or a function has been disabled)	Set values with error codes to null. Set LANE_STATUS of the appropriate lane to -9.
Erroneous timestamps and location IDs	Ignore records with invalid date, time, or location values. Set LANE_STATUS to -8.
Inconsistent elapsed time between records (polling period length drifting, polling cycle missing, controller accumulating data)	If polling period length is inconsistent, volume-based quality check rules should use volume flow rate instead of absolute count. Set LANE_STATUS to -7.
Duplicate records (caused by errors in data archiving logic or software process)	Ignore duplicate records. Set LANE_STATUS to -7.
Consecutive identical volume, occupancy, speed values (caused by hardware failure)	Set VOLUME, OCCUPANCY, SPEED to null. Set LANE_STATUS to -7.
No vehicles present	Set SPEED to null. Set LANE_STATUS to -6.
Extremely high-traffic volume	Set VOLUME to null. Set LANE_STATUS to -5.
Extremely high-occupancy value	Set VOLUME, OCCUPANCY, SPEED to null. Set LANE_STATUS to -5.
Extremely low or extremely high speed	Set SPEED to null. Set LANE_STATUS to -5.
Zero speed (when volume and occupancy are non-zero)	
Zero volume (when speed is non-zero)	Set VOLUME, OCCUPANCY, SPEED to null. Set LANE_STATUS to -5.
Extremely high density (improbable combination of volume and speed)	

The data quality procedure during the 30-second to 5-minute aggregation also attempts to impute bad or missing data and then recalculate aggregate values using the imputed data. Interpolation is initially attempted using upstream and/or downstream data. For example, if upstream and downstream data are of good quality, an average of the two is used for the imputed values. If no good downstream data are present but good upstream and next downstream are available or if no good upstream data are present but good downstream and previous upstream data are available, an average of the two is used for the imputed values. If only downstream (or upstream) data of good quality are available, the imputed values are set to equal these data. If interpolation is not possible or does not yield good data, the flagged data are set to a 2-year historical average, assuming the data are good; otherwise, the data require manual adjustment. The best values for each lane, whether originally measured, interpolated, historical, or manual, are used to recalculate aggregate station values. Records that include data imputed via the procedure described above are still flagged for bad data quality. However, it is not known how many of the bad-quality records correspond to manual entries, entries adjusted by interpolation, or historical data. Only the total number of imputed entries is available via the data portal.

2.2 Accuracy of Wavetronix Radar Sensors

KC Scout employs roadside-mounted microwave radar sensors to collect lane-specific traffic volume, speed, and vehicle classification data along the Kansas City area freeway network. Microwave radar sensors are advantageous because they are nonintrusive, can gather traffic flow data on multiple lanes, measure speed directly, and are not sensitive to lighting conditions or inclement weather such as rain, snow, and fog (Mimbela & Klein, 2007).

Several previous studies have evaluated the accuracy of microwave radar sensors. Coifman (2005) compared the performance of a side-fire traffic microwave sensor to loop detectors on I-80, north of Oakland, California. The study calculated the average absolute percent error in aggregated flow, occupancy, and speed for each of five eastbound lanes at 30-second and 5-minute aggregation levels. Results showed that microwave sensor flow measurements overestimated flow but were within 10% of loops for every lane except the lane nearest the median. Overcounting on that lane was attributed to “echoes off the concrete barrier.” In addition, aggregate sampling intervals (e.g., 5-minute intervals) generally produced a smaller percent error than disaggregate 30-second sampling intervals.

A subsequent study by the same researcher (Coifman, 2006) found 4.8% missed vehicle detections and 5.6% false vehicle detections, primarily due to occlusion and reflections, respectively. Lane-to-lane variation in count accuracy was also observed.

Klein and Kelley (1996) performed a comprehensive comparison of various traffic detection technologies for the Intelligent Vehicle Highway Systems (IVHS) study funded by FHWA. A total of 19 detectors were evaluated, including five microwave radars, three ultrasonic, two infrared, five video image processing (VIP), one acoustic, one inductive loop, and one magnetometer. Field testing sites in three states with diverse climates (Arizona, Florida, and

Minnesota) were selected to test detector performance under a variety of environmental conditions. Inductive loops produced the most accurate vehicle count results, with a 99% accuracy under both low- and high-volume conditions, but microwave radar sensors performed the best for nonintrusive methods. Microwave radars were also most impervious to inclement weather conditions, showing no difference in performance during rain, snow, high wind, and extreme cold or heat.

In a study for the Minnesota Department of Transportation (MnDOT), Kotzenmacher, Minge, and Hao (2005) evaluated three portable, side-fire, nonintrusive detectors: the SmartSensor by Wavetronix, the Remote Traffic Microwave Sensor (RTMS) by Electronic Integrated Systems (EIS), and the SAS-1 by SmarTek. Although all three detectors had similar, multiple-lane detection capabilities for volume, speed, and length-based vehicle classification, they were based on different technologies: digital radar for the SmartSensor, analog radar for the RTMS, and passive acoustic for the SAS-1. The Wavetronix SmartSensor exhibited the highest accuracy, with a per-lane volume detection error ranging from 1.4% to 4.9%.

Minge, Kotzenmacher, and Peterson (2010) conducted another detector accuracy study for MnDOT, testing five detector technologies: Wavetronix SmartSensor HD (microwave radar), Global Traffic Technologies (GTT) Canoga Microloops (magnetometer), PEEK AxleLight (laser), TIRTL (infrared; "The Infra-Red Traffic Logger"), and Miovision (video). The study found that the Wavetronix SmartSensor had a volume absolute percent error of less than 1.6% during free flow and congested conditions, as well as during extreme cold, rain, and snow. Fog conditions resulted in a greater volume error, although the error was still less than 5%. A per-vehicle analysis revealed occlusion when slow-moving trucks in the lane nearest the sensor blocked subsequent lanes, especially in periods of heavy congestion.

Performance of the Wavetronix SmartSensor microwave radar detector was also investigated by researchers at the University of Nebraska-Lincoln. Zhang (2006) found a 1.4% mean volume error and a 3.2% mean absolute volume error for 15-minute aggregation periods, while Grone (2012) reported a -5.1% mean volume error and an 8.2% absolute volume error for more disaggregate 1-minute periods. Lighting conditions or rain did not affect sensor accuracy in either study.

2.3 FHWA Highway Performance Monitoring System (HPMS) Data Requirements

The Highway Performance Monitoring System (HPMS) is a national information system that provides data that reflect "the extent, condition, performance, use, and operating characteristics of the nation's highways" (FHWA, 2016a). HPMS data are used to evaluate highway system performance according to FHWA's strategic planning process and to allocate federal funds to the states. HPMS data are also included in annual highway statistics publications and utilized by state agencies, Metropolitan Planning Organizations (MPOs), and transportation professionals.

States are required to annually collect and report the following data according to HPMS requirements:

- Full extent: limited data on all public roads, including data in the federal-aid system;
- Sample panel: detailed data for designated sections of the federal-aid system; and
- Statewide summary: information on travel, system length, and vehicle classification by functional system and area type.

Table 2.2 lists all traffic data required for the full extent or sample panel sections. Summary data include system length, vehicle miles traveled (VMT), and vehicle classification by functional system and area type (rural, small urban, and individual urbanized, non-attainment, and maintenance areas).

Table 2.2: Traffic Data Items Reported for HPMS

Data Item	Extent	
Annual Average Daily Traffic	Full Extent including ramps	
Single-Unit Truck & Bus AADT	Full Extent for some functional systems	Some Sample Panel Sections
Percent Peak Single-Unit Trucks and Buses		All Sample Panel Sections
Combination Truck AADT	Full Extent for some functional systems	Some Sample Panel Sections
Percent Peak Combination Trucks		All Sample Panel Sections
K-Factor		All Sample Panel Sections
Directional Factor		All Sample Panel Sections
Future AADT		All Sample Panel Sections
Capacity		All Sample Panel Sections

Source: FHWA (2016a)

2.4 FHWA Traffic Volume Trends (TVT) Data Requirements

In addition to annual HPMS reports, KDOT must submit a monthly Traffic Volume Trends (TVT) report to FHWA. Based on per-lane hourly traffic count data collected at more than 4,000 continuous traffic counting locations nationwide and the annual average daily traffic (AADT) by road segment from HPMS (FHWA, 2016b), TVT reports VMT on all public roads and estimates the percent change in traffic between the current month and the same month in the previous year for each functional roadway classification (FHWA, 2017). To collect data for the TVT reports, KDOT utilizes select permanent count locations across the state and four locations in the Kansas City area, as shown in Figure 2.1 (KDOT, 2017).

Kansas Department of Transportation — Significant Facts and Trends Report — Bureau of Transportation Planning — Page: 3

Month of October, 2017 — Sparse Matrix Model

Date: 14-NOV-2017

Route Description	Location Description	County	Average Daily Traffic for the Month 2015	2016	2017	Percent Change Month 2015	2016	Maximum Hour Volume	Hour	DOW	Day	Maximum Day Volume	DOW	Day	7 Day Avg	5 Day Avg	7 Day / 5 Day Avg
Kansas City Stations																	
US-69	0.4 MILES S. OF 135TH ST.	JOHNSON	56,911	58,495	61,377	7.8	4.9	6.227	4- 5 pm	TUE	31	72,683	FRI	13	61.377	65.646	.93
135TH	135TH E. OF MUR-LEN	JOHNSON	33,659	35,997	34,341	2.0	- 4.6	3.255	4- 5 pm	FRI	27	39,415	FRI	13	34.341	35.944	.96
BLACKBOB	BLACKBOB RD. S. OF 139TH	JOHNSON	22,093	21,844	21,606	- 2.2	- 1.1	2.131	5- 6 pm	FRI	6	23,781	FRI	6	21.606	22.144	.98
I 435	0.8 MILES S. OF I - 70	WYANDOTTE	71.507	72.538	73.000	2.1	0.6	8.516	5- 6 pm	THU	19	88.945	FRI	20	73.000	77.233	.95
Percent of Change for Kansas City Stations						3.3	.8										

Figure 2.1: KDOT October 2017 Fact and Trends Report, Kansas City Area Permanent Count Stations

Source: KDOT (2017)

Chapter 3: Data Collection

3.1 Location Selection

After consulting with KDOT, six locations of interest for obtaining data in the Kansas City metropolitan area were identified. Table 3.1 presents each location's description and coordinates, as well as the name and identification numbers of corresponding KC Scout detectors.

Table 3.1: Study Sites

Location Description	Coordinates	Detector ID	Location Reference	Detector Name
I-70 West of N 72nd St	39.106192, -94.748603	8277	K070EBVR-17	I-70 E @ 72nd Street
		8278	K070WBVR-17	I-70 W @ 72nd Street
I-435 North of Kansas Ave	39.094402, -94.808634	8307	K435SBVR-31	I-435 S @ North of Kansas Ave
		8308	K435NBVR-31	I-435 N @ North of Kansas Ave
I-635 North of Shawnee Dr	39.055389, -94.679583	7619	K635NBVR-03	I-635 N @ Shawnee Dr
		7620	K635SBVR-03	I-635 S @ Shawnee Dr
I-35 West of Cambridge Circle	39.075145, -94.611071	7726	K035NBV-57	I-35 N @ Cambridge Circle
		7732	K035SBV-57	I-35 S @ Cambridge Circle
I-35 at W 95th St	38.956160, -94.733229	7478	K035NBVR-05	I-35 N @ 95th St
		7672	K035SBVR-05	I-35 S @ 95th St
I-35 at Prairie St	38.889840, -94.789524	8239	K035SBVR-77	I-35 S @ Prairie Street
		8240	K035NBVR-77	I-35 N @ Prairie Street

3.2 Ground Truth Data

Ground truth data collection occurred in two phases. During the first phase, 6 continuous hours of field data were collected at each study location between June 29 and July 16, 2017. The second data collection phase, which occurred between October 28 and November 19, 2017, validated the counts collected during the previous phase and verified the accuracy of the investigated sensors. During the second data collection phase, data were obtained at all six study locations for 3 continuous hours. Table 3.2 presents the data collection information for each location and both phases.

