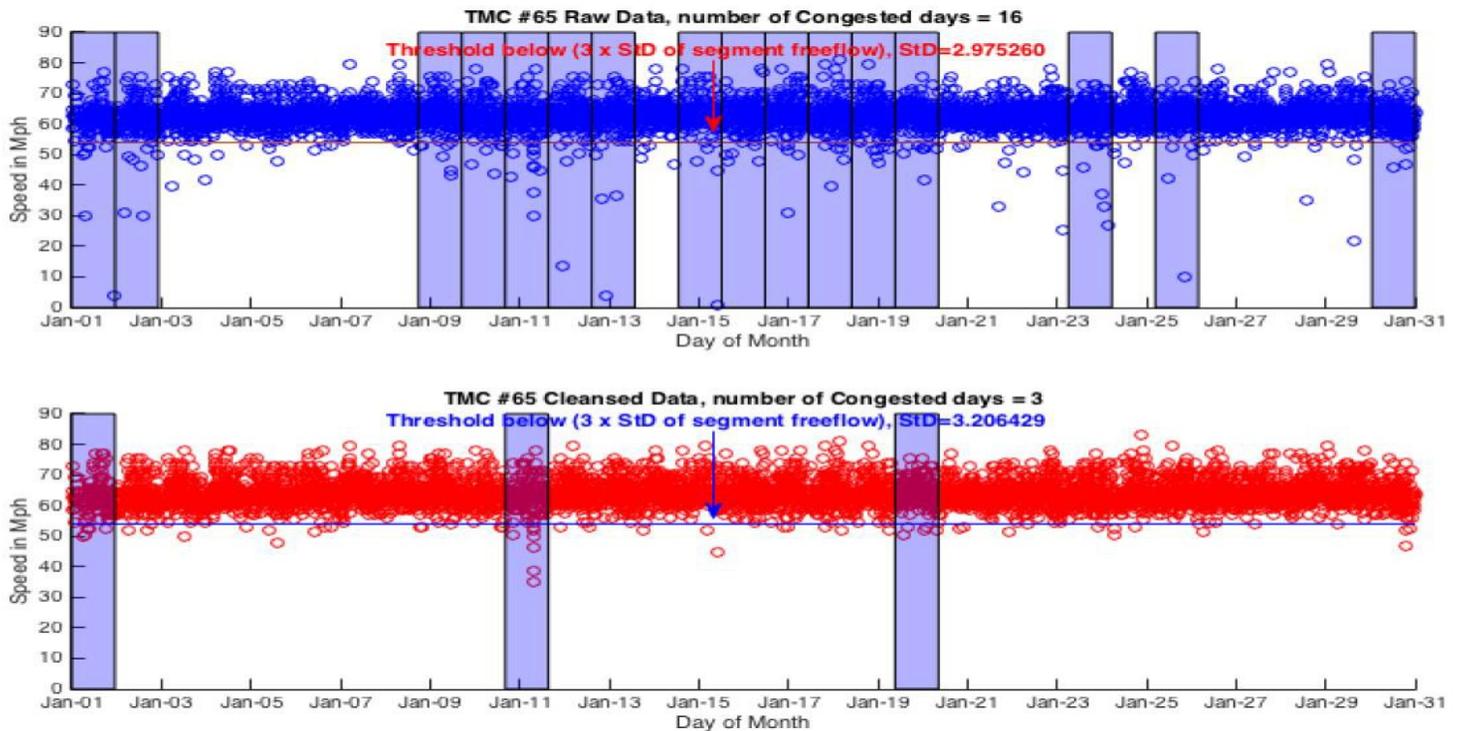


Data extraction, preprocessing, and statistical analysis were performed on the dataset and a comprehensive study of dataset characteristics, influencing variables, outliers and anomalies was conducted. In addition, a study on filtering and removing speed data outliers across multiple road segments was developed, and a comparative analysis of raw baseline speed data and cleansed data was performed. A method for improved congestion detection was investigated and developed. Identification and a computational comparison analysis of travel time reliability performance metrics for both raw and cleansed datasets was conducted, as illustrated below. An outlier removal framework was formulated, and a cleansed and complete version of NPMRDS v.1 was generated. Finally, a validation analysis on the cleansed dataset was presented.



The research affirms that understanding domain-specific characteristics is vital for filtering data outliers and anomalies of this dataset, which in turn is key for calculating accurate performance measurements. Models serve as tools for validating, designing, and understanding the characteristics of travel time measurement systems. Recommendations for improving accuracy and alleviating data anomalies in the NPMRDS v.1 were reported. Research affirms careful consideration of system capture time granularity and segment length must be considered, as the interaction between the two—coupled with the speed of vehicles on the road—could result in generating anomalous data. Statistical analysis confirms that while summary statistics of data averaged over the course of a month is not highly effected by outliers, granular time periods are. Mean and variance statistics exhibited a difference of around 3-5 mph when summarization was done over a period of one day. Finally, in regard to congestion detection, the effect of outlier removal was observed in the reduction of false-positive congestion rates for both the variance and thresholding detection methods. Also, the principle effect of outlier removal was found to be in relation to travel time reliability measures, such as travel time index, buffer time index, and planning time index. Thus, careful consideration for outlier removal must be done when computing these measurements.

POTENTIAL BENEFITS Future ITS systems are expected to manage and resolve the arduous challenges of maintaining and improving roadway performance faced by today's transportation engineers and agencies. This project can aid ODOT in its pursuit of systems incorporating intelligence, coupled with the ability to ingest highly heterogeneous data in real-time for performing various types of inferences (i.e., analysis, diagnosis, exploration, and predictions) that allow insight and knowledge to be extracted and optimal solutions to be employed. This solution will maximize existing transportation infrastructure capacity and improve efficiency.