**Highway Incident Detection Timeline**

**Work Order (WO) 009**

**Contract No. 4400011166**

**Report on Task 3-Deliverable 3.1:**

**Analysis of 911 Call Data and PennDOT Highway Road Closure Data**

**By:**

**Temple University Research Team**

Principal Investigator: Joseph Thomas Coe, Jr., Ph.D.

Assistant Professor

Co- Principal Investigator: Bechara E. Abboud, Ph.D., P.E.

Associate Professor

Co- Principal Investigator: Joseph Picone, Ph.D.

Professor

Graduate Research Assistant: Siavash Mahvelati

Undergraduate Research Assistant: Aaron M. Gross

**07/14/2017**

Contents

[1 Task 3: Analysis of 911 Call Data and PennDOT Highway Road Closure Data 3](#_Toc487839912)

[1.1 Introduction 3](#_Toc487839913)

[1.2 Summary of Data Acquisition & Pre-Processing Efforts 3](#_Toc487839914)

[1.3 Integrated Framework & Graphical User Interface (GUI) 6](#_Toc487839915)

[1.4 Data Analysis and Discussion 11](#_Toc487839916)

[1.4.1 Distribution of Data and Measure of Central Tendency 11](#_Toc487839917)

[1.4.2 Geospatial Distribution of Results 17](#_Toc487839918)

[1.4.3 Match Rate 17](#_Toc487839919)

[1.5 Future Efforts 18](#_Toc487839920)

[Appendix A: Histograms and Cumulative Distribution Plots By County A-1](#_Toc487839921)

# Task 3: Analysis of 911 Call Data and PennDOT Highway Road Closure Data

## Introduction

The purpose of this task is to evaluate the timeline for incidents requiring highway closures along Primary Interstate Highway I-76, I-80, I-81, and I-95, and Auxiliary Interstate Highway I-78 and I-83 throughout the Commonwealth of Pennsylvania. This task addresses the general objective of TEM Work Order (WO) 009, which is to estimate the latency between when the Pennsylvania Department of Transportation (PennDOT) is notified of incidents along the aforementioned highways relative to when 911 call centers receive notification. Previous tasks involved acquiring data from 911 call centers throughout the Commonwealth of Pennsylvania and from the PennDOT Road Condition Reporting System (RCRS). The data was normalized and post-processed so that an integrated framework could be developed that allowed matching data between the two databases and statistical analysis of the matches to evaluate time difference. The following sections present a brief review of data acquisition and post-processing, followed by an in-depth discussion of efforts involved to develop statistical estimates of notification latency and to analyze the results.

## Summary of Data Acquisition & Pre-Processing Efforts

As noted in previous task deliverable reports, significant efforts were involved in acquiring data from 911 call centers, pre-processing the resulting data, and in developing an integrated framework for pairing the data from the 911 call center and PennDOT RCRS databases. PennDOT provided electronic copies of the RCRS logs for the highways of interest early on within the project timeline (Tuesday, November 22, 2016), so data collection efforts for this project primarily focused on contacting the 911 call centers. Based on the original RCRS data provided by PennDOT, a total of 20,950 entries occurred along Primary Interstate Highway I-76, I-80, I-81, and I-95, and Auxiliary Interstate Highway I-78 and I-83 between the dates of 01/01/2013 and 11/22/2016. These entries corresponded to multiple event “types”, including Lane Restriction, Shoulder Closed, Traffic Disruption, Ramp Restriction, Ramp Closure, Closed, Open, and Residual Delays. Given this initial database from PennDOT and a review of the routes of the project highways, a total of 37 counties (of the 67 counties throughout Pennsylvania) were identified as pertinent to this project (Table 1). The Temple project team contacted these counties based on direct emails/phone calls to contacts provided by PennDOT as well as submission of right-to-know (RTK) requests directly with the counties. PennDOT had also already reached out to the aforementioned contacts and advised them of Temple’s involvement in this project. Based on conversations between the Temple research team and the PennDOT project team at an in-person meeting (Friday, February 3rd, 2017), filter criteria was established to focus on the RCRS incident types most important to PennDOT (i.e., entries with status of “Closed”, “Lane Restriction”, “Ramp Closure”, and “Ramp Restriction”). This changed the total number of pertinent PennDOT RCRS entries from 20,950 entries to 9,009 entries and reduced the number of pertinent counties to 30 (Table 1). Efforts have continued since that time to acquire as much data as possible for the statistical analysis in this project. In some cases (Philadelphia, Delaware, Columbia, Franklin, and Monroe counties), this has included interfacing with contacts at the Pennsylvania State Police (PSP) and submitting RTK requests through them since their agency handles 911 calls regarding highways in the aforementioned counties.

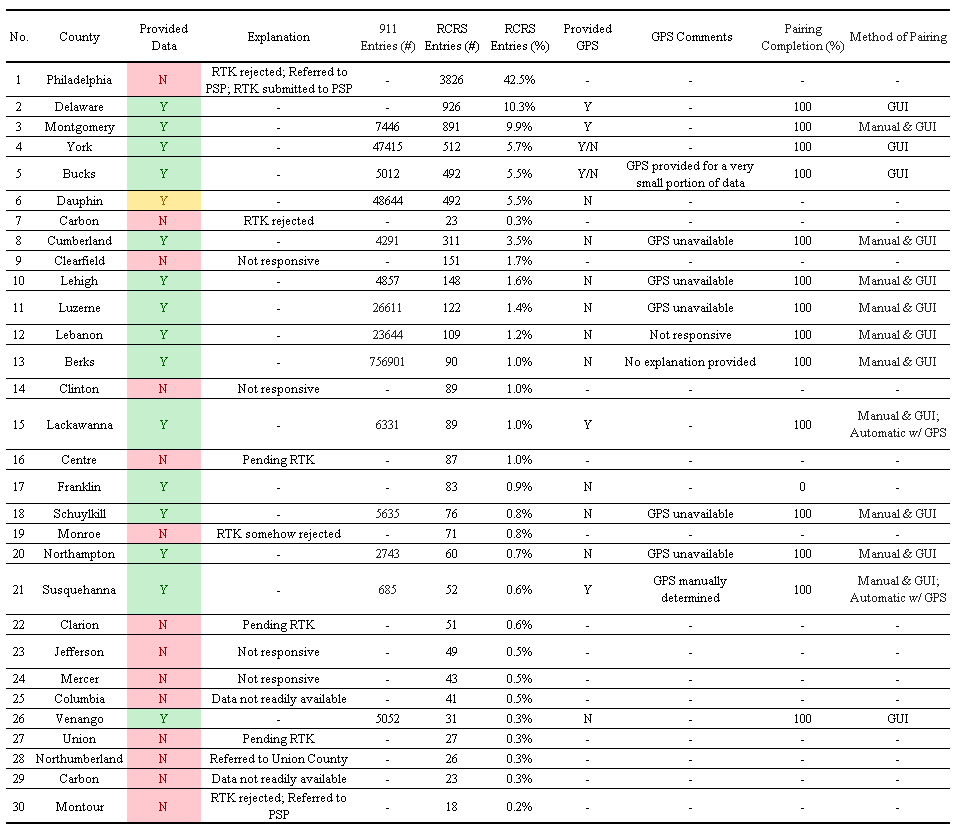


Table . Current status for all counties.

As of the writing of this deliverable report, the Temple research team has acquired pertinent 911 call data from 15 counties out of the 30 counties with entries in the filtered RCRS database (Table 1). As a whole, these counties combined to provide a total of 945,267 entries in the 911 call database over a time period ranging from the start of 2013 to late 2016. Based on pairing of the data (as will be discussed in later sections of this report), this acquired 911 data represented 43.4% of the total number of filtered RCRS entries (i.e., 3,909 of 9,009). The bulk of the remaining data is composed of 911 calls from Philadelphia county (42.5% of the RCRS entries), which is incapable of directly providing the data because highway-related calls are dispatched by PSP in that county. Efforts to acquire the data from PSP have been unsuccessful thus far and an RTK request is pending in addition to direct communication with the local Troop K within Philadelphia County. Apart from Philadelphia County, the Temple research team was able to acquire 911 data from what amounts to 75.4% of the pertinent RCRS entries (i.e., 3,909 out of 5,183, excluding RCRS entries from Philadelphia County). Philadelphia County has a high population density relative to other Pennsylvania counties and its urban setting provides a more extensive traffic camera network. However, nearby counties where the Temple research team was able to acquire data (e.g., Delaware and Montgomery) share some similarities to Philadelphia County regarding population and traffic camera density. So despite the lack of data from Philadelphia County, the Temple research team expects that the overall results from this study will be generally representative of the study area in this project and conditions across the Commonwealth of Pennsylvania as a whole.