Table 3.2: Data Collection Periods

Location	Data Collection Phase I			Data Collection Phase II		
	Date	Time	Weather Conditions	Date	Time	Weather Conditions
I-70 West of N 72nd St	06/29/2017	9:35–15:35	Heavy rain	10/29/2017	7:00–10:00	Good weather
I-435 North of Kansas Ave	07/16/2017	14:45–20:45	Good weather	11/18/2017	14:50–17:50	Good weather
I-635 North of Shawnee Dr	07/15/2017	11:00–17:00	Good weather	11/19/2017	8:25–11:25	Good weather
I-35 West of Cambridge Cir	06/30/2017	10:20–16:20	Good weather	10/29/2017	10:25–13:25	Good weather
I-35 at W 95th St	07/16/2017	7:50–13:50	Good weather	10/28/2017	13:05–16:05	Good weather
I-35 at Prairie St	07/01/2017	10:10–16:10	Good weather	10/28/2017	9:30–12:30	Good weather

For both data collection phases, the research team used video recording equipment (digital camcorder) operated on site from within a parked vehicle or placed on a tripod outside the roadside clear zone. Figures 3.1 through 3.6 depict positions of the recording equipment and the corresponding mapped locations of the KC Scout detectors. Adverse weather conditions (heavy rain) were present at the beginning of the first data collection period (on June 29) on I-70, but no precipitation occurred during the remaining periods. For the purposes of this study and to ensure the safety of the recording team, a highway permit was obtained from KDOT to allow use of the highway right-of-way. The permit covered all six locations, but the highway right-of-way was utilized in only three of the cases (I-435, I-635, and I-25 at 95th Street); available public parking spaces were used for the remaining three cases, as shown in Figures 3.1 through 3.6.

