Data Science for Transportation (2023) 5:9
https://doi.org/10.1007/s42421-023-00071-0

RESEARCH

Boxes-Based Representation and Data Sharing of Road Surface Friction for CAVs

Liming Gao[1] · Juliette Mitrovich[1] · Craig Beal[2] · Wushuang Bai[1] · Satya Prasad Maddipatla[1] · Cindy Chen[3] · Kshitij Jerath[4] · Hossein Haeri[4] · Lorina Sinanaj[3] · Sean Brennan[1]

Received: 11 November 2022 / Revised: 22 April 2023 / Accepted: 15 May 2023 / Published online: 1 June 2023

Abstract

Vehicles can easily lose control unexpectedly when encountering unforeseen hazardous road friction conditions. With automation and connectivity increasingly available to assist drivers, vehicle performance can significantly benefit from a road friction preview map, particularly to identify where and how friction ahead of a vehicle may be suddenly decreasing. Although many techniques enable the vehicle to measure the local friction as driving upon a surface, these encounters limit the ability of a vehicle to slow down before a low-friction surface is already encountered. Using the connectivity of connected and autonomous vehicles (CAVs), a global road friction map can be created by aggregating information from vehicles. A challenge in the creation of these global friction maps is the very large quantity of data involved, and that the measurements populating the map are generated by vehicle trajectories that do not uniformly cover the grid. This paper presents a road friction map generation strategy that aggregates the measured road-tire friction coefficients along the individual trajectories of CAVs into a road surface grid. In addition, through clustering the friction grids further, an insight of this work is that the friction map can be represented compactly by rectangular boxes defined by a pair of corner coordinates in space, a friction value, and a confidence interval within the box. To demonstrate the method, a simulation is presented that integrates traffic simulations, vehicle dynamics and on-vehicle friction estimators, and a highway road surface, where friction is changing in space, particularly over a bridge segment. The experimental results indicate that the road friction distribution can be measured effectively by collecting and aggregating the friction data from CAVs.

Keywords Connected and autonomous vehicles · Friction map · Clustering · Database

✉ Juliette Mitrovich
jfm5876@psu.edu

Liming Gao
lug358@psu.edu

Craig Beal
cbeal@bucknell.edu

Wushuang Bai
wxb41@psu.edu

Satya Prasad Maddipatla
szm888@psu.edu

Cindy Chen
cindy_chen@uml.edu

Kshitij Jerath
Kshitij_Jerath@uml.edu

Hossein Haeri
hossein_haeri@student.uml.edu

Sean Brennan
snb10@psu.edu

[1] Mechanical Engineering, The Pennsylvania State University, University Park, PA, USA

[2] Mechanical Engineering, Bucknell University, Lewisburg, PA, USA

[3] Computer Science, University of Massachusetts Lowell, Lowell, MA, USA

[4] Mechanical Engineering, University of Massachusetts Lowell, Lowell, MA, USA

Introduction

In the field of transportation, significant research, development, and performance improvements have been enabled by digitally mapping the topology and geometry of the road network. These maps can aid drivers: for example, with the help of a digital road network and traffic map, drivers can easily find the fastest route to their destinations. As well, drivers can aid map development: the data from the drivers' phones and traffic sensors can be connected and utilized to generate a live traffic map for choosing the best route (Jain et al. 2019).

In impacting vehicle stability and safety, few data are as important as maps of road surface conditions and the friction between tires and the road surface. Some research has been conducted in this area (Chen et al. 2017; Merritt et al. 2015), yet actionable road-friction maps are not yet widely available. Vehicles are highly likely to spin out or skid unexpectedly when encountering unforeseen hazardous road friction conditions, such as snow, ice, rain, etc. (Hebden et al. 2004). It is well known that these low-friction conditions can even result in traffic crashes (Alhasan et al. 2018; Pu et al. 2020). The Federal Highway Administration reports that approximately 20% of all crashes occur in adverse road conditions. Drivers and driving algorithms can compensate by assuming worst-case conditions and driving slowly, yet this practice can lead to a low level of service (LOS), where the throughput of the highway is not fully utilized. It is known, for example, that light rain or snow can reduce average traffic flow volume by 5–10% on the highway (Highway 2020).

Research has proved that prior estimation of friction allows significant improvements in vehicle chassis control systems (Falcone et al. 2007). Specifically, a vehicle can proactively plan appropriate paths and velocities with the preview of friction distribution, particularly where and how friction ahead of a vehicle may be suddenly decreasing (Cao et al. 2017; Gao et al. 2021a, b). In addition, the availability of friction previews could enhance the performance and reliability of driver-assist systems, such as stability control, adaptive cruise control, and electronic braking. The potential benefit of a friction map to mitigate the impact of road surface conditions on traffic safety and mobility via providing a preview is obvious, but mapping road friction is a challenging problem. Thus, an efficient and cost-effective road surface friction mapping and data-sharing methodology are needed.