Given that the goal of the study was to match 911 data to PennDOT RCRS entries, it was important to understand what information in the 911 data logs most effectively paired with the PennDOT RCRS records. However, as anticipated, the 911 call center data was provided in a number of different file formats (e.g., Excel files, pdf files, etc.) and with a wide range of information included in the entry fields (Figure 1). Therefore, extra care was necessary to normalize data in a manner that would allow the Temple research team the ability to definitively match entries between the two databases. As data was initially acquired, the Temple research team performed a pilot study with a small subset of the data to manually pair data between the PennDOT RCRS records and 911 call data from Lehigh, Luzerne, Lebanon, Berks, Lackawanna, Schuylkill, Northampton, and Susquehanna counties. This study allowed the Temple research team to analyze the nature of the 911 call data so that the amount of necessary pre-processing and normalization could be better characterized. Based on the manual pairing efforts from the initial pilot study, an integrated framework was proposed based on the following parameters that were noted as critical to confidently identify matches between PennDOT RCRS and county 911 entries: Date/timestamp, GPS coordinates, location, and incident type. As discussed in previous deliverable reports, all data files provided by the counties were then reviewed and subjected to a normalization sequence that ensured consistency between files. Figure 2 presents an example of a normalized data file for Susquehanna County. For the normalization process, all data from a county was manually sorted into the same Excel® file, which included converting files in .pdf or .txt format in some cases. The programming language Python was used to develop scripts that manipulate Excel® files and format the pertinent columns within each spreadsheet into a consistent manner such that Date/timestamp, location information, and incident type were included for each entry in the 911 record.

|  |
| --- |
|  |
|  |
|  |

**Figure 1. Examples of logs provided by (top) Dauphin County (middle) Montgomery County, and (bottom) Cumberland County**

|  |
| --- |
|  |

**Figure 2. Example of normalized logs for Susquehanna County.**

## Integrated Framework & Graphical User Interface (GUI)

Once the 911 data was normalized, an integrated framework was developed to pair this data with entries in the PennDOT RCRS database. This integrated framework was continually revised as the project progressed and the data available to the Temple research team evolved. Generally, all attempts at developing this integrated framework resulted in a similar overall approach: (1) Identify an RCRS entry of interest and identify the county in which the entry was located; (2) Load all records from county of interest; (3) Pre-filter county 911 records to remove any entries unrelated to highway incidents; (4) Identify a threshold timeframe to filter county 911 entries and remove unlikely matches based on time alone (e.g., any entries greater than three hours and earlier than one hour from the RCRS entry timestamp); (5) Use location information (e.g., GPS, mile marker information, location descriptors, etc.) to identify most likely matches to RCRS entry; (6) Develop final criterion to select among most likely matches (if more than a single county 911 entry is a likely candidate). The main difference between different iterations of the framework was the manner in which location information was integrated into the pairing process. Initial efforts with manual pairing highlighted the level of inefficiency present in this overall framework when GPS coordinates were not available. Matching between the databases was largely slowed in this case by the need to manually interpret mile markers and/or location descriptor information to deduce the location of the 911 entry and compare to the location information from GPS coordinates in the RCRS data. This often necessitated conversion of the GPS coordinates into a more qualitative descriptor that could be manually matched to the mile marker and/or location description information field in the 911 data or vice-versa. Additionally, the RCRS and 911 data were still in different Excel® spreadsheets, which necessitated having to switch windows repeatedly during the manual pairing process.

Given the inefficiency of manual pairing, the Temple research team pursued an integrated framework based on an automated algorithm using the Python programming language. This approach used scripts to parse through the county 911 and PennDOT RCRS for time, location, and incident type information. The tool developed a list of potential matching county 911 entries for each RCRS entry. This list was then manually reviewed for selection of the most likely match between county 911 entry and RCRS entry. The algorithm implemented GPS coordinates to filter the potential county 911 matching entries based on location. Direct comparison of GPS coordinates allowed rapid consideration of a distance threshold for filtering potential matching entries in the 911 data for a given RCRS entry. PennDOT RCRS records provided GPS information and two counties (Susquehanna and Lackawanna) also included it with their initial responses to the 911 call data requests. Comparison of the resulting matches using this version of the automated framework with GPS coordinates matched well with efforts from manual pairing for the aforementioned counties. However, the majority of the 911 call centers were unable to provide GPS data, despite repeated efforts to procure such data. As a result, the GPS-based iteration of the integrated framework was not implemented on the remaining counties to match records between 911 and RCRS databases.

Given the unavailability of GPS data for the vast majority of the county 911 records, the Temple research team focused their efforts on developing an iteration of the integrated framework that increased the efficiency of the manual pairing process. Several design constraints affected the development of this framework, all of which were intended to improve the efficiency with which manual pairing could take place. The resulting tool needed to address the following items: (1) be simple to use, with an intuitive user interface; (2) present the data on a single screen with a uniform format so that it is easy to compare records from dissimilar sources; (3) automate as much of the pairing process as possible, leaving the end user with little additional efforts after the tool is implemented for a given set of RCRS entries; and (4) store the paired results in a consistent and easily decipherable format. It was decided that the most effective approach for this version of the framework was a graphical user interface (GUI) developed using the PyQT4 Python bindings for the Qt cross-platform GUI/XML/SQL C++ framework. As before, the pairing tool actually used to match records was built in Python 2.7 using the library Pandas to load, store, and manipulate data. These technologies were chosen to keep the resulting tool cross platform.

A video is provided on the project website that demonstrates the functionality of the resulting GUI tool. Additionally, Figure 3 provides a series of screenshots during operation. The GUI features a toolbar at the top which can be used to specify time difference and distance settings for filtering records (Fig. 3a). Generally, for optimization purposes these filtering criteria were adjusted on a county by county basis. However, a time difference of 2-4 hours and a distance difference of approximately 5 miles were often used as a starting point during the matching process. When multiple 911 entries were potential matches, the earliest record was typically selected to counteract the issue of multiple 911 entries corresponding to a single RCRS entry. Below the GUI toolbar are two scrolling lists of entries. The left column lists the source records and the right column lists potentially matching destination records. In the Fig. 3 example, source records are from the RCRS system and destination records are from the corresponding county 911 database. However, this GUI tool allows pairing to occur in the other direction as well. The currently selected source record is indicated by a thick outline (Fig. 3a). The user can left click on a destination record to indicate that it is a match to the currently selected source record, turning both green (Fig. 3b). Alternatively, the user can right click on any destination record to indicate that the currently selected source record has no matches in the county records, turning the source record red (Fig. 3c). Based on the data present in Fig. 3, the first destination record is likely a match to the selected source record based on the mile marker posts specified in both location fields. This can be marked as a match (Fig. 3d) and the user can proceed to the next source record in the left-hand column by left clicking it. This will bring up a new list of potential matches in the right-hand column (Fig. 3e). Based on mile marker and time, the second record in the destination column is a match for the new source record and can be left clicked to indicate a match (Fig. 3f). The user continues in this manner until all source records have been processed. The resulting pairings are saved automatically to a file specified by the user.

|  |  |
| --- | --- |
| **(a)** |  |
| **(b)** |  |
| **(c)** |  |

**Figure 3. Example of using GUI to locate matching county 911 and PennDOT RCRS records.**

|  |  |
| --- | --- |
| **(d)** |  |
| **(e)** |  |
| **(f)** |  |

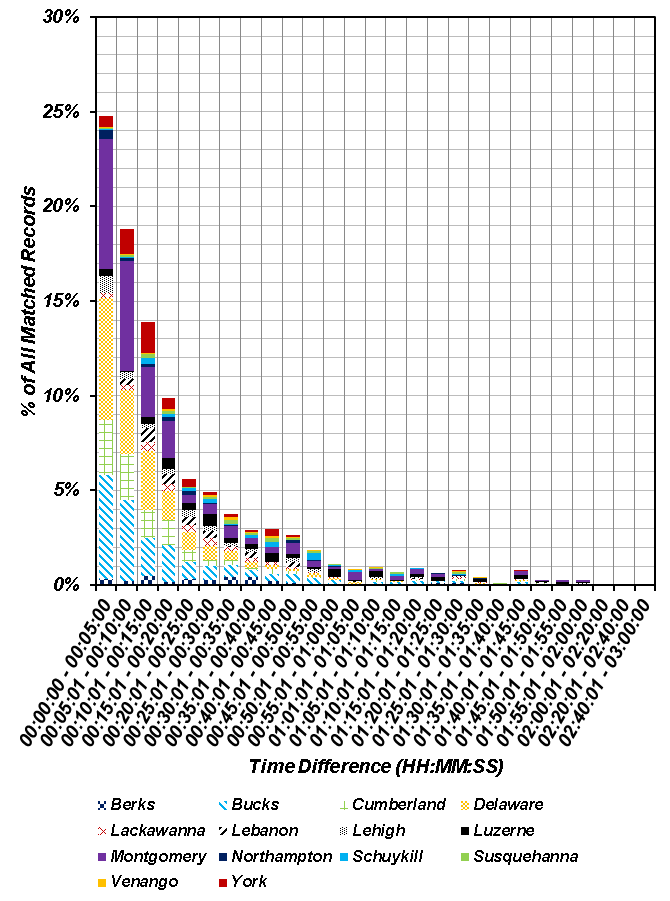
**Figure 3 (cont.). Example of using GUI to locate matching county 911 and PennDOT RCRS records.**

## Data Analysis and Discussion

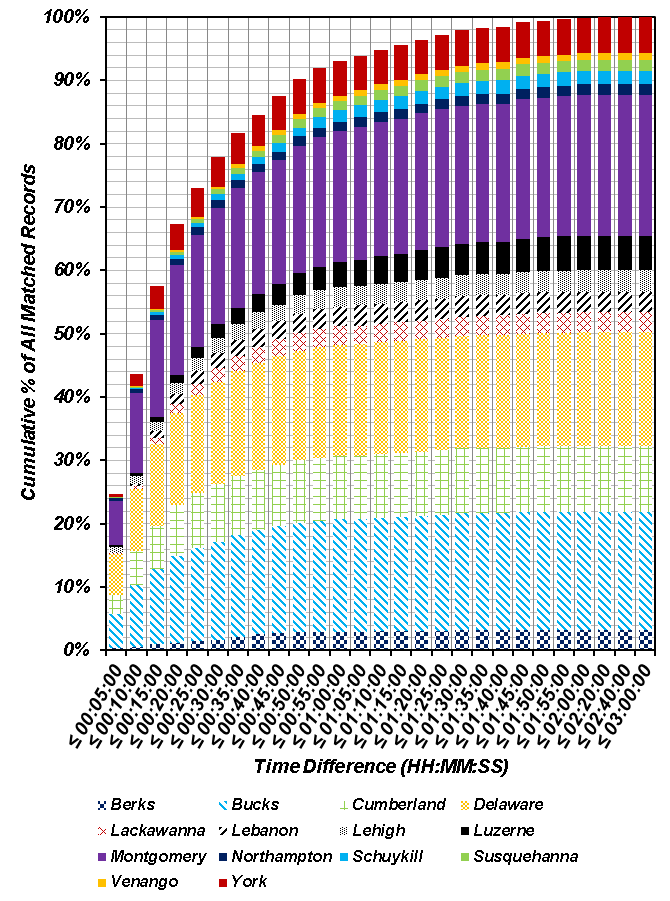
After development of the GUI-based tool, the Temple research team implemented it on all the datasets that had been previously paired (Table 1) to ensure compatibility of results and evaluate the improvements in the amount of time necessary to pair data. Though not as automated and rapid as the GPS-based framework, the GUI drastically improved the efficiency of the process over a fully manual approach. This allowed the Temple research team to complete the data pairing process based on all but two of the county datasets obtained as of July 7, 2017. The Franklin County data was only recently received and contained issues related to data normalization that prevented implementation of the pairing process with the GUI. Additionally, the Dauphin County data received did not allow even manual pairing to RCRS entries (i.e., an overwhelming majority of location descriptors were labeled as “NULL”). As was previously noted, the resulting database included 75.4% of the pertinent RCRS entries outside of Philadelphia County (i.e., 3,909 out of 5,183, excluding RCRS entries from Philadelphia County). Moreover, given the similarities between Philadelphia County and other nearby counties for which the Temple research team acquired data (i.e., Delaware and Montgomery counties), the resulting statistical analysis is expected to be highly representative of conditions throughout the project study area. The following sections discuss major aspects of the resulting statistical analyses performed after pairing was complete for each county.