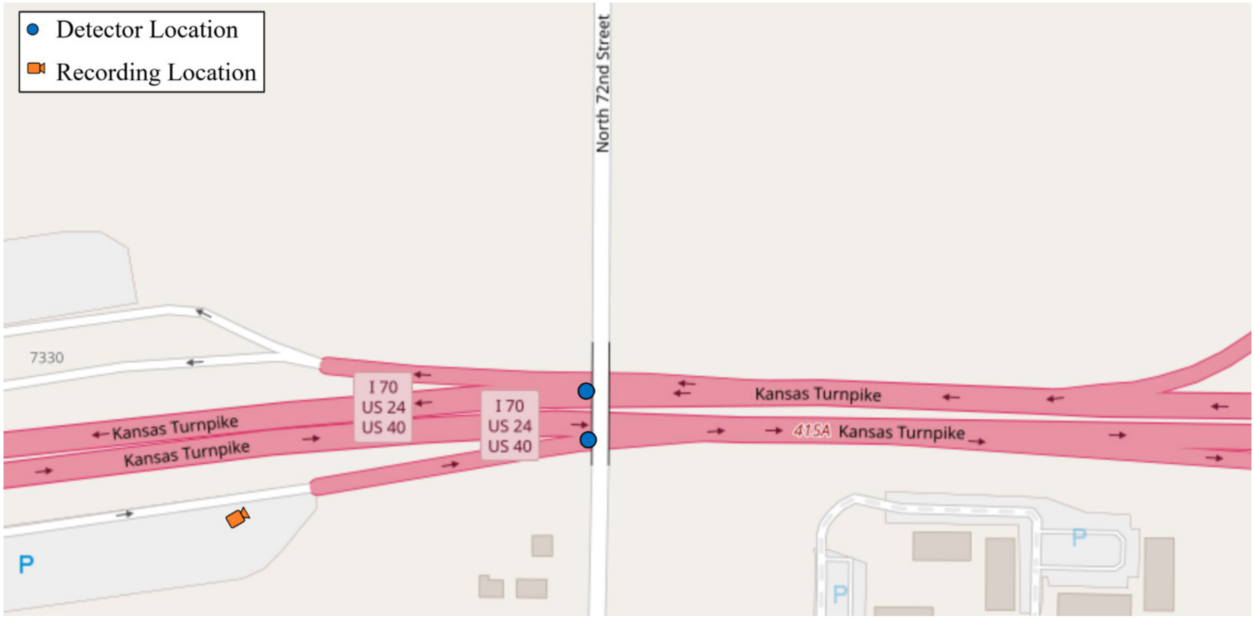


Figure 3.1: I-70 West of North 72nd Street, Detector and Recording Locations

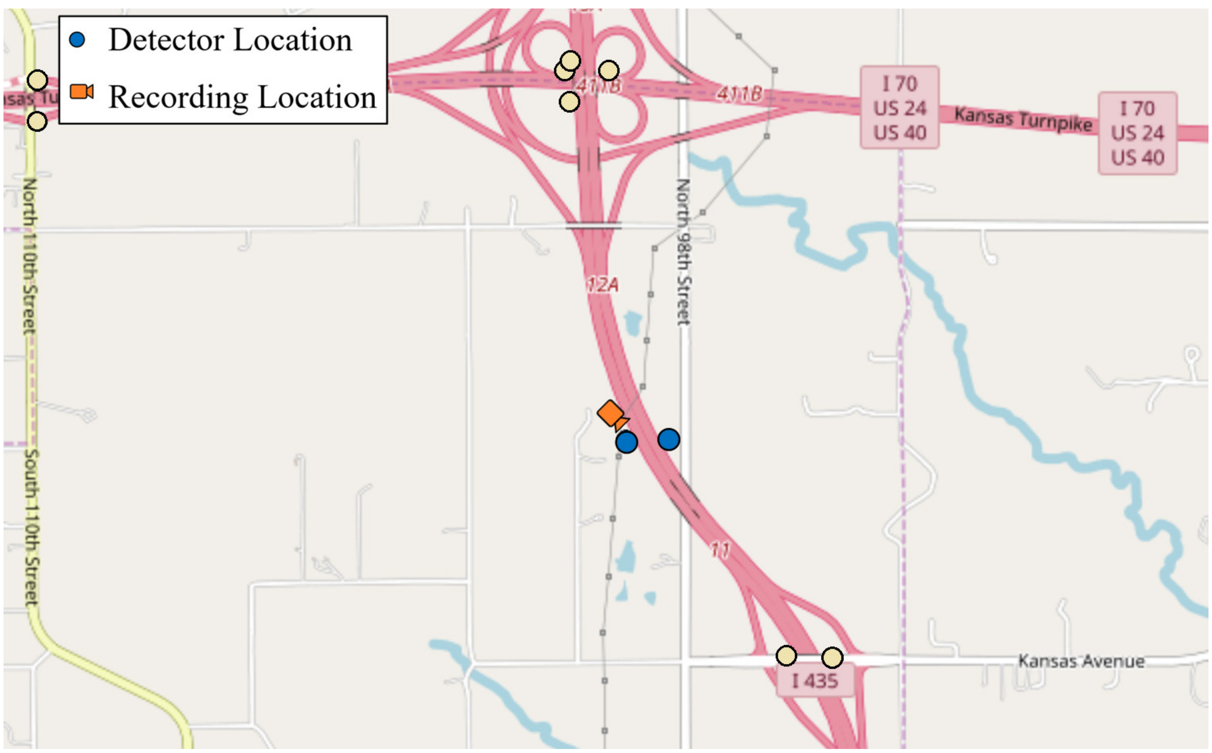


Figure 3.2: I-435 North of Kansas Avenue, Detector and Recording Locations

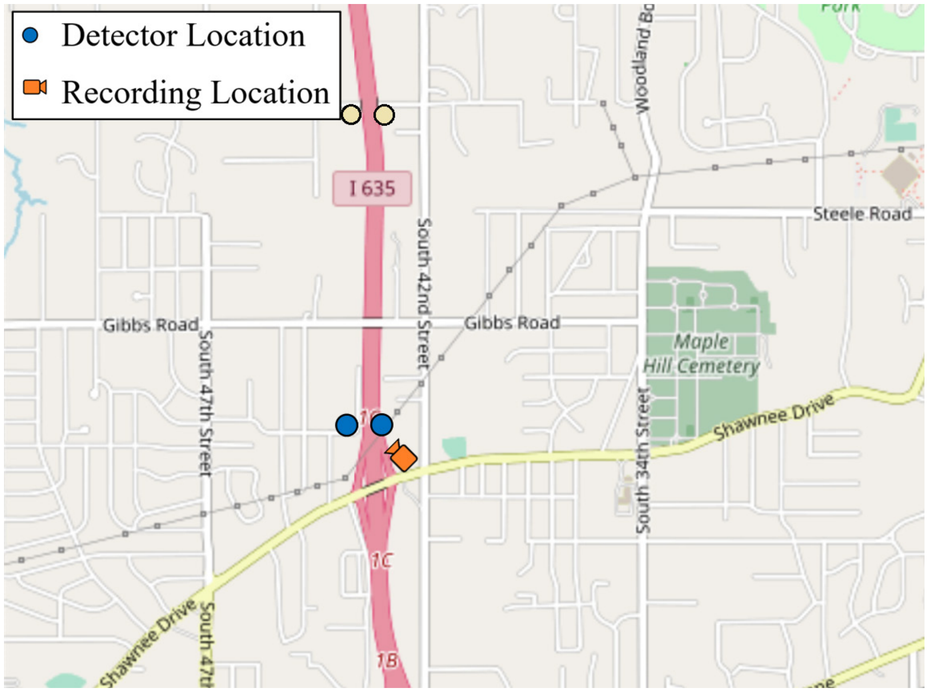


Figure 3.3: I-635 North of Shawnee Drive, Detector and Recording Locations

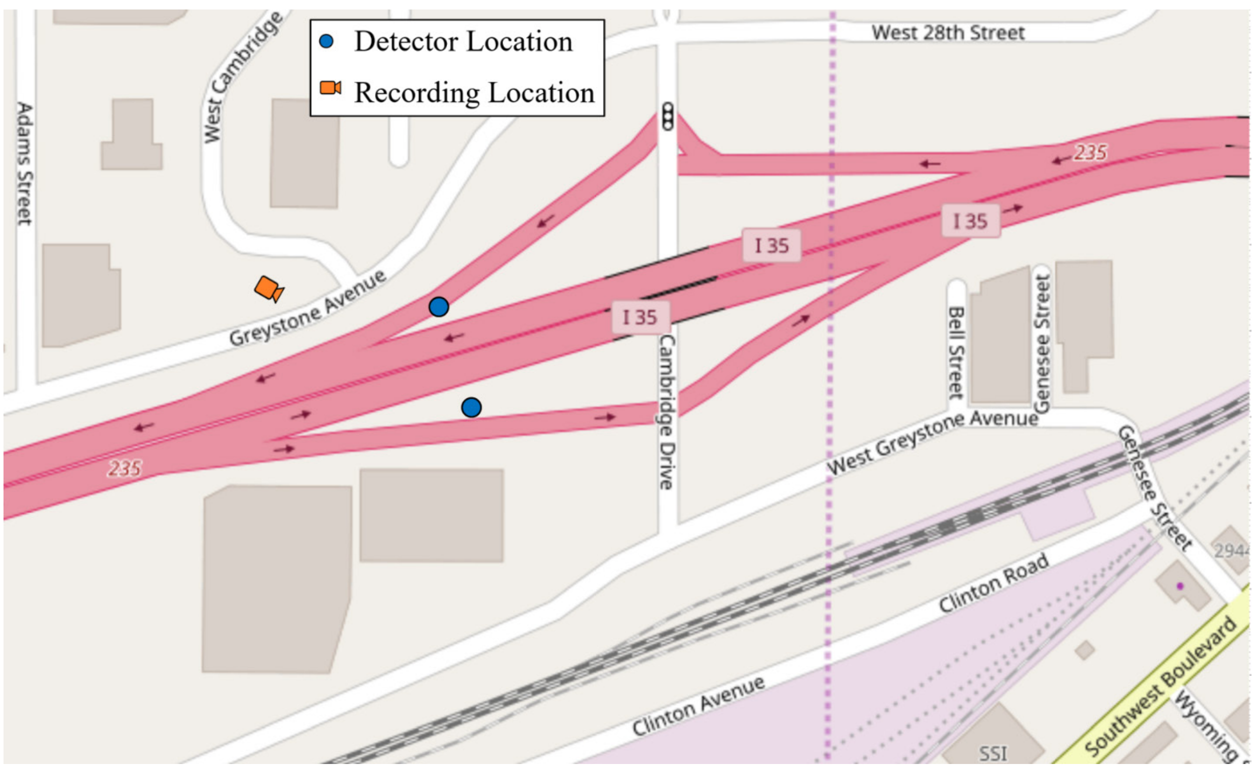


Figure 3.4: I-35 West of Cambridge Circle, Detector and Recording Locations

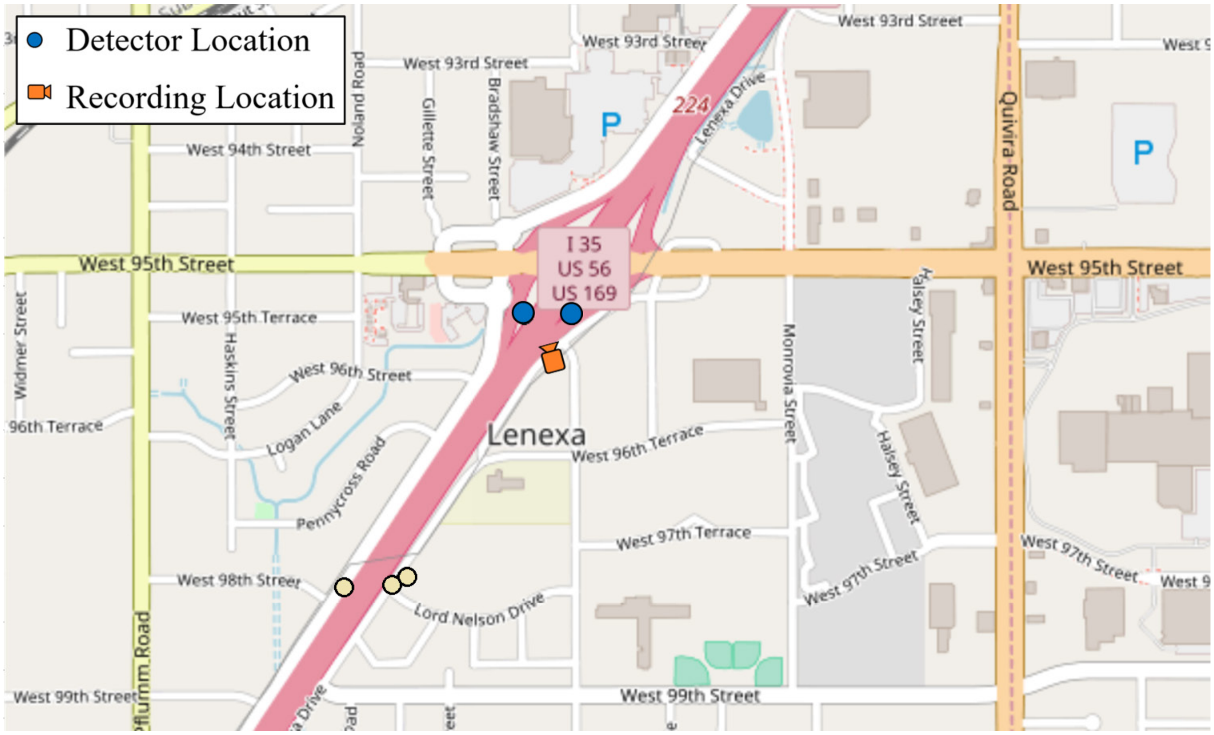


Figure 3.5: I-35 at West 95th Street, Detector and Recording Locations

Figure 3.6: I-35 at Prairie Street, Detector and Recording Locations

3.3 Sensor Data

For the 9-hour time periods (6 hours during first phase and 3 hours during second phase) of the field data collections (Table 3.2), raw (30 second) traffic data records for both directions at each location were obtained from the KC Scout database. The raw detector counts were compared to the ground truth data to evaluate each sensor's accuracy. To compare the data quality reported by KC Scout with the data quality calculated based on raw data (Section 5.2), 5-minute reports for the same time periods were also obtained.

3.4 System Reliability Data

Overall system reliability was estimated using 1 year (November 1, 2016, to October 31, 2017) of raw (30 second) traffic data records obtained from the KC Scout database for each of the 12 detectors in Table 3.1. A step-by-step guide for obtaining raw or aggregate sensor data from the KC Scout portal is provided in Appendix A. Using the quality control checks of Table 2.1, data records were manually flagged for bad quality according to the following criteria:

- Reported system errors,
- Reported lane status errors,
- Missing time intervals,
- Consecutive duplicate values,
- Extremely high volume (greater than 25 veh/lane/30-sec interval),
- Extremely high speed (greater than 90 mph),
- Extremely high occupancy (greater than 90),
- No volume but speed of non-zero,
- Speed of zero but volume of non-zero, and
- Extremely high density (greater than 250 veh/lane/mile).

This study obtained the listed thresholds from the literature (Nihan, Jacobson, Bender, & Davis, 1990; Jacobson, Nihan, & Bender, 1990; Turner, Margiotta, & Lomax, 2004; Elefteriadou et al., 2011) since the exact values used by KC Scout were not known. The annual percentages of good-quality and bad-quality data for each detector were calculated.

Chapter 4: Analysis of Data Accuracy

4.1 Comparison Methodology

The traffic volume for each direction (ground truth field data) was obtained via manual observation of the video recordings. These volumes were then compared to the equivalent raw 30-second volume counts available in the KC Scout database (detector data) to evaluate sensor accuracy. To avoid errors due to bad synchronization between the KC Scout clock and ours, we compared volumes per hour of data collection. Thus, both the manual traffic counts and the raw detector data were aggregated in 1-hour periods. The difference between the field data and the KC Scout detector data in each 1-hour period was expressed as a percentage:

$$Difference\ (\%) = \frac{Field\ Volume - Detector\ Volume}{Field\ Data\ Volume} * 100\% \qquad \textbf{Equation 4.1}$$

In addition, the absolute hourly difference was calculated as:

$$Abs.Difference\ (\%) = \left|\frac{Field\ Volume - Detector\ Volume}{Field\ Data\ Volume}\right| * 100\% \qquad \textbf{Equation 4.2}$$

Finally, the average difference for each detector during each data collection period was calculated as the mean of the respective hourly differences, and the average absolute difference was the mean of the absolute hourly differences.

Although the data reporting requirements of the TVT report include hourly traffic count data by lane, lane-specific accuracy was not evaluated for the following reasons: first, the video footage, obtained from near-ground level on the side of the road, could not be used in most cases to distinguish the lane of each vehicle, and second, most observation locations were marginally downstream or upstream of the actual detectors and any lane changing occurring between the two spots would not be taken into account.

4.2 Summarized Results

4.2.1 First Data Collection Phase

As shown in Table 4.1, for the first data collection phase, the average difference between the sensor data and the field data was less than 5% for seven directional sites. The detectors underestimated the volumes of three directional sites by approximately 6%, 10%, and 14%, respectively, but the detectors overestimated the volumes of the remaining two directional sites by approximately 18% and 30%, respectively. However, this last overestimation was deemed invalid because the quality of the detector's volume data was nearly zero during the data collection period; thus, the detector's measurements could not be compared to the ground truth data.

Table 4.1: Comparison of Field Data and KC Scout Data for Phase I

KC Scout Detector Location		Data Collection Period		Average Difference	Average Absolute Difference
ID	Station Name	Date	Time		
8277	I-70 E @ 72nd St	06/29/2017	9:35–15:35	-17.95%	17.95%
8278	I-70 W @ 72nd St			-29.83%	29.83%
7726	I-35 N @ Cambridge Cir	06/30/2017	10:20–16:20	-2.41%	2.41%
7732	I-35 S @ Cambridge Cir			-0.70%	0.87%
8239	I-35 S @ Prairie St	07/01/2017	10:10–16:10	-1.87%	2.84%
8240	I-35 N @ Prairie St			5.56%	5.56%
7478	I-35 N @ 95th St	07/16/2017	7:50–13:50	14.09%	14.09%
7672	I-35 S @ 95th St			-1.57%	1.61%
8307	I-435 S, North of Kansas Ave	07/16/2017	14:45–20:45	0.56%	0.63%
8308	I-435 N, North of Kansas Ave			0.88%	0.91%
7619	I-635 N @ Shawnee Dr	07/15/2017	11:00–17:00	-1.22%	1.61%
7620	I-635 S @ Shawnee Dr			9.87%	9.87%

4.2.2 Second Data Collection Phase

The radar sensors were determined to be configured incorrectly (e.g., missing lanes, wrong lanes assigned, etc.) at locations with high percentage differences. Once KC Scout corrected the sensor configurations, the research team repeated the data collection at all six locations for a 3-hour period. Table 4.2 summarizes the average percent difference between detector and field-measured counts for the second data collection phase. The results showed that the average absolute difference between the sensor data and the field data was less than 3%, and in many cases as low as approximately 1%, for all 12 directional sites.

Table 4.2: Comparison of Field Data and KC Scout Data for Phase II

KC Scout Detector Location		Data Collection Period		Average Difference	Average Absolute Difference
ID	Station Name	Date	Time		
8277	I-70 E @ 72nd St	10/29/2017	7:00–10:00	-0.92%	1.56%
8278	I-70 W @ 72nd St			-0.81%	0.81%
7726	I-35 N @ Cambridge Cir	10/29/2017	10:25–13:25	-1.95%	1.95%
7732	I-35 S @ Cambridge Cir			-0.74%	1.16%
8239	I-35 S @ Prairie St	10/28/2017	9:30–12:30	2.22%	2.42%
8240	I-35 N @ Prairie St			-0.06%	2.15%
7478	I-35 N @ 95th St	10/28/2017	13:05–16:05	-1.10%	1.10%
7672	I-35 S @ 95th St			-1.52%	1.52%
8307	I-435 S, North of Kansas Ave	11/18/2017	14:50–17:50	0.76%	0.76%
8308	I-435 N, North of Kansas Ave			0.82%	0.82%
7619	I-635 N @ Shawnee Dr	11/19/2017	8:25–11:25	-1.05%	1.05%
7620	I-635 S @ Shawnee Dr			-1.51%	1.51%

4.3 Detailed Results

4.3.1 Eastbound I-70 at 72nd Street

Table 4.3 compares hourly volumes measured by direct observation (ground truth field data) and hourly volumes measured by corresponding detectors for the eastbound direction of I-70 at 72nd Street for both data collection phases. The significant difference observed in Phase I occurred because the wrong lanes were assigned to the detector. Once the lane assignment was corrected in Phase II, the hourly difference was less than 2%.

Table 4.3: Comparison of Field Data and KC Scout Data at I-70 E at 72nd Street

Phase	Date	Time Period	Hourly Volume		Hourly Difference	Absolute Hourly Difference
			Field Data	Detector		
I	06/29/2017	9:35–10:35	1714	2031	-18.49%	18.49%
I	06/29/2017	10:35–11:35	1722	2064	-19.86%	19.86%
I	06/29/2017	11:35–12:35	1853	2167	-16.95%	16.95%
I	06/29/2017	12:35–13:35	1941	2286	-17.77%	17.77%
I	06/29/2017	13:35–14:35	2019	2394	-18.57%	18.57%
I	06/29/2017	14:35–15:35	2209	2564	-16.07%	16.07%
First Data Collection Phase Average Hourly Difference					-17.95%	17.95%
II	10/29/2017	7:00–8:00	598	610	-2.01%	2.01%
II	10/29/2017	8:00–9:00	877	892	-1.71%	1.71%
II	10/29/2017	9:00–10:00	1353	1340	0.96%	0.96%
Second Data Collection Phase Average Hourly Difference					-0.92%	1.56%

4.3.2 Westbound I-70 at 72nd Street

Table 4.4 compares hourly volumes measured by direct observation and hourly volumes measured by corresponding detectors for the westbound direction of I-70 at 72nd Street for both data collection phases. The significant difference observed in Phase I occurred because the wrong lanes were assigned to the detector and because the reported data quality was very low. Once the lane assignment was corrected in Phase II, the hourly difference was less than 1.5%.