Even as vehicles are increasingly equipped to produce friction data to populate friction maps for transportation networks, an additional challenge arises in the representation and sharing of this friction data effectively. This paper aims to address this challenge by proposing a compact two-dimensional (2D) data representation of location-specific road friction information, accounting for several data properties that are road-specific, specifically: that data tends to align in the direction of vehicle travel, and that data regions are most usefully defined by boundaries allowing entry/exit conditions of vehicle stability to be analyzed. The proposed strategy is demonstrated by the authors in ongoing work using Connected and Autonomous Vehicles (CAVs) to aggregate the measured road-tire friction coefficients from a fleet of CAVs.

The data flow of this strategy is illustrated in Fig. 1. In the framework, the estimated raw road friction data streaming from a fleet of vehicles driving on a road segment is shared with a database; then, the raw data are aggregated into a friction map and shared back to vehicles for appropriate driving planning. A naive approach is to finely grid the road surface in the data aggregation step at levels similar to the fidelity

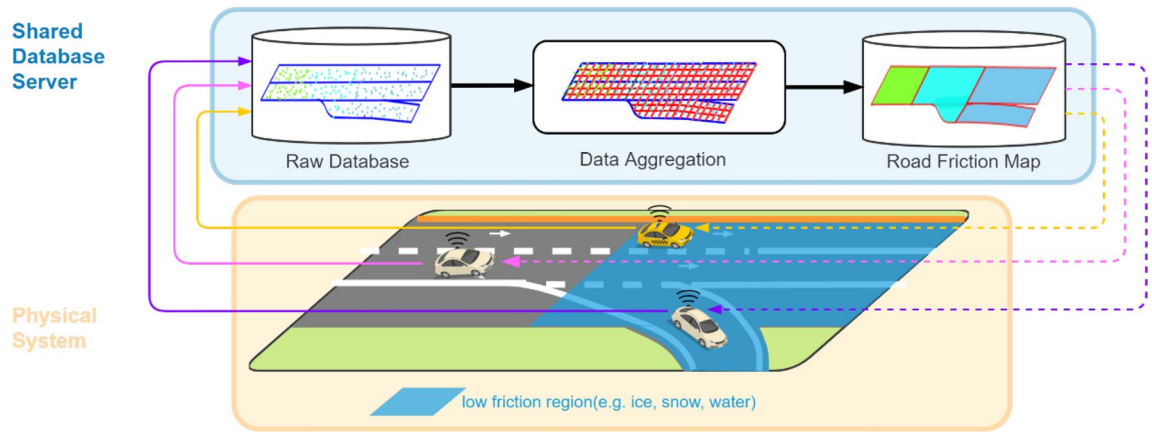

Fig. 1 Road friction map generation strategy by aggregating the measured road-tire friction coefficients along the individual trajectories of CAVs through a shared roadside database

of tire position accuracy and contact patch size, generally on the order of 10 cm; however, if one wishes to maintain this resolution then such databases, even for modest-sized roads, rapidly become prohibitively large. This insight motivated efforts in this paper, such that the friction map can be represented compactly by rectangular boxes defined by a pair of corner coordinates in space and a friction value within the box.

More specifically, this paper demonstrates a data aggregation process to generate a full road surface friction map. This process is explained in detail in the sections that follow, but to summarize: the road surface is first tiled into static roadway-aligned grids with a spatial resolution of 10 cm by 10 cm in a careful choice of the tiling coordinate system. Then, the measured raw friction data are associated with each grid cell and the friction value in each grid is represented by the average value of associated raw friction coefficient measurements, the number of associated measurement points, and the confidence interval. In this way, a grid-based road friction map with a high spatial resolution is generated. However, the data size of this friction map is too large for fast sharing. To represent the map more compactly, the grids with similar friction values and confidence intervals are clustered into a few regions using the K-Means method. The regions are further partitioned into a collection of rectangular boxes to represent them efficiently. As a result, the road surface condition map can be generated quickly and represented compactly as the partition of axis-aligned rectangular boxes with associated friction coefficient values and confidence intervals. To present this process in more detail, the remainder of this paper is organized as follows: related prior works and theoretical background are discussed in "Related Work ". Sample road patterns and friction data acquisition are introduced in "Friction Data Measurement and Collection". The data aggregation and friction map representation are presented in "Friction Data Aggregation". The discussions for the methodology performance are reported in "Analysis and Discussion". Finally, the conclusions and future work of the study are given in "Conclusions and Future Work".