### Distribution of Data and Measure of Central Tendency

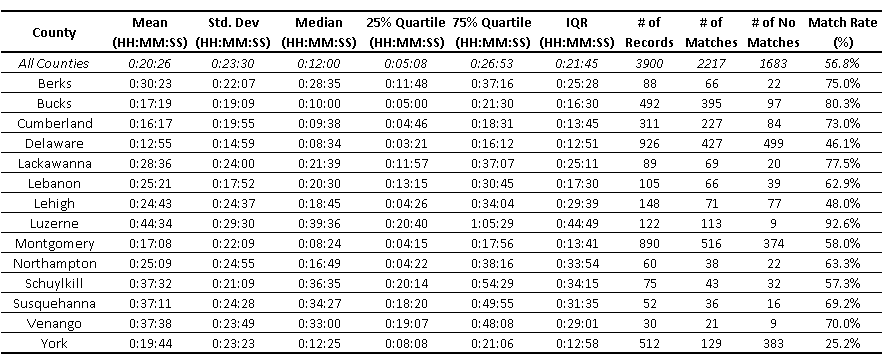
The results from statistical analysis are presented herein using a number of formats to facilitate discussion and interpretation. The main parameter of interest was a measurement of the time difference between when an incident is reported to 911 dispatch personnel and when PennDOT receives notification of a highway incident. Figures 4 and 5 present stacked histograms and cumulative distribution plots for all matched county records. Versions of these figures are replicated in Appendix A for each county separately. All histograms and cumulative distribution plots have been normalized to the number of matched records for a given county (or all counties in the case where that result is plotted in the same figure). Figure 6 plots measurements of the central tendency of the time difference results based on county. Finally, Table 2 provides a tabular summary of the statistical results.



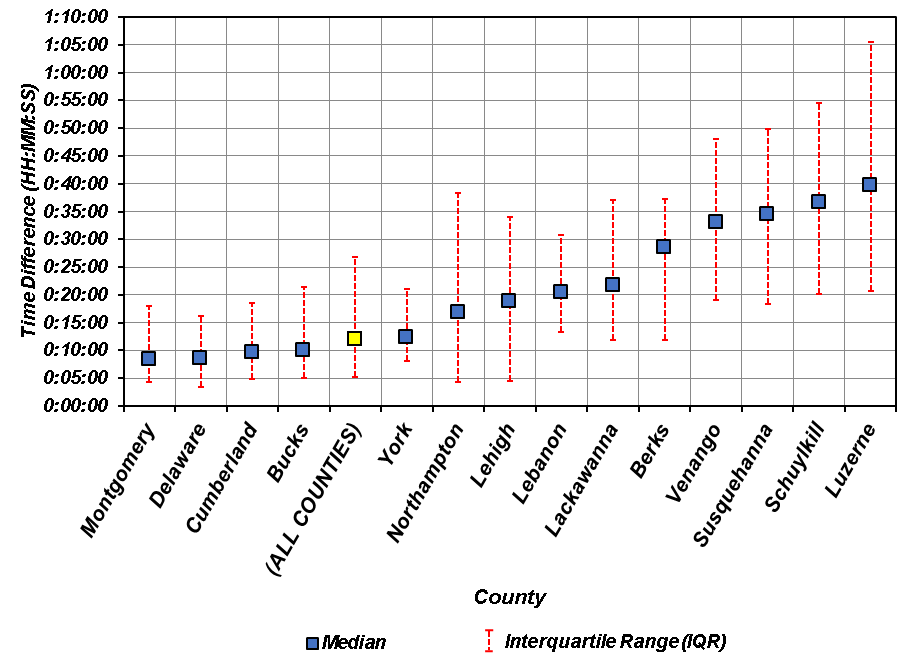
**Figure 4. Stacked histogram of time difference for all county matched records.**



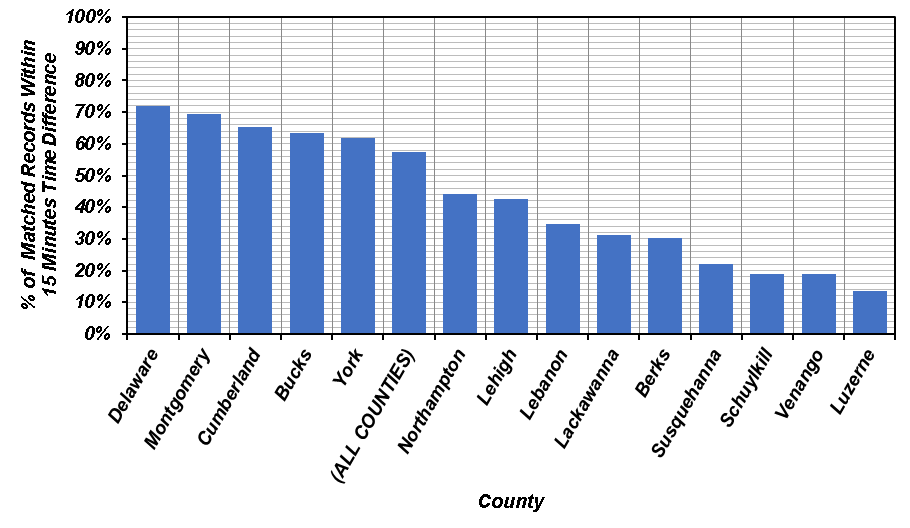
**Figure 5. Stacked cumulative distribution plot of time difference for all county matched records.**



**Table 2. Summary of statistical results based on pairing PennDOT RCRS and county 911 entries.**



**Figure 6. Median time difference and IQR for matched records based on county.**



**Figure 7. Comparison of counties based on percentage of matched records within a 15 minute time difference.**

A number of observations are notable in the aforementioned figures and tables, particularly related to the central tendency of the time difference results. An examination of Figs. 4 and 5 highlights that the resulting time difference is not normally distributed. The distribution is heavily skewed towards the shorter time differences and exhibits a significantly long tail. For example, nearly 70% of all matched records exhibit a time difference of less than 20 minutes. However, approximately 10% of all matched records push that latency to at least an hour or more. This represents a factor of three increase in the time difference. This overall pattern is similar to either an exponential distribution or a Pareto distribution, which is a skewed, heavy-tailed distribution that is sometimes used to model the distribution of incomes and other financial variables. The majority of the counties exhibit a similar distribution when viewed separately from each of the other counties (e.g., see figures in Appendix A). Given the non-normally distributed nature of the time difference results, it is important to use a suitable measurement to represent “average” time difference for all the counties so that overall performance can be evaluated. Generally, in descriptive statistics the “average” value of a parameter highlights the central tendency of its distribution. This central tendency is combined with some measurement of dispersion/variability of a parameter to provide a systematic manner with which to compare and contrast distributions. Typically, the arithmetic mean (*μ*) is often used as a measure of the central tendency of a distribution and standard deviation (*σ*) is calculated to represent variability. However, these parameters are inherently tied to normal distributions of data and can be ineffective or misleading in describing distributions with significant skew (as is the case in the time difference results from this study). The reason for that is that large outliers will tend to disproportionately affect *μ* and the resulting *σ* may describe nonsensical or impossible outcomes. For example, an examination of the statistical descriptors in Table 2 demonstrates that the arithmetic mean of the time difference for all counties was approximately 20 minutes. The standard deviation was 23.5 minutes. A confidence interval of ±1*σ* leads to a negative time difference. Multiple counties suffer from this issue (e.g., Bucks, Cumberland, Delaware, etc.), particularly those that tend to exhibit shorter latency times.