Table 4.4: Comparison of Field Data and KC Scout Data at I-70 W at 72nd Street

Phase	Date	Time Period	Hourly Volume		Hourly Difference	Absolute Hourly Difference
			Field Data	Detector		
I	06/29/2017	9:35–10:35	1818	2578	-41.80%	41.80%
I	06/29/2017	10:35–11:35	1551	1794	-15.67%	15.67%
I	06/29/2017	11:35–12:35	1921	2772	-44.30%	44.30%
I	06/29/2017	12:35–13:35	2260	2713	-20.04%	20.04%
I	06/29/2017	13:35–14:35	1978	2804	-41.76%	41.76%
I	06/29/2017	14:35–15:35	2217	2558	-15.38%	15.38%
First Data Collection Phase Average Hourly Difference					-29.83%	29.83%
II	10/29/2017	7:00–8:00	675	684	-1.33%	1.33%
II	10/29/2017	8:00–9:00	861	865	-0.46%	0.46%
II	10/29/2017	9:00–10:00	1132	1139	-0.62%	0.62%
Second Data Collection Phase Average Hourly Difference					-0.81%	0.81%

4.3.3 Northbound I-35 at Cambridge Circle

Table 4.5 compares hourly volumes measured by direct observation and hourly volumes measured by corresponding detectors for northbound I-35 west of Cambridge Circle for both data collection phases. The hourly difference was less than 3% for all nine 1-hour periods.

Table 4.5: Comparison of Field Data and KC Scout Data at I-35 N at Cambridge Circle

Phase	Date	Time Period	Hourly Volume		Hourly Difference	Absolute Hourly Difference
			Field Data	Detector		
I	06/30/2017	10:20–11:20	2719	2793	-2.72%	2.72%
I	06/30/2017	11:20–12:20	2919	3006	-2.98%	2.98%
I	06/30/2017	12:20–13:20	3181	3234	-1.67%	1.67%
I	06/30/2017	13:20–14:20	3007	3067	-2.00%	2.00%
I	06/30/2017	14:20–15:20	3148	3242	-2.99%	2.99%
I	06/30/2017	15:20–16:20	3482	3555	-2.10%	2.10%
First Data Collection Phase Average Hourly Difference					-2.41%	2.41%
II	10/29/2017	10:25–11:25	1956	1979	-1.18%	1.18%
II	10/29/2017	11:25–12:25	2151	2200	-2.28%	2.28%
II	10/29/2017	12:25–13:25	2464	2523	-2.39%	2.39%
Second Data Collection Phase Average Hourly Difference					-1.95%	1.95%

4.3.4 Southbound I-35 at Cambridge Circle

Table 4.6 compares hourly volumes measured by direct observation and hourly volumes measured by corresponding detectors for southbound I-35 west of Cambridge Circle for both data collection phases. The hourly difference was less than 2% for all nine 1-hour periods.

Table 4.6: Comparison of Field Data and KC Scout Data at I-35 S at Cambridge Circle

Phase	Date	Time Period	Hourly Volume		Hourly Difference	Absolute Hourly Difference
			Field Data	Detector		
I	06/30/2017	10:20–11:20	2616	2603	0.50%	0.50%
I	06/30/2017	11:20–12:20	2953	2971	-0.61%	0.61%
I	06/30/2017	12:20–13:20	3084	3124	-1.30%	1.30%
I	06/30/2017	13:20–14:20	3140	3200	-1.91%	1.91%
I	06/30/2017	14:20–15:20	3587	3598	-0.31%	0.31%
I	06/30/2017	15:20–16:20	4020	4044	-0.60%	0.60%
First Data Collection Phase Average Hourly Difference					-0.70%	0.87%
II	10/29/2017	10:25–11:25	1684	1712	-1.66%	1.66%
II	10/29/2017	11:25–12:25	2105	2130	-1.19%	1.19%
II	10/29/2017	12:25–13:25	2335	2320	0.64%	0.64%
Second Data Collection Phase Average Hourly Difference					-0.74%	1.16%

4.3.5 Southbound I-35 at Prairie Street

Table 4.7 compares hourly volumes measured by direct observation and hourly volumes measured by corresponding detectors for southbound I-35 at Prairie Street for both data collection phases. The recording location for this site was approximately 1,000 ft from the detector in the southbound direction. The hourly difference barely exceeded 5% for a 1-hour period and was less for the remaining eight periods.

Table 4.7: Comparison of Field Data and KC Scout Data at I-35 S at Prairie Street

Phase	Date	Time Period	Hourly Volume		Hourly Difference	Absolute Hourly Difference
			Field Data	Detector		
I	07/01/2017	10:10–11:10	2752	2725	0.98%	0.98%
I	07/01/2017	11:10–12:10	3147	3086	1.94%	1.94%
I	07/01/2017	12:10–13:10	3012	3132	-3.98%	3.98%
I	07/01/2017	13:10–14:10	3103	3175	-2.32%	2.32%
I	07/01/2017	14:10–15:10	2995	3140	-4.84%	4.84%
I	07/01/2017	15:10–16:10	2966	3054	-2.97%	2.97%
First Data Collection Phase Average Hourly Difference					-1.87%	2.84%
II	10/28/2017	9:30–10:30	2483	2358	5.03%	5.03%
II	10/28/2017	10:30–11:30	2736	2744	-0.29%	0.29%
II	10/28/2017	11:30–12:30	3060	3001	1.93%	1.93%
Second Data Collection Phase Average Hourly Difference					2.22%	2.42%

4.3.6 Northbound I-35 at Prairie Street

Table 4.8 compares hourly volumes measured by direct observation and hourly volumes measured by corresponding detectors for northbound I-35 at Prairie Street for both data collection phases. The 5–6% difference occurred because the detector missed one lane (counting only four of the five lanes). Once that oversight was corrected, however, the difference was approximately 3%, even when accounting for the aforementioned inconvenient distance between the recording location and the detector.

Table 4.8: Comparison of Field Data and KC Scout Data at I-35 N at Prairie Street

Phase	Date	Time Period	Hourly Volume		Hourly Difference	Absolute Hourly Difference
			Field Data	Detector		
I	07/01/2017	10:10–11:10	3759	3545	5.69%	5.69%
I	07/01/2017	11:10–12:10	3771	3568	5.38%	5.38%
I	07/01/2017	12:10–13:10	3654	3464	5.20%	5.20%
I	07/01/2017	13:10–14:10	3430	3256	5.07%	5.07%
I	07/01/2017	14:10–15:10	3381	3181	5.92%	5.92%
I	07/01/2017	15:10–16:10	3417	3209	6.09%	6.09%
First Data Collection Phase Average Hourly Difference					5.56%	5.56%
II	10/28/2017	9:30–10:30	3227	3126	3.13%	3.13%
II	10/28/2017	10:30–11:30	3467	3543	-2.19%	2.19%
II	10/28/2017	11:30–12:30	3639	3680	-1.13%	1.13%
Second Data Collection Phase Average Hourly Difference					-0.06%	2.15%

4.3.7 Northbound I-35 at 95th Street

Table 4.9 compares hourly volumes measured by direct observation and hourly volumes measured by corresponding detectors for northbound I-35 at West 95th Street for both data collection phases. The 14% difference occurred because the detector missed one lane (counting only three of the four lanes) due to improper recalibration after a fourth lane was added. Once that oversight was corrected, however, the difference was approximately 1.5%.

Table 4.9: Comparison of Field Data and KC Scout Data at I-35 N at 95th Street

Phase	Date	Time Period	Hourly Volume		Hourly Difference	Absolute Hourly Difference
			Field Data	Detector		
I	07/16/2017	7:50–8:50	948	842	11.18%	11.18%
I	07/16/2017	8:50–9:50	1373	1172	14.64%	14.64%
I	07/16/2017	9:50–10:50	1717	1489	13.28%	13.28%
I	07/16/2017	10:50–11:50	1840	1562	15.11%	15.11%
I	07/16/2017	11:50–12:50	2090	1781	14.78%	14.78%
I	07/16/2017	12:50–13:50	2062	1741	15.57%	15.57%
First Data Collection Phase Average Hourly Difference					14.09%	14.09%
II	10/28/2017	13:05–14:05	2412	2430	-0.75%	0.75%
II	10/28/2017	14:05–15:05	2535	2561	-1.03%	1.03%
II	10/28/2017	15:05–16:05	2399	2436	-1.54%	1.54%
Second Data Collection Phase Average Hourly Difference					-1.10%	1.10%

4.3.8 Southbound I-35 at 95th Street

Table 4.10 compares hourly volumes measured by direct observation and hourly volumes measured by corresponding detectors for southbound I-35 at West 95th Street for both data collection phases. The hourly difference was less than 4% for all nine 1-hour periods and less than 2% for all but two of the periods.

Table 4.10: Comparison of Field Data and KC Scout Data at I-35 S at 95th Street

Phase	Date	Time Period	Hourly Volume		Hourly Difference	Absolute Hourly Difference
			Field Data	Detector		
I	07/16/2017	7:50–8:50	861	860	0.12%	0.12%
I	07/16/2017	8:50–9:50	1221	1237	-1.31%	1.31%
I	07/16/2017	9:50–10:50	1574	1588	-0.89%	0.89%
I	07/16/2017	10:50–11:50	1735	1775	-2.31%	2.31%
I	07/16/2017	11:50–12:50	2050	2078	-1.37%	1.37%
I	07/16/2017	12:50–13:50	2050	2125	-3.66%	3.66%
First Data Collection Phase Average Hourly Difference					-1.57%	1.61%
II	10/28/2017	13:05–14:05	2477	2521	-1.78%	1.78%
II	10/28/2017	14:05–15:05	2239	2274	-1.56%	1.56%
II	10/28/2017	15:05–16:05	2438	2468	-1.23%	1.23%
Second Data Collection Phase Average Hourly Difference					-1.52%	1.52%

4.3.9 Southbound I-435, North of Kansas Avenue

Table 4.11 compares hourly volumes measured by direct observation and hourly volumes measured by corresponding detectors for southbound I-435, north of Kansas Avenue for both data collection phases. The hourly difference was less than 1% for all but one 1-hour period; that one period was slightly greater than 2%.

Table 4.11: Comparison of Field Data and KC Scout Data at I-435 S, North of Kansas Avenue

Phase	Date	Time Period	Hourly Volume		Hourly Difference	Absolute Hourly Difference
			Field Data	Detector		
I	07/16/2017	14:45–15:45	1946	1949	-0.15%	0.15%
I	07/16/2017	15:45–16:45	1841	1836	0.27%	0.27%
I	07/16/2017	16:45–17:45	1757	1718	2.22%	2.22%
I	07/16/2017	17:45–18:45	1396	1397	-0.07%	0.07%
I	07/16/2017	18:45–19:45	1190	1182	0.67%	0.67%
I	07/16/2017	19:45–20:45	1013	1009	0.39%	0.39%
First Data Collection Phase Average Hourly Difference					0.56%	0.63%
II	11/18/2017	14:50–15:50	2231	2209	0.99%	0.99%
II	11/18/2017	15:50–15:50	2015	1997	0.89%	0.89%
II	11/18/2017	16:50–17:50	2195	2186	0.41%	0.41%
Second Data Collection Phase Average Hourly Difference					0.76%	0.76%

4.3.10 Northbound I-435, North of Kansas Avenue

Table 4.12 compares hourly volumes measured by direct observation and hourly volumes measured by corresponding detectors for northbound I-435, north of Kansas Avenue for both data collection phases. The hourly difference was less than 1% for six 1-hour periods, and less than 2.5% for all nine 1-hour periods.