Related Work

The challenge of friction map development involves road friction estimation, raw data aggregation, and map data delivery. Significant research has proven the viability and performance of vehicle-based road friction estimation techniques, which can be broadly categorized as non-contact and contact techniques. Non-contact methods generally utilize special camera-based sensors and computer algorithms to recognize road texture and thereby estimate friction (Roychowdhury et al. 2018; Santini et al. 2021), and such sensors can be installed in a vehicle (Teconer 2020) or a fixed station (Vaisala 2021) to provide measurement and preview. The contact methods measure the interplay force between tires and pavement surface and thereby estimate the friction coefficient based on vehicle dynamics (Acosta et al. 2017; Beal 2019). These techniques enable the individual vehicle to measure the local friction when driving upon a surface or provide some friction preview to other vehicles immediately following, but this approach limits the ability of an individual vehicle to slow down before a sudden decrease in surface friction, especially when vehicles are operating on hills or sharp corners. For one vehicle's data to benefit the performance of another vehicle that is not in direct communication, methods are needed to collect and aggregate friction data from individual vehicles to thereby map road surface conditions. Furthermore, these methods need to be extensible for large-scale traffic systems.

By utilizing the data collected by existing sensing technologies, several data aggregation and representation models for road surface condition-related parameters description have been developed in the literature. Nordic countries including Finland, Norway, Estonia, etc. have developed online road surface condition maps (Mobile Road Condition Map 2021) with data from roadside road condition sensors. In (Panahandeh et al. 2017) machine learning models were trained to estimate and predict the friction class (slippery or non-slippery) for specific road segments using the friction data from connected Volvo cars and data from weather stations. The time-series features from station friction sensing data have been utilized in prior work to improve friction predictive performance (Pu et al. 2020). However, in these studies, the road and thereby friction distributions are represented as one-dimensional line segment features, and common heterogeneous lateral friction features, for example, the "snow rut" scenario (Hussein 2016), are thus neglected. In addition, the station sensors cannot provide the full road network surface map, as they are located on main roads and usually at far distances from each other (Zhan et al. 2020). This work is the first contribution toward providing an accurate 2D spatially dependent friction mapping for an entire road surface through the aggregation of the measured road-tire friction coefficients from a fleet of CAVs.

Friction Data Measurement and Collection

This section presents the friction coefficient data acquisition method from a fleet of connected vehicles when they are driving on a sample highway road segment, where friction changes in space. The data acquisition process is simulated based on the micro-simulation framework proposed in (Gao et al. 2021a, b), wherein traffic simulation tools are used to generate traffic-like trajectories of a large number

of vehicles, and these trajectories are further refined via a secondary layer of chassis dynamic simulations, wherein simulated drivers follow the trajectories, while friction estimation is occurring via chassis measurements.

Description of Testing Road Segment

In this work, a highway segment where friction changes in space, particularly over a bridge, is selected as a sample to demonstrate the friction map generation process. The geolocation and geometry of the road segment are selected from the Interstate-99 interstate highway, outside State College, PA, USA, shown in Fig. 2a. It is a 496 m double-lane highway segment with a 3.8 m lane width.

Figure 2b shows a typical actual road friction distribution pattern with the occurrence of ice as well as real-world photographs of bridges showing lateral snow ruts. A bridge is exposed to cold air more than a normal road surface with a solid roadbed. As a result, when it snows and the temperature drops, a bridge tends to cool and accumulate snow and ice much faster than the surrounding pavement. After a while, the bridge may be completely covered in snow, while the adjacent roads are completely snow-free (Robinson 2021). The transition from the clear road to snow can take place in mere centimeters of travel. The most dangerous threat of this road condition is that it can be abrupt and unforeseen, which catches drivers off guard when they are driving at full speed as the rest of the road is dry or slightly wet. Consequently, vehicles with high speed may lose control unexpectedly when encountering the bridge with this unforeseen sudden decrease in friction. Moreover, snow ruts may occur along the wheel paths when vehicles drive through the snowy bridge. Rutting is a transverse pavement surface condition characteristic, which has a different friction feature from the adjacent road (Hussein 2016; Martin and Schaefer 1996). Consequently, lane changing on roads with snow ruts presents a potential threat to vehicle stability and control. Thus, a road surface map that can provide a preview of this scenario could enable human drivers or autonomous vehicles to plan appropriate driving strategies proactively to avoid loss of control on this type of road segment.