Given the preceding discussion, a better measurement of the central tendency is the median value. The median of a distribution is determined by identifying the middle value when the data is ranked from largest to smallest. In that manner there are an equal number of data points larger and smaller than the median, and the median represents the middle quartile of the data (i.e., 50th percentile). A typical confidence interval that is used to represent data variability with the median value is the interquartile range (IQR). IQR represents the difference between the upper quartile (i.e., 75th percentile) and lower quartile (25th percentile) of the data. IQR is not affected by extreme values and is therefore often used together with the median when the distribution is skewed. Table 2 presents the median values and IQR for each of the counties in addition to *μ* and *σ*. The median values are generally quite different from *μ* as would be expected for distributions with significant skew. Figure 6 plots the median time difference and IQR as the confidence interval to allow a comparison of latency by county. Note that the use of IQR prevents the prediction of negative time difference in the confidence interval surrounding the median value. Based on the data in Table 2 and Figure 6, the overall median time difference between matched RCRS and 911 entries was 12 minutes. Based on the upper quartile estimate, 75% of all matched records have a time difference less than approximately 27 minutes. When viewed at an individual county level, Montgomery, Delaware, Cumberland, and Bucks counties all exhibit time differences shorter than the median computed for all counties. In fact, of those counties, all but Bucks County exhibit a time difference in single digit minutes (with Bucks recording a median of 10 minutes exactly). Generally, a trend could be noted where counties with smaller time differences tended to exhibit smaller variability in their distributions (i.e., smaller IQR). For example, Luzerne County, which exhibited the largest median time difference also had the largest IQR. Overall, there was large scatter in the results when comparing extremes in county-level results. For example, Fig. 7 highlights the county-level differences in how often PennDOT was informed of a highway incident within 15 minutes. A value of 15 minutes for the time difference is roughly between the arithmetic mean and the median for time difference across all counties, so it serves as a useful gauge of overall county statistical behavior. In this case, there is at least a factor of five difference between the county with the highest percentage of matched records with less than 15 minutes time difference and the county with lowest. The largest median time difference was also nearly five times slower than the smallest median time difference (Table 2 and Fig. 6). Finally, the county-level difference in IQR (i.e., data variability) between smallest and largest was nearly as high at approximately 3.5 times (Table 2 and Fig. 6). These results suggest appreciable variability in the various factors affecting 911 reporting practices across the Commonwealth of Pennsylvania

### Geospatial Distribution of Results

In addition to examining the statistical descriptors of the time differences across counties, it is also useful to visually examine the geospatial distribution of these time differences across the Commonwealth of Pennsylvania. This may reveal patterns related to geography that may prove useful for PennDOT practice (e.g., one particular stretch of one of the highways in this study may be significantly different than nearby stretches). Figure 7 presents a choropleth map of counties across the Commonwealth of Pennsylvania with a single hue progression color scheme representing the median time difference computed in this study. Also included for each county is a circular representation of the IQR to highlight variability in the results. An examination of Fig. 7 yields a number of items worthy of discussion. Another observation is that due to issues with data acquisition, much of the RCRS entries along the I-80 corridor across Pennsylvania were not matched to 911 county records. Areas of the Commonwealth traversed by I-95, I-83, and (where data is available) I-76 generally exhibited smaller time differences between RCRS entries and 911 records. Conversely, the sections of I-81 heading north from Lebanon County (and the only counties for which 911 data was procured along I-80 in this study) exhibited larger time differences. There are a number of potential factors that may explain these observations. For example, differences in 911 reporting procedures, allocation of responsibility for incidents along highways, population density, density of the 511PA traffic camera network, and 911 call volume may all play a role in the geospatial distribution highlighted in Fig. 8.

### Match Rate

One item of interest from Table 2 is the rate with which RCRS entries were matched to county 911 records. As noted in Table 2, the average match rate for all counties was nearly 60%. Some counties exhibited nearly complete matching rates (e.g., Luzerne County with nearly 93% of entries matched) while others had less favorable matching rates (e.g., York County with only approximately 25% of records matched). There are a number of explanations for these discrepancies. In an ideal scenario, all RCRS entries correspond to at least one entry in databases maintained by 911 call centers. However, there may be some situations where highway incidents are not reported to 911 call centers or are otherwise unavailable within their CAD systems. For example, PennDOT personnel monitoring traffic cameras may respond to an incident and generate a lane closure without any 911 calls being generated. In some counties, highway incidents may be handled by another nearby agency/call center (e.g., Northumberland County referred the Temple research team to Union County for records related to I-80 in Northumberland County) or by PSP. In many cases, the CAD systems used by the 911 call centers cannot maintain records longer than a specified time period (e.g., one year). In the case of this study where RCRS entries extended back three years, the 911 data provided to the Temple research team was incomplete for certain counties, which may explain the non-matching RCRS entries. Therefore the match rate column in Table 2 highlights the difficulties caused by differences in 911 call center operations, reporting procedures, and CAD systems across the Commonwealth. One lesson from this study is that increased integration of datasets among the various stakeholders involved with highway incidents can begin to address some of the issues noted in this study and improve operational emergency management of highways in the Commonwealth of Pennsylvania.

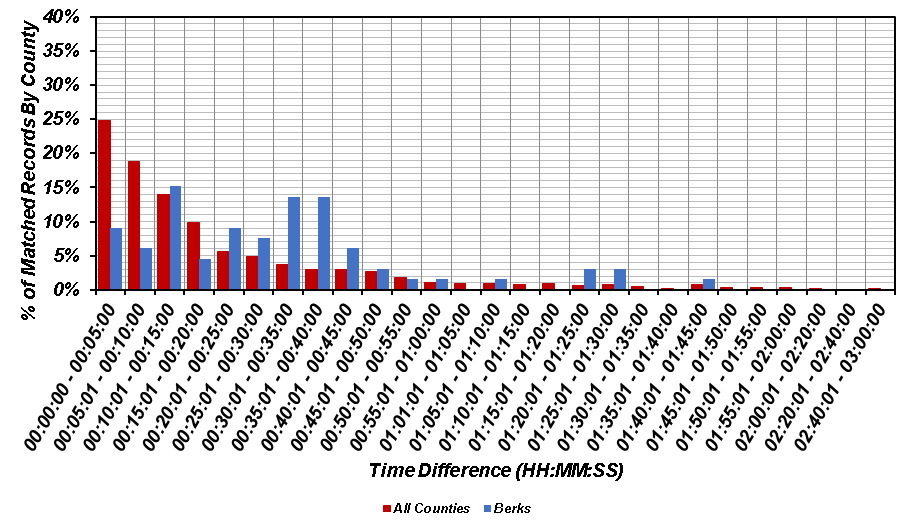


**Figure 8. Geospatial distribution of median time difference and IQR.**

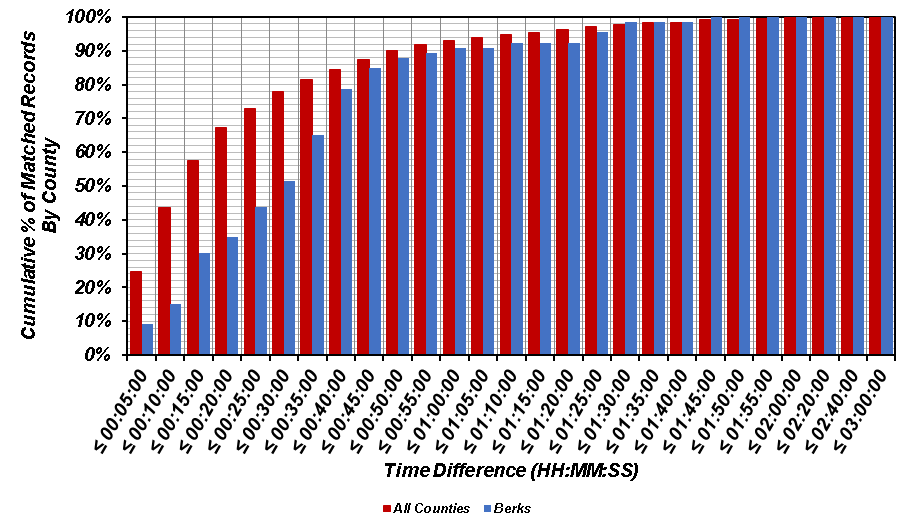
## Future Efforts

The statistical results from this study can aid PennDOT in developing best practices for policy and procedural decisions related to traffic incident management, which can improve operation at the statewide, regional, and district traffic management centers. Estimates of the time necessary for PennDOT to receive notification of highway incidents across the Commonwealth of Pennsylvania is the first step in minimizing the time gaps for highway closures in response to emergencies. Future efforts in this study by the Temple research team will ensure that the data post-processing and statistical framework will be archived. This will include any programming codes/scripts, documentation of research efforts, and outputs from the statistical analysis. All data will be systematically archived as part of curation efforts with the project website. Finally, any errant data that the Temple research team obtains prior to the final task deliverable will be normalized, curated, post-processed, analyzed, and incorporated into the final report.

# Appendix A: Histograms and Cumulative Distribution Plots By County

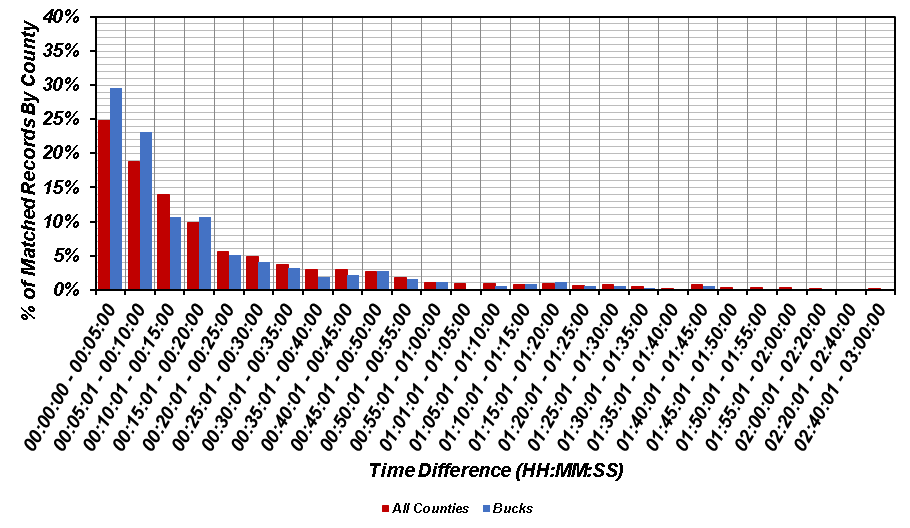


**(a)**

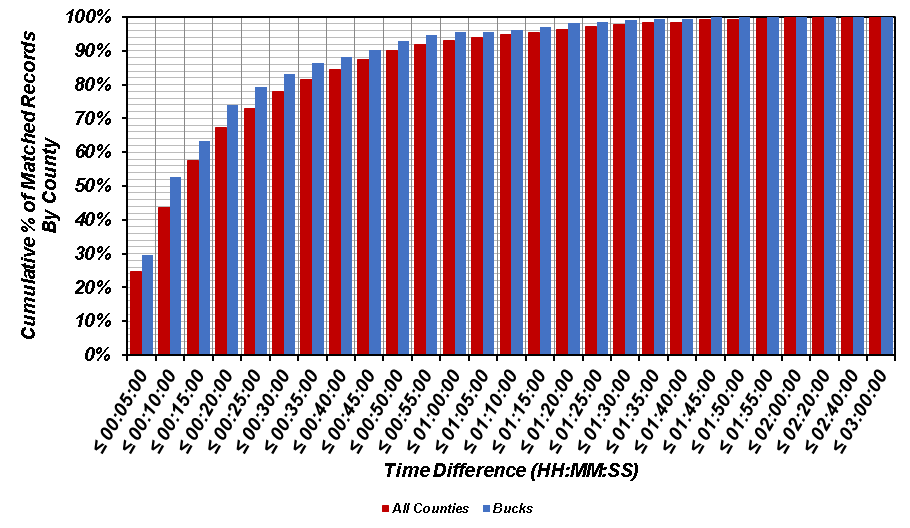


**(b)**

**Figure A.1. Comparison of time difference for matched records in Berks County versus all counties: (a) histogram; and (b) cumulative distribution.**