Table 4.12: Comparison of Field Data and KC Scout Data at I-435 N, North of Kansas Avenue

Phase	Date	Time Period	Hourly Volume		Hourly Difference	Absolute Hourly Difference
			Field Data	Detector		
I	07/16/2017	14:45–15:45	2060	2055	0.24%	0.24%
I	07/16/2017	15:45–16:45	2035	2023	0.59%	0.59%
I	07/16/2017	16:45–17:45	2072	2035	1.79%	1.79%
I	07/16/2017	17:45–18:45	1900	1854	2.42%	2.42%
I	07/16/2017	18:45–19:45	1434	1435	-0.07%	0.07%
I	07/16/2017	19:45–20:45	1179	1175	0.34%	0.34%
First Data Collection Phase Average Hourly Difference					0.88%	0.91%
II	11/18/2017	14:50–15:50	2447	2435	0.49%	0.49%
II	11/18/2017	15:50–15:50	2350	2332	0.77%	0.77%
II	11/18/2017	16:50–17:50	2462	2432	1.22%	1.22%
Second Data Collection Phase Average Hourly Difference					0.82%	0.82%

4.3.11 Northbound I-635 at Shawnee Drive

Table 4.13 compares hourly volumes measured by direct observation and hourly volumes measured by corresponding detectors for northbound I-635 north of Shawnee Drive for both data collection phases. The hourly difference was less than 5% all nine 1-hour periods and less than 2% for seven of the periods.

Table 4.13: Comparison of Field Data and KC Scout Data at I-635 N at Shawnee Drive

Phase	Date	Time Period	Hourly Volume		Hourly Difference	Absolute Hourly Difference
			Field Data	Detector		
I	07/15/2017	11:00–12:00	1894	1876	0.95%	0.95%
I	07/15/2017	12:00–13:00	2084	2106	-1.06%	1.06%
I	07/15/2017	13:00–14:00	1997	2052	-2.75%	2.75%
I	07/15/2017	14:00–15:00	2083	2078	0.24%	0.24%
I	07/15/2017	15:00–16:00	2201	2301	-4.54%	4.54%
I	07/15/2017	16:00–17:00	2128	2131	-0.14%	0.14%
First Data Collection Phase Average Hourly Difference					-1.22%	1.61%
II	11/19/2017	8:25–9:25	864	873	-1.04%	1.04%
II	11/19/2017	9:25–10:25	1225	1236	-0.90%	0.90%
II	11/19/2017	10:25–11:25	1470	1488	-1.22%	1.22%
Second Data Collection Phase Average Hourly Difference					-1.05%	1.05%

4.3.12 Southbound I-635 at Shawnee Drive

Table 4.14 compares hourly volumes measured by direct observation and hourly volumes measured by corresponding detectors for southbound I-635 north of Shawnee Drive for both data collection phases. A 10% difference occurred because the manual counts included the volume of the exit-only ramp lane, but the detector did not count exiting vehicles. In Phase II, however, the manual count ignored all exiting vehicles, so the difference was less than 2%.

Table 4.14: Comparison of Field Data and KC Scout Data at I-635 S at Shawnee Drive

Phase	Date	Time Period	Hourly Volume		Hourly Difference	Absolute Hourly Difference
			Field Data	Detector		
I	07/15/2017	11:00–12:00	2158	1981	8.20%	8.20%
I	07/15/2017	12:00–13:00	2241	2066	7.81%	7.81%
I	07/15/2017	13:00–14:00	2248	2071	7.87%	7.87%
I	07/15/2017	14:00–15:00	2391	2091	12.55%	12.55%
I	07/15/2017	15:00–16:00	2316	2042	11.83%	11.83%
I	07/15/2017	16:00–17:00	2278	2028	10.97%	10.97%
First Data Collection Phase Average Hourly Difference					9.87%	9.87%
II	11/19/2017	8:25–9:25	992	1001	-0.91%	0.91%
II	11/19/2017	9:25–10:25	1363	1388	-1.83%	1.83%
II	11/19/2017	10:25–11:25	1740	1771	-1.78%	1.78%
Second Data Collection Phase Average Hourly Difference					-1.51%	1.51%

Chapter 5: Assessment of System Reliability

This section presents reliability analysis of count data at the selected sensor locations. Downtime frequency, percentage of missing or invalid data, and percentage of imputed values were analyzed and reported for a 1-year period (from November 1, 2016, to October 31, 2017) and for 9 hours of field data collection in 2017. The analysis utilized quality control checks described in Table 2.1 and threshold values listed in Section 3.4 to flag the raw (30 second, per lane) traffic data records for bad quality. Output of the process was the number and percentage of missing and bad-quality records per error category for each selected time period. Data quality reported by KC Scout was then compared with process results. The following sections describe analysis results for the 1-year period and the 9 hours of data collection.

5.1 One-Year Reliability Assessment

One-year reliability assessment for each detector was performed in four 3-month periods due to computational restrictions of voluminous data records. System reliability and consistency were also assessed within the 1-year period. Results showed that the following four data-quality criteria (Section 3.4) caused more than 99% of the flagged data records for all detectors:

- Missing time intervals,
- Duplicate records (consecutive identical volume, occupancy, and speed values caused by hardware failure or errors in data archiving logic or software process),
- Extremely high-traffic volume (greater than 25 veh/lane/30-sec interval), and
- Zero volume with non-zero speed.

Duplicate records and zero volume with non-zero speed errors were significant in eight of the 12 detectors, while high-traffic volume errors were significant in three detectors. A significant percentage of missing intervals was observed on only one pair of detectors. The remaining six criteria (i.e., system errors, lane status errors, zero speed with non-zero volume, extremely high speed, extremely high occupancy, and extremely high density) were not statistically significant for all detectors. Table 5.1 shows the 1-year data quality for each detector expressed as the percentage of data missing or flagged for bad quality, as well as the percentages corresponding to each of the four error criteria that were significant.

Table 5.1: One-Year Data Quality Analysis (November 2016 to October 2017)

Station Name	Percentage of Missing or Bad Quality Data	Error Criteria			
		Missing Intervals	Duplicate Records	High Volume	Volume=0 Speed≠0
I-70 E @ 72nd St	18.99%	1.35%	6.47%	0.18%	10.88%
I-70 W @ 72nd St	23.95%	0.52%	10.19%	3.23%	9.82%
I-35 N @ Cambridge Cir	2.41%	0.76%	1.05%	0.00%	0.09%
I-35 S @ Cambridge Cir	2.60%	0.91%	0.86%	0.00%	0.12%
I-35 S @ Prairie St	14.67%	0.70%	5.29%	0.00%	8.61%
I-35 N @ Prairie St	16.32%	0.67%	6.13%	0.02%	9.42%
I-35 N @ 95th St	40.27%	8.42%	15.73%	7.76%	8.27%
I-35 S @ 95th St	39.04%	8.41%	15.04%	8.21%	7.24%
I-435 S, North of Kansas Ave	27.17%	0.41%	10.67%	0.05%	15.67%
I-435 N, North of Kansas Ave	23.07%	0.42%	8.91%	0.00%	13.44%
I-635 N @ Shawnee Dr	2.03%	0.89%	1.10%	0.00%	0.00%
I-635 S @ Shawnee Dr	2.06%	0.88%	1.13%	0.00%	0.00%

Data were unavailable for the two detectors at I-35 at 95th Street from November 1, 2016, to November 29, 2016, due to construction. These 29 days were excluded from the data quality analysis of these two detectors; only the remaining 337 days of the year were used (the extra day being due to the leap year). This unavailability demonstrated that construction can negatively affect data quality. Therefore, detectors should be monitored more thoroughly if construction takes place close to their location.

Table 5.1 shows that data quality varied significantly between detectors, although most detector pairs at each location exhibited similar characteristics. Two pairs of detectors showed excellent year-long quality (approximately 98% good-quality data), while the other four pairs demonstrated approximately 85%, 79%, 75%, and 60% good-quality data, respectively.

The causes of bad-quality data varied among detector pairs. Flow-related criteria (last two columns of Table 5.1) failed on four detector pairs, potentially indicating that the detector needed calibration or maintenance. Duplicate records consistently accompanied the flow-related errors and were responsible for a significant percentage of bad-quality data (33–55%).

Data quality consistency throughout the four 3-month periods for each detector was investigated. Results shown in Table 5.2 indicate that quality was relatively consistent throughout the year, except for the detector at I-35 at 95th Street, where the data quality ranged from approximately 82% in the first 6 months to approximately 57% for the last 6 months of the year. The causes of error also varied between these two periods, with almost zero errors due to extremely high volume in the first six months but more than 200,000 high-volume errors (approximately 33% of all error causes) in the last 6 months. Duplicate value errors were also three to five times more common in the last 6 months. These results are potential indications of detector hardware malfunctions.

Table 5.2: Data Quality Analysis Throughout the Year

Station Name	Data Quality				
	11/01/2016–10/31/2017	11/01/2016–1/31/2017	2/01/2017–4/30/2017	5/01/2017–7/31/2017	8/01/2017–10/31/2017
I-70 E @ 72nd St	81.01%	79.71%	81.65%	82.55%	80.13%
I-70 W @ 72nd St	76.05%	79.66%	82.48%	59.78%	82.51%
I-35 N @ Cambridge Cir	97.59%	96.67%	97.68%	97.23%	97.85%
I-35 S @ Cambridge Cir	97.40%	96.74%	97.68%	97.25%	97.95%
I-35 S @ Prairie St	85.33%	83.58%	84.28%	86.85%	86.56%
I-35 N @ Prairie St	83.68%	81.79%	82.75%	85.15%	84.98%
I-35 N @ 95th St*	67.68%	79.79%	82.62%	52.85%	59.75%
I-35 S @ 95th St*	68.91%	81.46%	84.11%	53.57%	60.91%
I-435 S, North of Kansas Ave	72.83%	69.60%	71.72%	75.43%	74.54%
I-435 N, North of Kansas Ave	76.93%	74.56%	76.46%	78.50%	78.20%
I-635 N @ Shawnee Dr	97.97%	97.34%	98.10%	97.81%	98.65%
I-635 S @ Shawnee Dr	97.94%	97.31%	98.05%	97.78%	98.61%

* Excluded data from 11/1/2016 to 11/29/2016.

Table 5.3 shows data quality percentages reported by the KC Scout data portal. These percentages were calculated via aggregation of 5-minute quality completeness factors for each 3-month period.

Table 5.3: KC Scout Data Portal Quality Report Throughout the Year

Station Name	Data Quality				
	11/01/2016–10/31/2017	11/01/2016–1/31/2017	2/01/2017–4/30/2017	5/01/2017–7/31/2017	8/01/2017–10/31/2017
I-70 E @ 72nd St	82.66%	80.93%	83.35%	84.64%	81.74%
I-70 W @ 72nd St	78.93%	80.90%	83.41%	68.50%	83.09%
I-35 N @ Cambridge Cir	83.89%	82.40%	84.42%	84.16%	84.57%
I-35 S @ Cambridge Cir	86.79%	84.73%	87.02%	87.34%	88.06%
I-35 S @ Prairie St	84.50%	82.88%	83.51%	86.23%	85.38%
I-35 N @ Prairie St	82.35%	80.57%	81.39%	84.14%	83.27%
I-35 N @ 95th St*	73.63%	80.16%	83.99%	64.36%	68.34%
I-35 S @ 95th St*	74.03%	81.84%	85.43%	63.30%	68.30%
I-435 S, North of Kansas Ave	73.40%	70.23%	72.33%	75.97%	75.03%
I-435 N, North of Kansas Ave	78.15%	75.80%	77.79%	79.70%	79.26%
I-635 N @ Shawnee Dr	84.13%	82.48%	84.09%	85.24%	84.71%
I-635 S @ Shawnee Dr	80.83%	79.27%	80.60%	82.03%	81.40%

* Excluded data from 11/1/2016 to 11/29/2016.

Comparison of results in Tables 5.2 and 5.3 revealed three trends. First, in the two detector pairs in which manual analysis resulted in data quality higher than 97%, KC Scout reported data quality between 80% and 87%. Second, for the detectors and the 3-month periods with manual analysis resulting in data quality from approximately 70% to 85%, KC Scout quality results were similar, while 3-month periods with data quality between 50% and 60% showed that KC Scout's data quality ranged from 60% to 70%. The differences and similarities in the data quality analysis results between the manual method and the automated KC Scout data portal can be attributed to the reasons (quality check criteria) for low-quality data in each case. Analysis of the temporal distribution of data quality per day and within each day allowed inferences to be made even though percentages that each criterion contributes to low-quality data in the KC Scout reports could not be obtained.

KC Scout reported very similar quality among high-quality detectors (97%) during daily peak hours but only 30–60% quality during off-peak hours, especially during the night, resulting in an overall percentage of 80–87% of good-quality data. These results were attributed to KC Scout's treatment of 30-second intervals with zero vehicle counts. KC Scout flags all intervals with volumes of zero as bad-quality, even if the speed is also zero; therefore, KC Scout flags all time intervals with no vehicles, a low-volume situation that often occurs during late night and early morning hours, resulting in significantly lower reported data quality in the KC Scout database for those hours. Since mere lack of vehicles should not skew the data quality results, it appears that the KC Scout data quality checks process introduces false errors when flagging all zero-volume intervals. Manual analysis, on the other hand, correctly flagged only intervals with zero volume but non-zero speed, which cannot occur, thus indicating hardware or software malfunction.

The difference in flagging zero-vehicle time intervals should cause KC Scout results to always be 10–15% lower than manual analysis results. However, detectors demonstrating medium quality (70–85%) in the manual analysis exhibited similar quality levels in the KC Scout quality reports. Table 5.1 shows that detectors with 15–30% bad-quality records demonstrated approximately 10–15% bad-quality data due to time intervals where the volume was zero but the speed was non-zero and approximately 5–10% bad-quality data due to duplicate records. Distribution of these bad-quality intervals showed that both criteria occur overwhelmingly (98% of the time) during off-peak hours, especially during the night. Duplicate records and zero-volume with non-zero speed intervals had the same distribution as the zero-volume, zero-speed intervals incorrectly flagged by KC Scout during the quality check process of the high-quality (97%) detectors. These errors (duplicate records, zero volume with non-zero speed), though, were accounted for in both the manual analysis and the KC Scout quality checks of the medium-quality (70–85%) detectors. Therefore, peak-period quality results (95% or higher) and off-peak-period quality results (30–60%) for these detectors were consistent between the manual analysis and the KC Scout reports. The exact cause, though, of the errors (duplicate records, zero volume with non-zero speed) during off-peak hours should be investigated since it was not present in all detectors. For example, the high-quality (97%) detectors were free of such errors.

The average 10% higher quality that KC Scout reported for low-quality detectors (50–60%) compared to manual analysis could also be attributed to the cause of the bad-quality data. During the 3-month periods in which the detectors exhibited low-quality, duplicate values and zero volume, non-zero speed errors during off-peak hours were significant; however, the high-volume error was also prevalent. This error, unlike the zero volume, non-zero speed error, did not occur during specific hours in each day but was observed throughout the entire day for a period of many consecutive days, longer than a month in most cases. During these days, the manual analysis resulted in data quality near 0%, but the KC Scout data portal quality varied from 0% to 30%. The assumption was made that the threshold for extremely high volume in KC Scout may be higher than the threshold used for the manual analysis. However, high-volume errors in one or more lanes had a very high tendency (close to 100%) to coincide with duplicate records in all lanes of each 30-second time interval, indicating that these two error types may have a common triggering factor, a malfunction that causes both simultaneously. In addition, KC Scout inconsistently identified duplicate records, especially records that also occurred during high-volume intervals. Although KC Scout correctly flagged most duplicate records as low quality (where high-volume intervals also occurred for some lanes), it failed to flag about one-sixth of them and mistakenly reported good or excellent quality instead. Because no data pattern explained this inconsistency, the duplicate records quality check procedure of the data portal software should be examined for errors. Whether varying high-volume error thresholds, KC Scout's failure to identify a portion of duplicate record errors, or both causes were to blame, this difference in flagging the high-volume and duplicate-values records during high-volume days was the reason for the average 10% higher quality reported by KC Scout.

In conclusion, comparison of the data quality results of the manual analysis and the data quality checks performed by KC Scout during data aggregation showed that the results were generally consistent for the peak periods and off-peak periods when duplicate records and zero-volume, non-zero speed errors occurred. However, unlike the manual analysis, KC Scout flags zero-volume, zero-speed intervals, has a higher threshold for very-high-volume errors, and detects only a percentage of duplicate record errors.

5.2 Reliability Assessment During Field Data Collection

Data quality analysis was applied to the 6-hour and 3-hour time periods of the first and second data collection phases (Table 5.4). With the exception of one detector (I-70 W @ 72nd St) during Phase I, the quality of data for all detectors during both phases was much higher than the 1-year average. Missing intervals were only observed in the I-70 W @ 72nd St detector during Phase I, while for all the other detectors, bad-quality records were exclusively (100%) due to consecutive duplicate values and zero volume with non-zero speed. The low-quality detector (2.5% data quality) had 30 missing intervals (1.4% of the total errors); approximately half of the errors were due to consecutive duplicate values, and the other half were due to extremely high volumes. Data collection for this detector during Phase I occurred under heavy rain conditions.

Data quality reported by KC Scout for the corresponding time periods is shown in Table 5.5. As expected, since field data collection generally occurred during time periods with average-to-high volumes, the differences between the two methods are minimal in most cases. For the detector with only 2.5% quality in the manual analysis, KC Scout reported a higher percentage (10.8%) due to a higher threshold value for high-volume errors and failure to correctly identify one-sixth of the duplicate records.

Table 5.4: Field Data Quality Analysis

Station Name	Data Quality	
	Phase I (6 hours)	Phase II (3 hours)
I-70 E @ 72nd St	96.81%	85.28%
I-70 W @ 72nd St	2.50%	85.28%
I-35 N @ Cambridge Cir	99.72%	99.54%
I-35 S @ Cambridge Cir	99.86%	99.35%
I-35 S @ Prairie St	99.03%	98.61%
I-35 N @ Prairie St	99.26%	98.89%
I-35 N @ 95th St	94.07%	97.69%
I-35 S @ 95th St	94.35%	98.80%
I-435 S, North of Kansas Ave	85.83%	89.07%
I-435 N, North of Kansas Ave	93.61%	98.47%
I-635 N @ Shawnee Dr	98.94%	98.43%
I-635 S @ Shawnee Dr	98.70%	97.78%

Table 5.5: KC Scout Field Data Quality Report

Station Name	Data Quality	
	Phase I (6 hours)	Phase II (3 hours)
I-70 E @ 72nd St	99.31%	85.83%
I-70 W @ 72nd St	10.83%	86.67%
I-35 N @ Cambridge Cir	100.00%	100.00%
I-35 S @ Cambridge Cir	100.00%	99.72%
I-35 S @ Prairie St	99.86%	100.00%
I-35 N @ Prairie St	99.86%	100.00%
I-35 N @ 95th St	96.11%	99.44%
I-35 S @ 95th St	96.39%	99.72%
I-435 S, North of Kansas Ave	86.11%	90.83%
I-435 N, North of Kansas Ave	94.58%	99.72%
I-635 N @ Shawnee Dr	99.58%	94.17%
I-635 S @ Shawnee Dr	97.50%	91.67%

Chapter 6: Conclusions and Recommendations

An analysis of previous studies showed that KC Scout's radar sensors can provide overall reliable traffic volumes; however, because the primary application of this data is for travel time/speed measurements, caution is recommended for volume-related data applications. This research project collected field data from 12 sensor sites in two multi-hour phases and corroborated these findings, showing that, when calibrated and checked by KC Scout personnel, volume data accuracy was quite high. In fact, the average difference between manual counts and raw sensor counts was less than 2.2% for all sites during Phase II, which is an adequate result for KDOT specifications. However, a long-term data quality analysis revealed complications either with sensor performance or the automated data quality checks and aggregation processes of the KC Scout data portal.

This research analyzed raw sensor data from KC Scout to investigate missing or imputed values. Data quality checks for 1 year revealed two pairs of detectors with excellent 1-year quality (approximately 98% good-quality data), while qualities of the remaining four pairs ranged from 85% to 60%. The most frequent source of error was the presence of duplicate records. Other reasons for low quality included zero traffic counts with speed of non-zero, extremely high vehicle counts, or missing intervals. An analysis of the temporal distribution of the errors showed the following:

- Missing intervals were distributed randomly throughout the year and within each day.
  - For all but one pair of sensors, missing intervals represented a small percentage (0.5–1.5%) of the daily errors.
  - One pair of sensors was offline for 29 days (8% of the year) in November 2016 due to construction work at the sensor location.
- High-volume records occurred in continuous blocks of several days, indicating a malfunction of the sensor or data portal software.

- Duplicate records and records of zero traffic counts with non-zero speeds occurred almost exclusively during off-peak hours, especially late at night and early in the morning.
- Duplicate records were also present during peak hours if high-volume records also occurred.

A comparison of these results with the automated data quality checks performed by KC Scout during data aggregation showed differences stemming from the use of diverse threshold values or different definitions for error criteria. KC Scout used a higher threshold for extremely high volumes and failed to identify a percentage of duplicate records. No pattern was identified to explain why some duplicate records were not registered as errors by the automated quality checks. However, KC Scout flagged all zero-volume records for bad quality, regardless of whether the speed was zero. Due to the temporal distribution of the errors, data quality results of the manual analysis and data quality checks performed by KC Scout during data aggregation were only consistent with each other for most of the peak periods and for the off-peak periods when duplicate records and zero-volume, non-zero speed errors occurred. KC Scout data quality was up to 30% higher than manual analysis data quality during days with extremely high recorded volumes, and 30–70% lower during the night when zero-volume, zero-speed intervals were frequent.

Data quality analysis during field data collection showed that, in all but one case, data quality was above 85% (Table 5.3). In 20 of the 24 cases, quality was over 93%. One case during Phase I exhibited very low data quality (2.5%), indicating a malfunctioning sensor. In Phase II of the data collection, which took place because four radar sensors were configured incorrectly during Phase I (e.g., missing lanes, wrong lanes assigned, etc.), nine out of 12 cases exhibited quality higher than 97.5%, and all 12 cases had a quality above 85%.

Missing data records, as well as data records determined to be of low quality by the automated quality checks of the KC Scout data portal, were imputed during the data aggregation process. If upstream and downstream data records of good quality were available, interpolation was performed; otherwise, a 2-year historical average was used. If a historical average was not available, the imputed values were entered manually. The data portal did not provide the percentage of imputed values that were interpolation results, historical averages, or manual entries; only the total number of imputed values, equal to bad-quality data records, was known.

The errors and inconsistencies of the KC Scout data portal quality checks process, including a potentially poorly calibrated error threshold parameter, resulted in the imputation of valid data records and the validation and usage of data records with errors. Imputation of valid data records caused low-volume, off-peak hours to exhibit reported volumes 60% higher, on average, and up to 120% higher (or 100 to 250 more vehicles per hour) than actual traffic counts during those hours. Reported volumes should be discarded as invalid when sensor malfunction causes a combination of very high traffic counts and duplicate records.

The findings of this report suggest that under certain conditions, KC Scout sensors can record traffic counts of acceptable accuracy and quality. However, long-term reliability is inconsistent among the sensors and data quality within the year and within each day varied significantly in some cases, requiring sensors to be closely monitored for errors. In addition, errors found during data quality checks of the KC Scout data portal rendered it unreliable for use in its current state. If KDOT wishes to use traffic count data from the KC Scout data portal, the following steps are recommended:

- KDOT should monitor subject sensors for any maintenance or construction activity that could potentially decrease sensor accuracy and reliability.
- KDOT should communicate directly with KC Scout to verify that subject sensors are operating as expected and are configured properly.
- Year-long and daily data quality reliability of the subject sensors must be taken into account. This study showed that both year-long and daily data quality of the sensors varies. Some sensors exhibited exceptional quality for the entire year and during each day, some produced errors during low-volume, off-peak intervals every day of the year, and a few detectors had month-long malfunctioning periods. KC Scout should identify the causes of these errors and perform any maintenance necessary to prevent them or reduce their frequency.
- Since the exact sources of error are not available from the KC Scout data portal, the malfunctioning detectors can be identified by the data quality

check process used in this study. After downloading raw data from the KC Scout website (a step-by-step guide is shown in Appendix A), KDOT can perform quality checks to estimate the percentage of missing or erroneous records, including the specific sources of error, for each sensor. An Excel spreadsheet with an example calculation of data quality is available separately upon request.