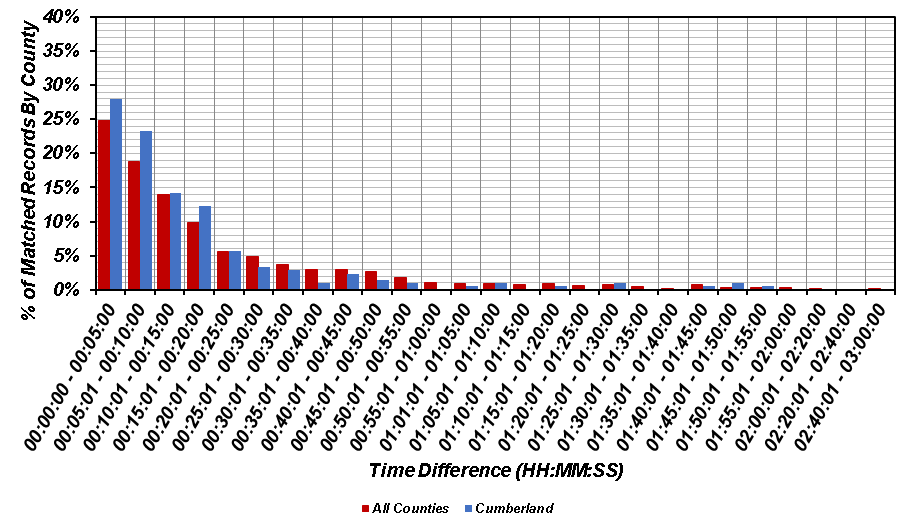


**(a)**

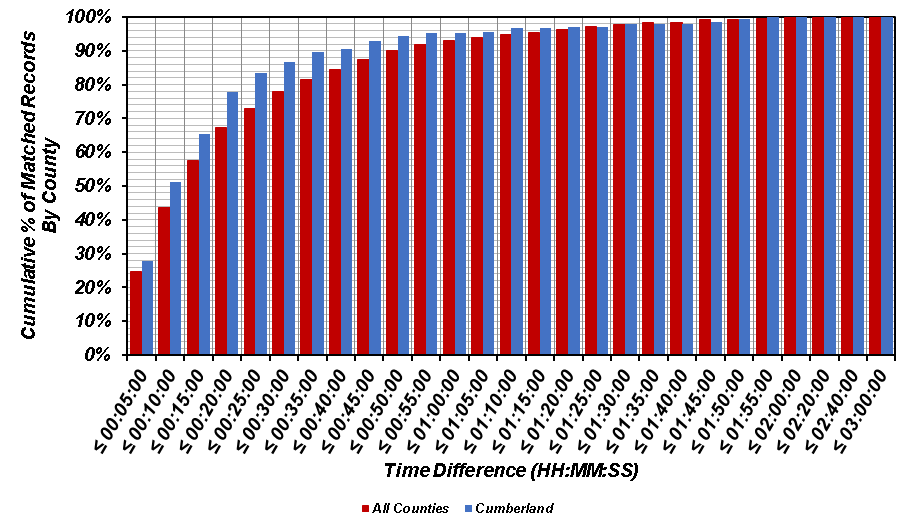


**(b)**

**Figure A.2. Comparison of time difference for matched records in Bucks County versus all counties: (a) histogram; and (b) cumulative distribution.**

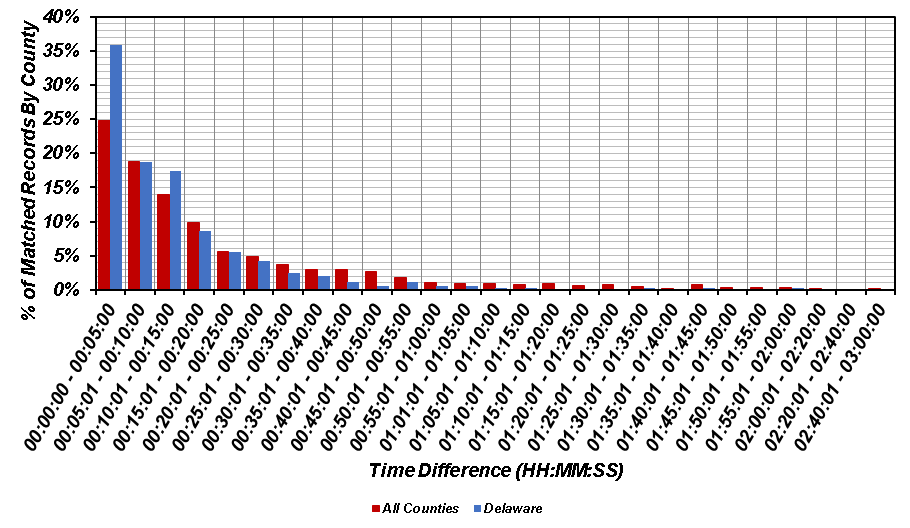


**(a)**

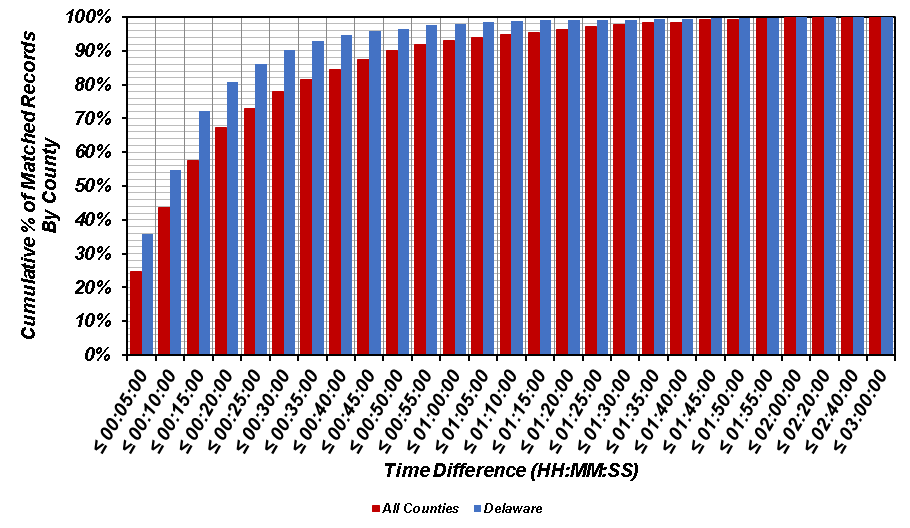


**(b)**

**Figure A.3. Comparison of time difference for matched records in Cumberland County versus all counties: (a) histogram; and (b) cumulative distribution.**

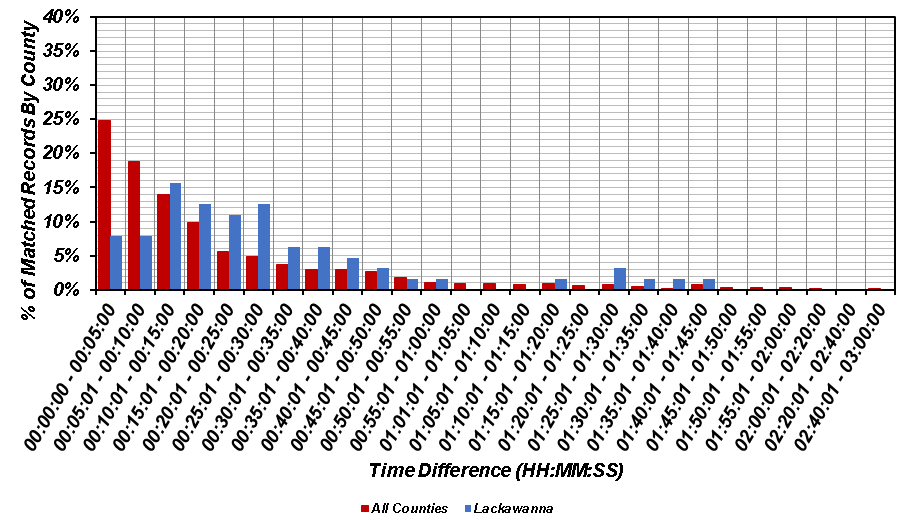


**(a)**

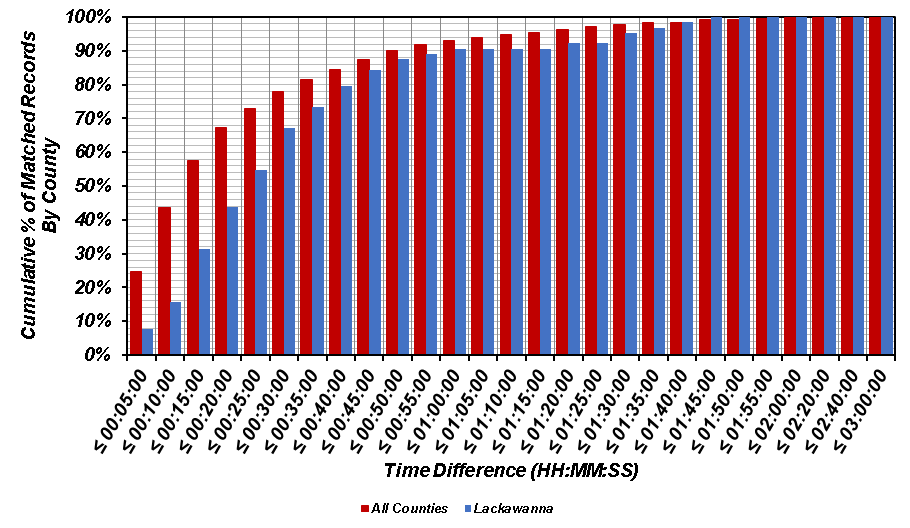


**(b)**

**Figure A.4. Comparison of time difference for matched records in Delaware County versus all counties: (a) histogram; and (b) cumulative distribution.**