- In addition, KC Scout should review and revise the automated data quality checks process of the data portal according to KDOT specifications to eliminate the errors and inconsistencies identified in this study, including unnecessary flagging of zero-volume, zero-speed intervals, failure to flag duplicate values, and error threshold values that may not be appropriately calibrated. If these revisions are performed, the automated data quality checks process of the data portal could replace the manual quality check process.

Final recommendations of this study include the following:

- If the percentage of good-quality records is acceptable in both the manual and the automated quality check process, KDOT can use aggregate volume data from KC Scout at the specific sensor for the specific time interval.
- If the percentage of good-quality records is acceptable in the manual process but not acceptable in the automated process, then KDOT should use raw traffic counts at the specific sensor for the specific time interval.
- If the percentage of good-quality records is not acceptable in the manual process, KDOT cannot obtain useful volume data at the specific sensor for the specific time interval.
- If the automated process is revised to remove the current errors and inconsistencies, the manual process is no longer required, and the percentage of good-quality records in the manual process is a sufficient indicator of data quality and whether the aggregate volume data can be used by KDOT for reporting purposes.

References

Coifman, B. (2005). Freeway detector assessment: Aggregate data from remote traffic microwave sensor. *Transportation Research Record, 1917*, 149–163.

Coifman, B. (2006). Vehicle level evaluation of loop detectors and the remote traffic microwave sensor. *Journal of Transportation Engineering, 132*(3), 213–226.

Elefteriadou, L., Kondyli, A., Washburn, S. S., Brilon, W., Lohoff, J., Jacobson, L.… Persaud, B. (2011). Proactive ramp management under the threat of freeway-flow breakdown. *Procedia Social and Behavioral Sciences, 16*, 4–14.

Federal Highway Administration (FHWA). (2016a). *Highway performance monitoring system field manual*. Washington, D.C.: Author.

Federal Highway Administration (FHWA). (2016b). *Traffic monitoring guide*. Washington, D.C.: Author.

Federal Highway Administration (FHWA). (2017). *Traffic volume trends: Frequently asked questions*. Retrieved from https://www.fhwa.dot.gov/ohim/tvtw/tvtfaq.cfm

Grone, B. W. (2012). *An evaluation of non-intrusive traffic detectors at the NTC/NDOR detector test bed* (Master's thesis). University of Nebraska-Lincoln, Lincoln, NE.

Jacobson, L. N., Nihan, N. L., & Bender, J. D. (1990). Detecting erroneous loop detector data in a freeway traffic management system. *Transportation Research Record, 1287*, 151–166.

Kansas Department of Transportation (KDOT). (2017, October). *KDOT facts and trends reports*. Retrieved from https://www.ksdot.org/bureaus/burtransplan/prodinfo/factsandtrends.asp

Klein, L. A., & Kelley, M. R. (1996). *Detection technology for IVHS Volume I: Final report* (Report No. FHWA-RD-95-100). McLean, VA: Federal Highway Administration.

Kotzenmacher, J., Minge, E., & Hao, B. (2005). *Evaluation of portable non-intrusive traffic detection system* (Report No. MN-RC-2005-37). St. Paul, MN: Minnesota Department of Transportation.

Mimbela, L. E. Y., & Klein, L. A. (2007). *A summary of vehicle detection and surveillance technologies used in intelligent transportation systems*. Washington, D.C.: Federal Highway Administration (FHWA).

Minge, E., Kotzenmacher, J., & Peterson, S. (2010). *Evaluation of non-intrusive technologies for traffic detection* (Report No. MN/RC 2010-36). St. Paul, MN: Minnesota Department of Transportation.

Nihan, N., Jacobson, L. N., Bender, J. D., & Davis, G. (1990). *Detector data validity* (Report No. WA-RD 208.1). Olympia, WA: Washington State Department of Transportation.

TransCore ITS, LLC. (2016). *TransSuite® Web-based Data Portal user manual* (Version 4.7). Retrieved from http://kcscout.net/KcDataPortal/docs/Data%20Portal%20User%20Manual-v4.7.pdf

Turner, S., Margiotta, R., & Lomax, T. (2004). *Monitoring urban freeways in 2003: Current conditions and trends from archived operations data* (Report No. FHWA-HOP-05-018). Washington, D.C.: Federal Highway Administration.

Zhang, L. (2006). *Comparison of non-intrusive traffic detection devices for freeway application* (Master's thesis). University of Nebraska-Lincoln, Lincoln, Nebraska.

Appendix A: Using the KC Scout Data Portal

This section describes the steps to extract monthly traffic data from KC Scout's database for submitting TVT report (TransCore ITS, 2016).

A.1 User Authentication and Navigation Interface

The KC Scout data portal is only accessible to registered users. Therefore, KC Scout administrators must create a user account for KDOT if one does not already exist. Once the account has been activated, the user should visit the data portal's webpage (http://kcscout.net/KcDataPortal/) and log in with credentials (user name and password).

The *Home* screen is displayed after user authentication is successfully completed. Depending on the role (or level of access) assigned to the user, various pages/areas are available from the navigation interface. Figure A.1 displays the *Home* screen for users with basic privileges, where the top navigation bar only includes *Home* and *Logout* buttons. System administrators and users with additional privileges can also access the *Data Edit* and *Administration* modules from the top navigation bar. The left navigation flow allows the user to browse the sub-modules of the data portal. In this guide only the first sub-module, *Detector Stations*, is relevant to the described process.

Figure A.1: Data Portal Home Screen

A.2 Detector Report Queries

The *Home* screen shown in Figure A.1 also serves as the *View Detector Reports* screen, which lists all existing detector report queries of the user. This screen also allows a user to search for a specific query or to edit, copy, delete, run, or create a new query. The following steps are applicable for creating, running, and downloading a detector report query:

1. Click the *Home* button or the *View Reports* menu item under the *Detector Stations* section to return to that screen if on a different screen.
2. Click the Add New button to navigate to the *Edit Detector Report* screen (Figure A.2).

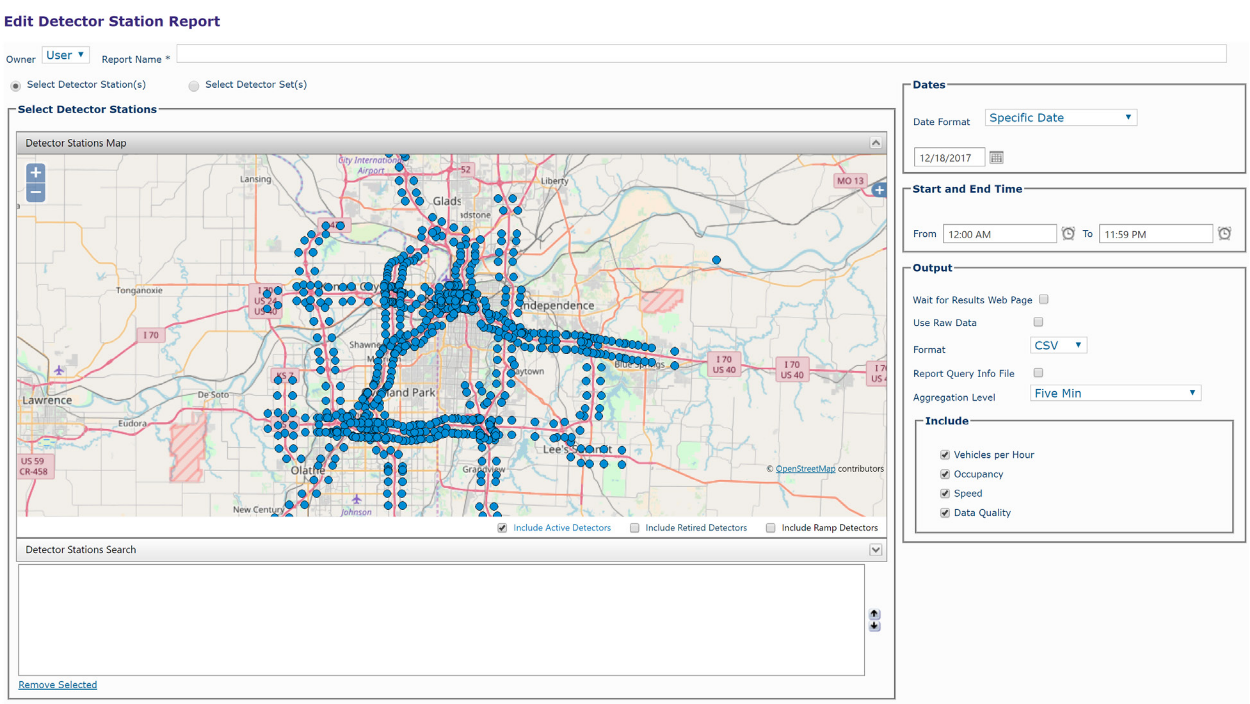
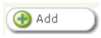


Figure A.2: Edit Detector Report Screen

3. Enter a name for the query in the *Query Name* input field.
4. Add detector stations of interest to the query by directly clicking on the detector station on the map under the *Select Detector Station(s)* option. The selected detectors will show up on the list box at the bottom of the screen. Alternatively, entire detector sets can be added at once using the *Select Detector Set(s)* option.

Existing detector station sets can be selected from the *Set* drop-down list in the *Detector Stations* search panel and added to the *Detector Stations* list box by clicking on the 

button.

5. Select *Date Range* from the *Date Format* drop-down list. Then enter a start and end day for the query in the *Start* and *End* input fields or by clicking on the calendar pop-up button. Check the box or boxes for the days of the week to include in the query, and check or uncheck the box to identify if holidays should be excluded from the query. According to FHWA reporting requirements, the date range should be the entire month being reported, including all days of the week and holidays.
6. Select a start and end time in the *From* and *To* input fields. The default values of 12:00 a.m. to 11:59 p.m. are sufficient for reporting requirements.
7. Select an aggregation level of 1 hour for the query from the *Aggregation Level* drop-down list. Other aggregation levels are available if needed. If the *Use Raw Data* box is selected, the 30-second raw data are retrieved instead of the aggregated data. This step only applies to the manual data quality procedure and not the traffic trends report.
8. Select the type of output file from the *Format* drop-down list.
9. Click the 

button to save the changes and to create the new query. The message

should be displayed at the bottom of the screen.
10. Click the 

button to submit the query for execution. By clicking the *View Report Status* menu item under the *Detector Stations* section, the user can see the status of the query and whether it is ready for download.
11. To download the report the user must click the *Retrieve Results* menu item under the *Detector Stations* section and then click on the download button ().
12. The downloaded report file contains hourly traffic volumes for all selected detectors for the chosen time period. Figure A.3 shows an example of the output file, with November 1 data for the I-435 northbound sensor used in the accuracy

analysis. Column I, which lists vehicle counts during each time interval (1 hour), is a simple sum of the raw, unadjusted, 30-second data counts for each respective hour, regardless of data quality. Column J is the volume per hour calculated by summing the adjusted aggregate values. Separate results for each lane are also available: Columns U, AG, AS, BE, and BQ are the raw vehicle counts for lanes 1, 2, 3, 4, and 5, respectively, while Columns V, AH, AT, BF, and BR are the adjusted aggregate volumes. Section 2.1 describes how data points flagged for low quality are adjusted via interpolation using historical data or manual adjustment. Volume M is the volume data quality percentage according to KC Scout criteria, rounded to the nearest multiple of 10. The hourly counts and vehicles per hour (VPH) columns have the same values only if volume data quality is 100%. Therefore, KDOT can use either column when KC Scout data quality is high; when data quality is low, however, the values of the counts column become more unreliable, and values of the VPH column are based more on assumptions (interpolations or historical data) than on field data. A further complication is caused by the observation (Section 5.1) that current KC Scout criteria tend to falsely attribute low quality to hours with sparse traffic (primarily in the middle of the night) because these criteria interpret all intervals with 0 vehicles as bad-quality records, even if speed is also 0, indicating that no vehicle was present during that time. This causes the data quality during those hours to be reported as lower (20–50%, Figure A.3) than it should be, and the values of the VPH column to fail to represent the actual ground truth data contained in the Counts column. To solve this complication, KDOT could work with KC Scout to define quality check thresholds according to KDOT specifications and eliminate bad-quality flags for intervals with 0 vehicles if the speed is also 0, or KDOT could determine data quality directly from the raw data based on the procedure described in Chapter 5 and in the following subsection of the Appendix (see the example spreadsheet, available separately upon request to KDOT#Research.Library@ks.gov) instead of using the quality reported by KC Scout. If the first method is followed, then the values in Column J (VPH) should be used; if the second method is used, Column I (Counts) should be used when quality is high and Volume J when quality is low.

	A	B	C	D	E	F	G	H	I	J	K
1	IntId	Station	State	Location	Timestamp	Interval	Lanes	Dir	Cnt	VPH	Occ
2	8308	I-435 N @ North of Kansas Ave		Edwardsville KS	11/1/2017 0:00	Hour	5	N	236	485	0.812
3	8308	I-435 N @ North of Kansas Ave		Edwardsville KS	11/1/2017 1:00	Hour	5	N	162	443	0.591
4	8308	I-435 N @ North of Kansas Ave		Edwardsville KS	11/1/2017 2:00	Hour	5	N	168	479	0.725
5	8308	I-435 N @ North of Kansas Ave		Edwardsville KS	11/1/2017 3:00	Hour	5	N	179	463	0.734
6	8308	I-435 N @ North of Kansas Ave		Edwardsville KS	11/1/2017 4:00	Hour	5	N	373	574	1.092
7	8308	I-435 N @ North of Kansas Ave		Edwardsville KS	11/1/2017 5:00	Hour	5	N	788	954	1.929
8	8308	I-435 N @ North of Kansas Ave		Edwardsville KS	11/1/2017 6:00	Hour	5	N	1708	1756	3.785
9	8308	I-435 N @ North of Kansas Ave		Edwardsville KS	11/1/2017 7:00	Hour	5	N	3227	3250	6.884
10	8308	I-435 N @ North of Kansas Ave		Edwardsville KS	11/1/2017 8:00	Hour	5	N	3099	3134	7.015
11	8308	I-435 N @ North of Kansas Ave		Edwardsville KS	11/1/2017 9:00	Hour	5	N	1850	1865	4.554
12	8308	I-435 N @ North of Kansas Ave		Edwardsville KS	11/1/2017 10:00	Hour	5	N	1551	1601	3.802
13	8308	I-435 N @ North of Kansas Ave		Edwardsville KS	11/1/2017 11:00	Hour	5	N	1636	1668	4.034
14	8308	I-435 N @ North of Kansas Ave		Edwardsville KS	11/1/2017 12:00	Hour	5	N	1727	1769	4.077
15	8308	I-435 N @ North of Kansas Ave		Edwardsville KS	11/1/2017 13:00	Hour	5	N	1679	1720	4.103
16	8308	I-435 N @ North of Kansas Ave		Edwardsville KS	11/1/2017 14:00	Hour	5	N	1889	1914	4.341
17	8308	I-435 N @ North of Kansas Ave		Edwardsville KS	11/1/2017 15:00	Hour	5	N	2676	2740	5.802
18	8308	I-435 N @ North of Kansas Ave		Edwardsville KS	11/1/2017 16:00	Hour	5	N	3813	3815	7.85
19	8308	I-435 N @ North of Kansas Ave		Edwardsville KS	11/1/2017 17:00	Hour	5	N	4273	4241	8.564
20	8308	I-435 N @ North of Kansas Ave		Edwardsville KS	11/1/2017 18:00	Hour	5	N	2320	2344	4.817
21	8308	I-435 N @ North of Kansas Ave		Edwardsville KS	11/1/2017 19:00	Hour	5	N	1289	1342	2.92
22	8308	I-435 N @ North of Kansas Ave		Edwardsville KS	11/1/2017 20:00	Hour	5	N	1041	1144	2.524
23	8308	I-435 N @ North of Kansas Ave		Edwardsville KS	11/1/2017 21:00	Hour	5	N	793	908	1.925
24	8308	I-435 N @ North of Kansas Ave		Edwardsville KS	11/1/2017 22:00	Hour	5	N	486	679	1.274
25	8308	I-435 N @ North of Kansas Ave		Edwardsville KS	11/1/2017 23:00	Hour	5	N	372	580	1.07

	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB
1	Spd	VQ	SQ	OQ	VC1 Cnt	VC2 Cnt	VC3 Cnt	VC4 Cnt	StationIntId1	Cnt1	VPH1	Occ1	Spd1	VQ1	SQ1	OQ1	VC1 Cnt1
2	74	40	90	30	1	131	36	59	8308	25	124	0.15	77	10	90	10	0
3	74	30	90	30	1	83	32	40	8308	19	129	0.041	81	0	80	0	0
4	72	20	100	20	2	60	39	59	8308	21	134	0.075	76	0	90	0	0
5	72	30	100	30	0	78	43	55	8308	17	123	0.125	74	10	100	10	0
6	76	50	100	50	1	233	84	54	8308	37	133	0.258	79	20	100	20	0
7	77	70	100	70	3	531	184	70	8308	122	214	0.65	81	50	100	40	0
8	78	90	100	90	4	1238	361	105	8308	331	369	1.675	83	80	90	80	0
9	77	90	90	90	10	2495	603	119	8308	715	723	3.616	82	90	90	90	0
10	77	90	90	90	10	2274	624	191	8308	692	705	3.608	82	90	90	90	0
11	77	100	90	100	3	1237	446	164	8308	379	394	2.016	82	90	90	90	0
12	77	90	90	90	2	1053	360	136	8308	260	290	1.433	81	90	90	80	0
13	77	90	100	90	1	1113	372	150	8308	277	304	1.508	81	90	90	80	0
14	77	90	100	90	3	1190	408	126	8308	293	321	1.591	81	90	90	80	0
15	76	90	100	90	3	1143	387	146	8308	270	311	1.433	80	80	100	80	0
16	77	90	100	90	4	1328	432	125	8308	373	389	2	81	90	90	90	0
17	76	90	90	90	10	1983	557	126	8308	662	696	3.391	81	90	90	90	0
18	76	90	90	90	28	2984	674	127	8308	1160	1154	5.958	81	90	90	90	0
19	75	90	90	90	26	3476	680	91	8308	1416	1367	7.283	80	90	90	90	1
20	77	90	100	90	6	1845	383	86	8308	536	557	2.658	82	90	100	90	0
21	76	90	100	90	2	1014	197	76	8308	246	282	1.275	80	80	100	80	0
22	75	80	100	80	4	799	156	82	8308	205	271	1.125	79	70	100	70	0
23	76	80	100	80	2	599	128	64	8308	119	184	0.666	80	60	100	50	0
24	74	60	90	60	1	356	81	48	8308	54	141	0.416	77	30	90	30	0
25	73	50	100	50	0	268	54	50	8308	44	147	0.266	77	20	90	20	0

Figure A.3: Report File Example

A.3 Data Quality Assessment

This subsection describes the steps to determine volume data quality based on raw data from KC Scout and the data quality example spreadsheet (available separately upon request).

1. Use the instructions in Section A.2 to create and download a raw data detector report for each detector and time period of interest. Raw data are retrieved by selecting the *Use Raw Data* check box (Step 7 in Section A.2).
2. Open the downloaded report file and the data quality example spreadsheet and copy the entire report spreadsheet into the first tab (*RAW*) of the data quality example spreadsheet (or copy only the lines of data corresponding to the time period of interest).
3. Calculate the number of minutes that correspond to the data in the *RAW* tab. For example, if the *RAW* tab contains data for all days of the month of November, the number of minutes should be 30 days * 24 hours * 60 minutes = 43,200 minutes. Or *RAW* tab data for one day should be 1,440 minutes. Actual lines of data in the *RAW* tab will probably be less than the calculated number due to missing intervals.

4. Match the number of lines above the summary statistics in the *Quality Checks* tab with twice the number of minutes calculated in Step 3 by adding or subtracting lines in the *Quality Checks* tab accordingly (Figure A.4). The example spreadsheet contains lines for a 3-month period, so lines typically need to be subtracted from that. (Columns B–E are an alternative for calculating the number of 30-second intervals.)
5. Find the summary statistics on data quality, categorized by type of error, in the *Summary* tab (Figure A.5).

	A	B	C	D	E	F	G	H	I	J	K	L	M
1		First Quarter	Second Quarter	Third Quarter	Fourth Quarter	TimeStamp		FLAGS	ALL LANES	"Interval"	"Status"	"RawStatus"	Elapsed Time
264947	264946	1/31/2017 23:52		7/31/2017 23:52	10/31/2017 23:52	1/0/1900 0:00	0	3	0	0	0	0	1
264948	264947	1/31/2017 23:53		7/31/2017 23:53	10/31/2017 23:53	1/0/1900 0:00	0	3	0	0	0	0	1
264949	264948	1/31/2017 23:53		7/31/2017 23:53	10/31/2017 23:53	1/0/1900 0:00	0	3	0	0	0	0	1
264950	264949	1/31/2017 23:54		7/31/2017 23:54	10/31/2017 23:54	1/0/1900 0:00	0	3	0	0	0	0	1
264951	264950	1/31/2017 23:54		7/31/2017 23:54	10/31/2017 23:54	1/0/1900 0:00	0	3	0	0	0	0	1
264952	264951	1/31/2017 23:55		7/31/2017 23:55	10/31/2017 23:55	1/0/1900 0:00	0	3	0	0	0	0	1
264953	264952	1/31/2017 23:55		7/31/2017 23:55	10/31/2017 23:55	1/0/1900 0:00	0	3	0	0	0	0	1
264954	264953	1/31/2017 23:56		7/31/2017 23:56	10/31/2017 23:56	1/0/1900 0:00	0	3	0	0	0	0	1
264955	264954	1/31/2017 23:56		7/31/2017 23:56	10/31/2017 23:56	1/0/1900 0:00	0	3	0	0	0	0	1
264956	264955	1/31/2017 23:57		7/31/2017 23:57	10/31/2017 23:57	1/0/1900 0:00	0	3	0	0	0	0	1
264957	264956	1/31/2017 23:57		7/31/2017 23:57	10/31/2017 23:57	1/0/1900 0:00	0	3	0	0	0	0	1
264958	264957	1/31/2017 23:58		7/31/2017 23:58	10/31/2017 23:58	1/0/1900 0:00	0	3	0	0	0	0	1
264959	264958	1/31/2017 23:58		7/31/2017 23:58	10/31/2017 23:58	1/0/1900 0:00	0	3	0	0	0	0	1
264960	264959	1/31/2017 23:59		7/31/2017 23:59	10/31/2017 23:59	1/0/1900 0:00	0	3	0	0	0	0	1
264961	264960	1/31/2017 23:59		7/31/2017 23:59	10/31/2017 23:59	1/0/1900 0:00	0	3	0	0	0	0	1
264962	264961					EXISTING	**89062**	**532551**	**176003**	**175898**	**175898**	**175898**	**105**
264963	264962					TOTAL	264960						
264964	264963					MISSING	**175898**						
264965	264964						FINAL:	**532551**	**105**	**0**	**0**	**0**	**105**
264966	264965												

Figure A.4: *Quality Checks* Tab in the Data Quality Example Spreadsheet, Including Summary Statistics Lines

	A	B	C	D
1				
2		**Overall Records**	794880	
3		**Bad Records (Flags)**	532551	
4		**Good Quality Data %**	33.00%	
5				
6				
7		**By Category**	**Bad Records (Flags)**	% By Category
8		System Errors	0	0.00%
9		Lane Status Errors	0	0.00%
10		Elapsed Time Errors	315	0.06%
11		Consecutive Duplicate Values	2659	0.50%
12	SUM	All System Errors	2974	0.56%
13				
14	SUM	Missing Intervals	527694	99.07%
15				
16		High Volume	0	0.00%
17		High Speed	730	0.14%
18		High Occupancy	0	0.00%
19		0 veh, >0 speed	343	0.06%
20		0 speed, >0 veh	902	0.17%
21		High Density	2	0.00%
22	SUM	All Measurement Errors	1977	0.37%

Figure A.5: *Summary* Tab in the Data Quality Example Spreadsheet

K-TRAN

KANSAS TRANSPORTATION RESEARCH AND NEW-DEVELOPMENT PROGRAM

AD ASTRA PER ASPERA

Kansas

Department of Transportation