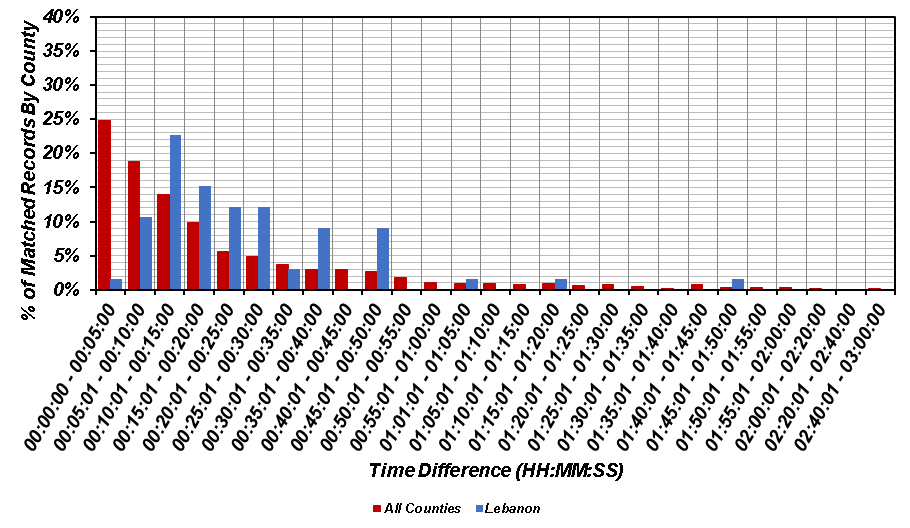


**(a)**

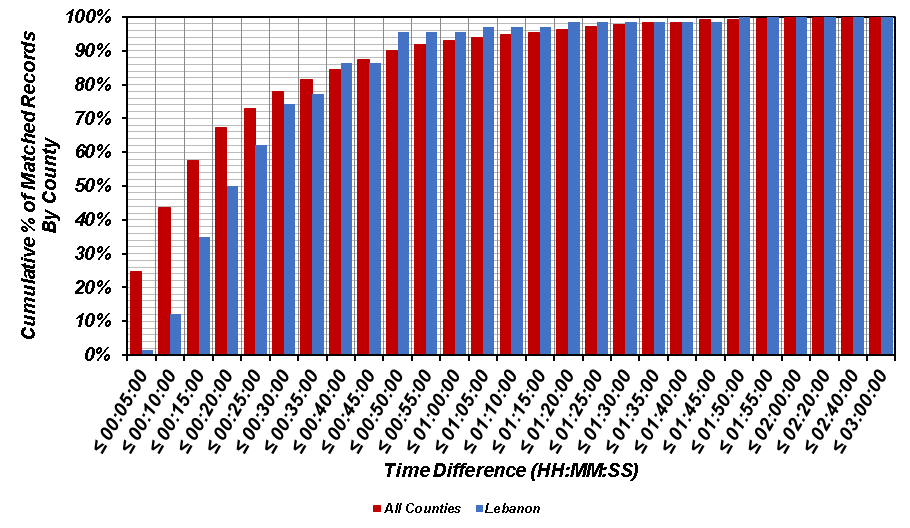


**(b)**

**Figure A.5. Comparison of time difference for matched records in Lackawanna County versus all counties: (a) histogram; and (b) cumulative distribution.**

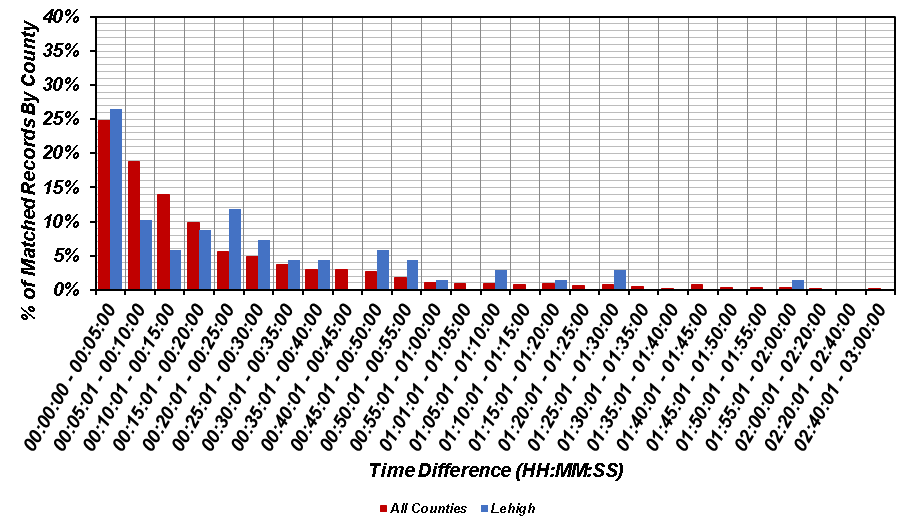


**(a)**

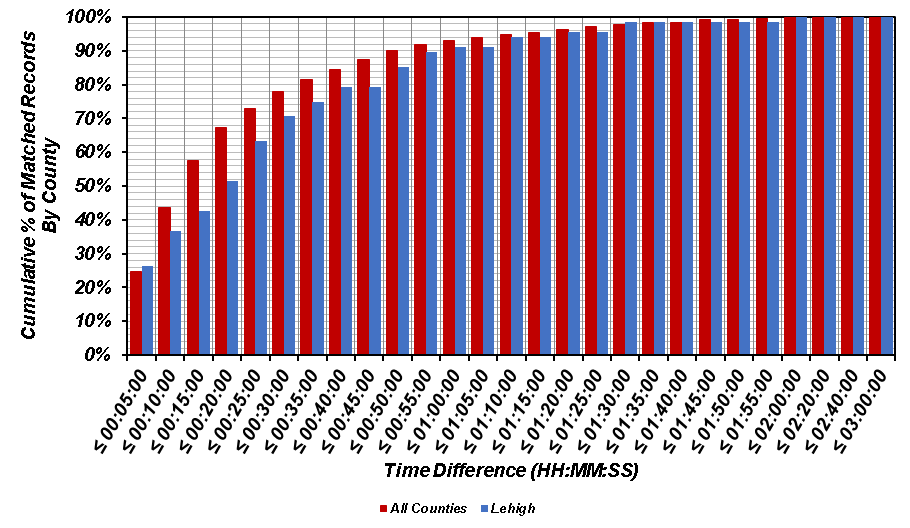


**(b)**

**Figure A.6. Comparison of time difference for matched records in Lebanon County versus all counties: (a) histogram; and (b) cumulative distribution.**

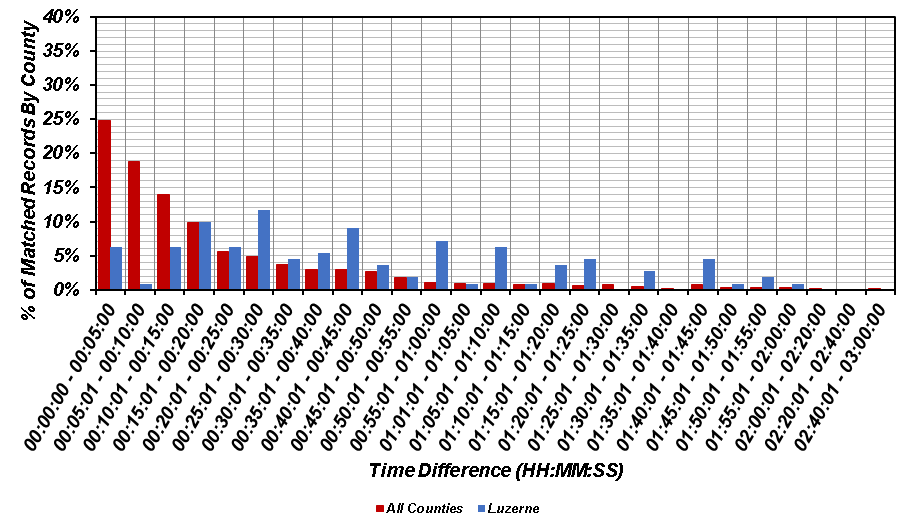


**(a)**

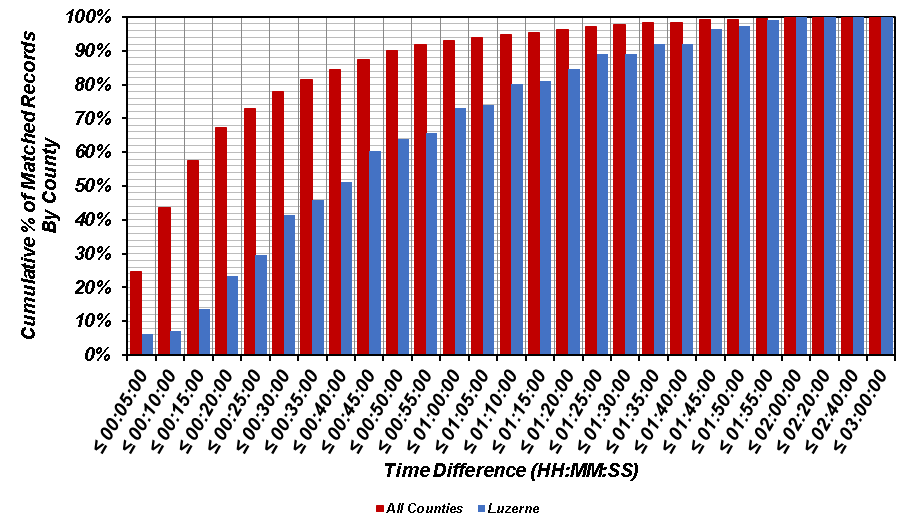


**(b)**

**Figure A.7. Comparison of time difference for matched records in Lehigh County versus all counties: (a) histogram; and (b) cumulative distribution.**

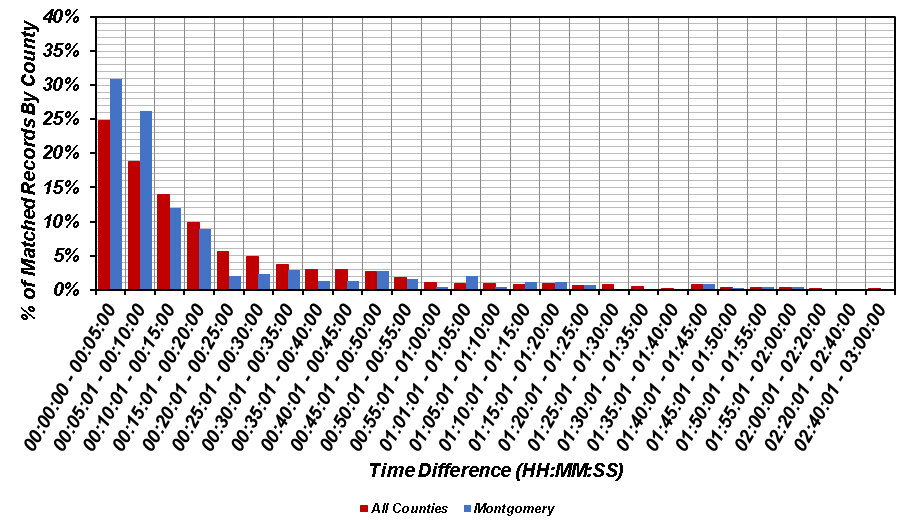


**(a)**

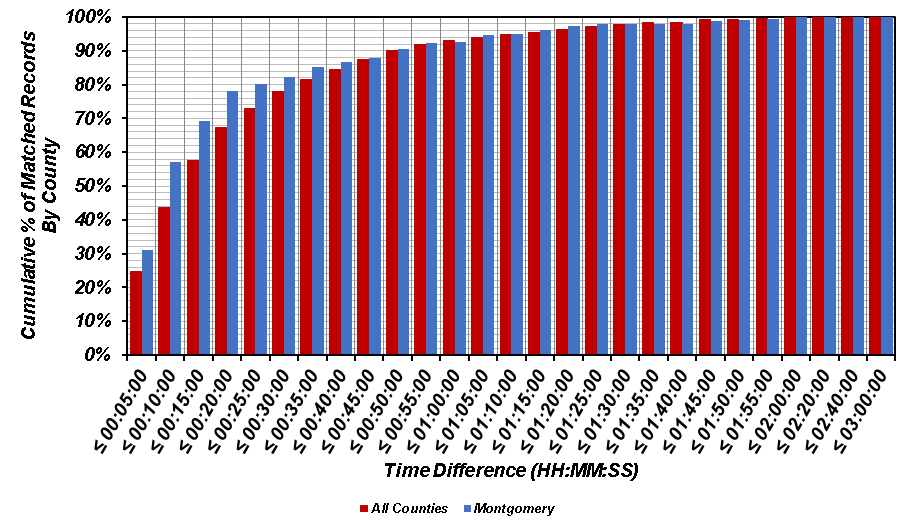


**(b)**

**Figure A.8. Comparison of time difference for matched records in Luzerne County versus all counties: (a) histogram; and (b) cumulative distribution.**

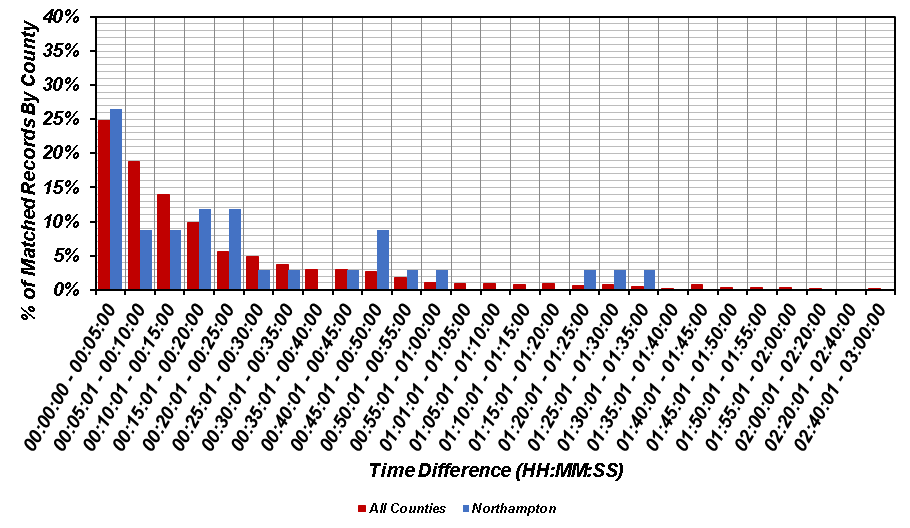


**(a)**

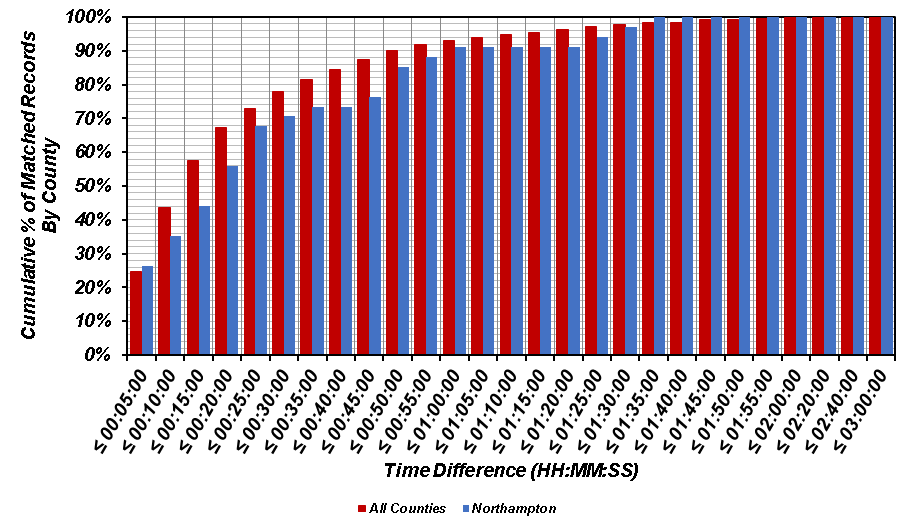


**(b)**

**Figure A.9. Comparison of time difference for matched records in Montgomery County versus all counties: (a) histogram; and (b) cumulative distribution.**

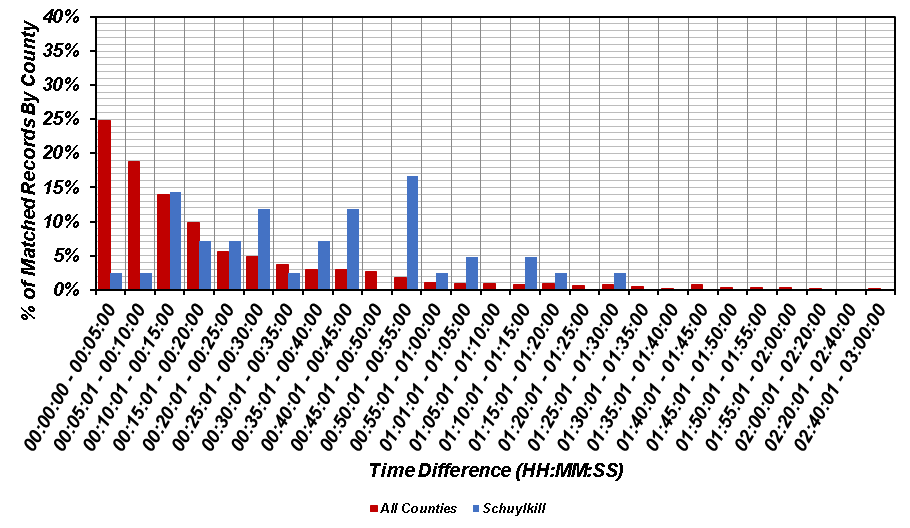


**(a)**

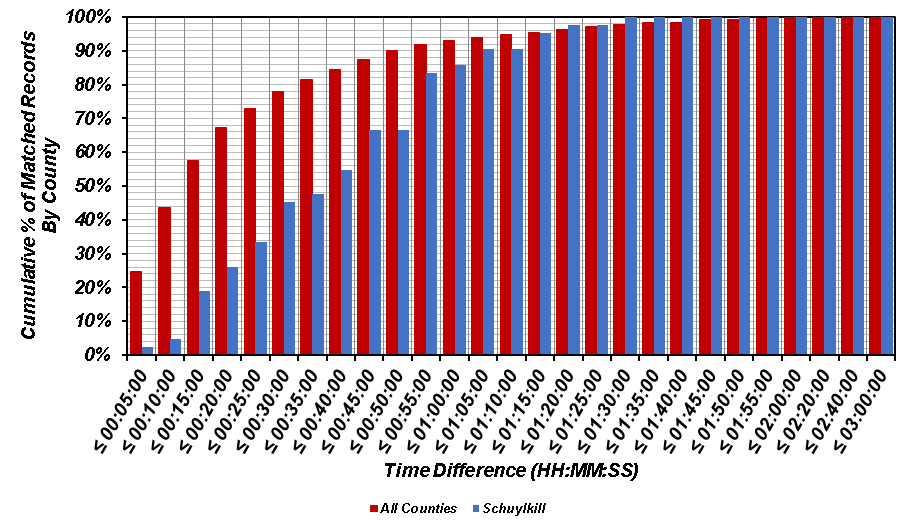


**(b)**

**Figure A.10. Comparison of time difference for matched records in Northampton County versus all counties: (a) histogram; and (b) cumulative distribution.**

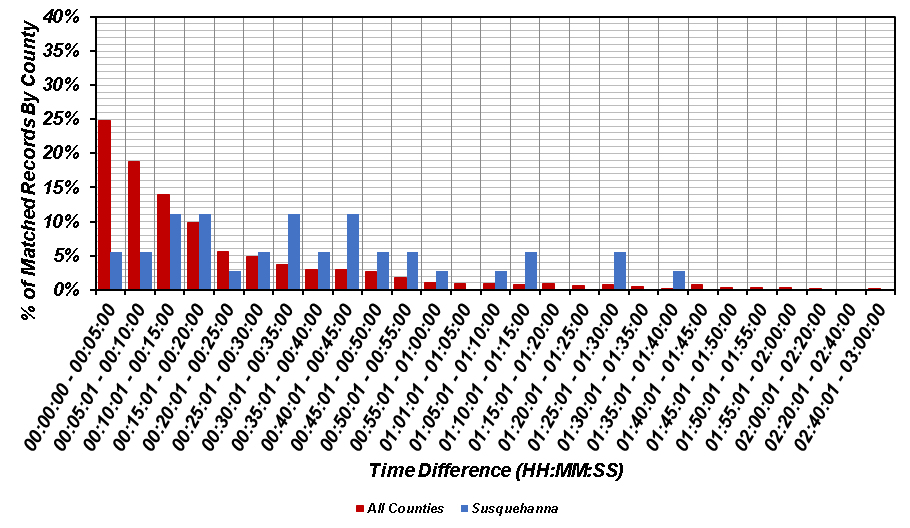


**(a)**

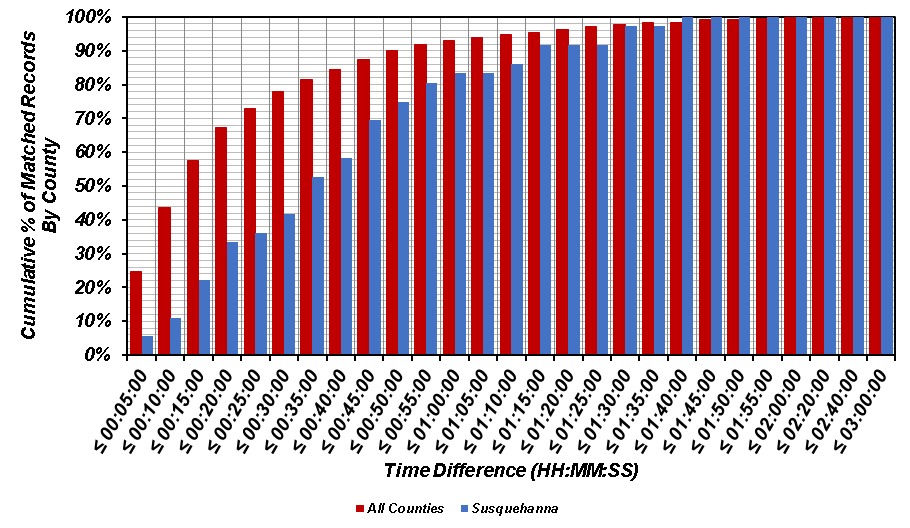


**(b)**

**Figure A.11. Comparison of time difference for matched records in Schuylkill County versus all counties: (a) histogram; and (b) cumulative distribution.**

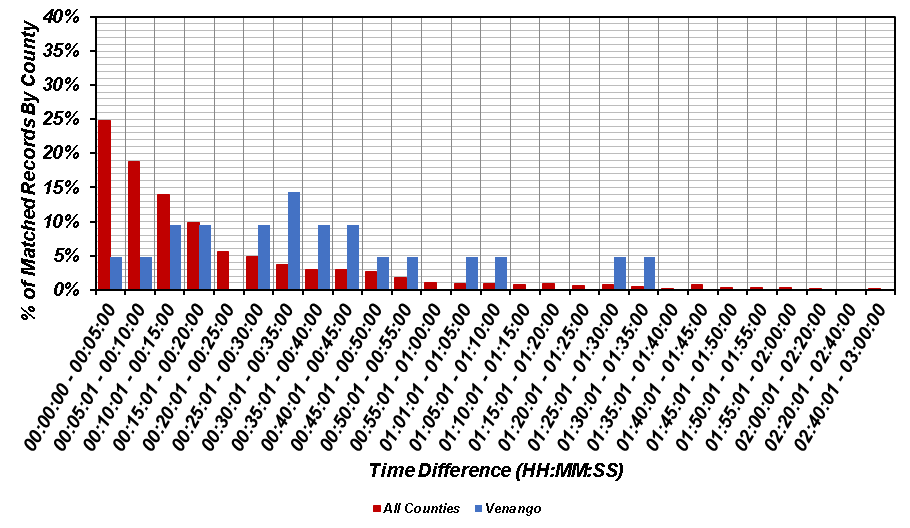


**(a)**

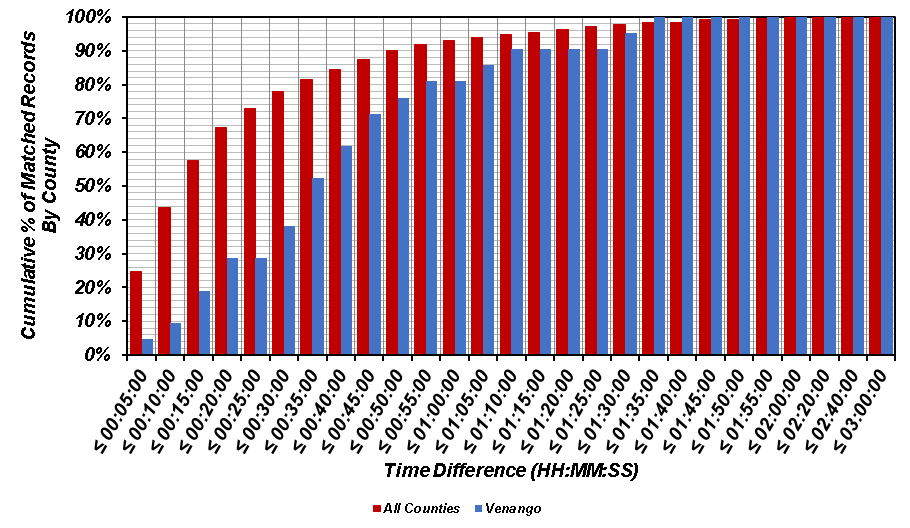


**(b)**

**Figure A.12. Comparison of time difference for matched records in Susquehanna County versus all counties: (a) histogram; and (b) cumulative distribution.**

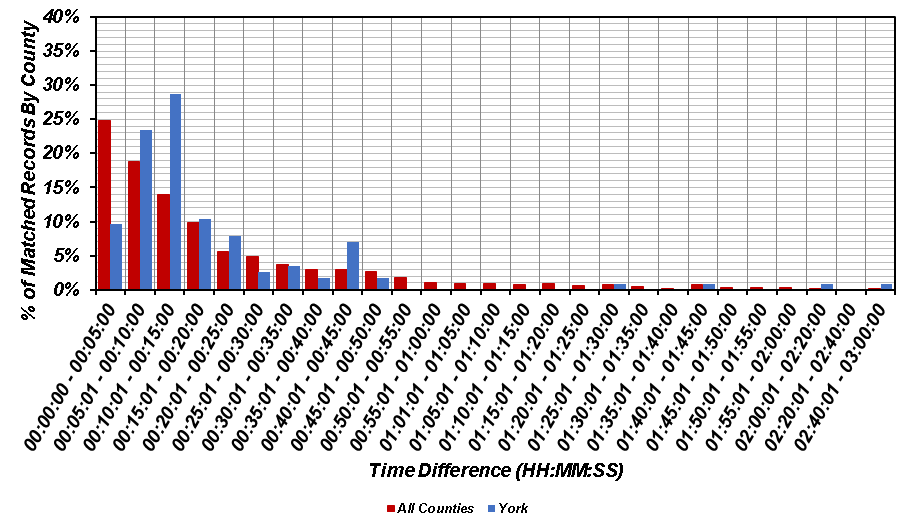


**(a)**

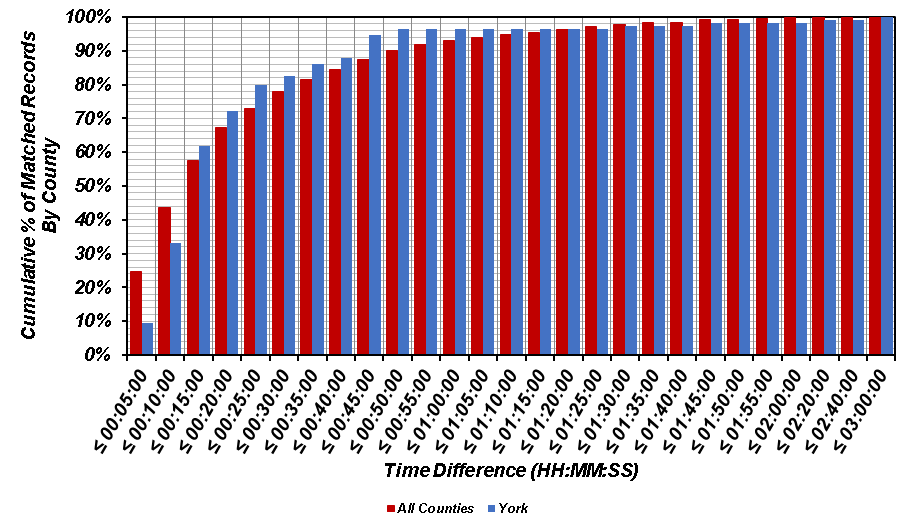


**(b)**

**Figure A.13. Comparison of time difference for matched records in Venango County versus all counties: (a) histogram; and (b) cumulative distribution.**



**(a)**



**(b)**

**Figure A.14. Comparison of time difference for matched records in York County versus all counties: (a) histogram; and (b) cumulative distribution.**