Without loss of the generality, to represent the snowy bridge and snow ruts, we synthetically generate a friction distribution numerically as the "true" road surface condition in this work. These ground-truth data are shown in Fig. 2c and d. To further align the data in directions that are likely similar in friction, the location of the road segment is first converted from latitude–longitude–altitude (LLA) geographic coordinates into the east–north-up (ENU) cartesian coordinates. Next, to align the data with the local roadway directions, the ENU coordinates are converted into station-transverse-height (STH) curvilinear coordinates. In

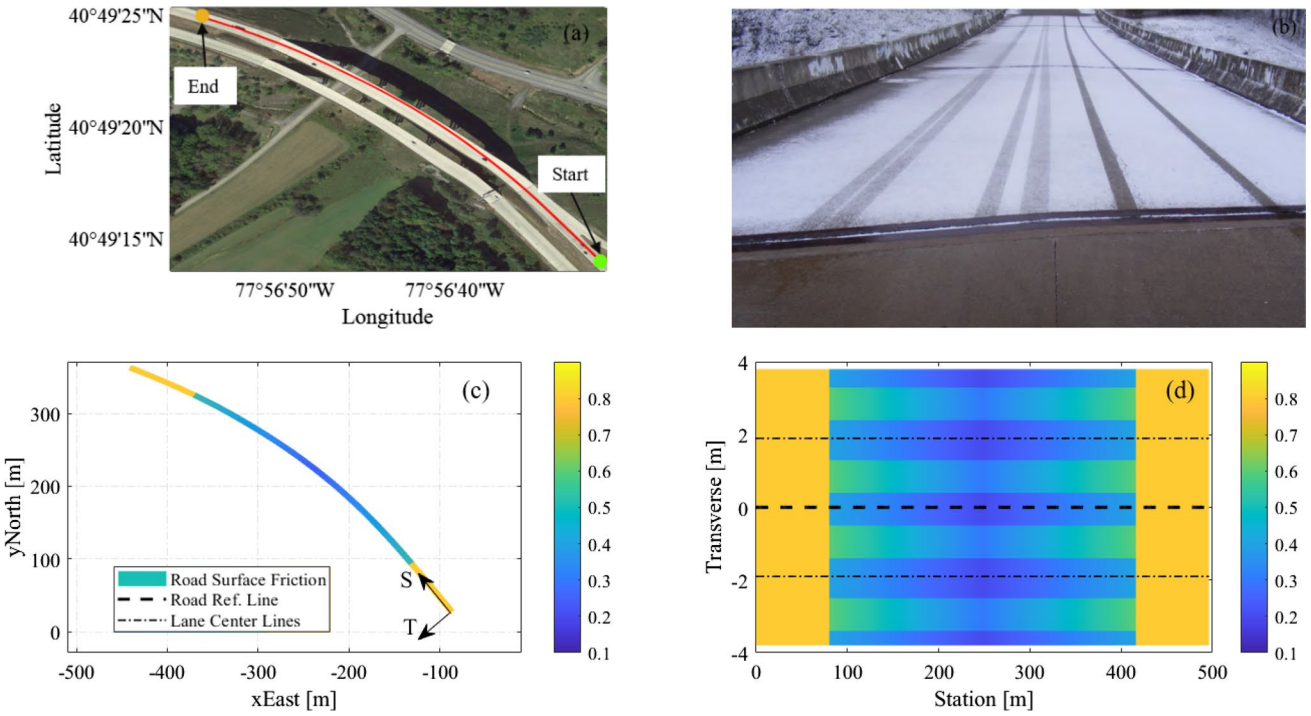

Fig. 2 Testing road segment and surface friction distribution: **a** the geolocation of the road segment, **b** an actual friction distribution pattern over a bridge, **c**, **d** the numerical representation of the road surface friction pattern in ENU and STH coordinates, respectively. The color bar indicates the friction coefficient values

the curvilinear STH coordinates system, the station is the distance traveled along the road segment reference line and the transverse is orthogonal offset from the reference line. The definition of STH is detailed further in "Grid-based road surface representation". This process ensures that the geometric length of road cells and travel distance are preserved; however, the resulting road grid introduces artifacts such as geometric distortions for networks that span a significant curvature of the earth's sphere (~ 100 km or more). The geolocational corrections necessary for data alignment within very large-scale transportation networks are possible with the careful design of road network databases, but this is outside the scope of this paper.

Friction Data Acquisition Simulation

A micro-simulation framework proposed by Gao et al. (2021a; b) is used in this study to simulate the friction coefficient data acquisition process from a fleet of connected vehicles. The simulation framework integrates traffic simulation, chassis simulation, and a shared database for the simulation of a database-mediated CAV study. Using the tool, a highway traffic scenario was simulated for 30 min with highway flow volumes during a winter day morning reported by PennDOT at a traffic data collection site near the I99 segment (Pennsylvania Department of Transportation 2021) shown in Fig. 1a. During this time, 1038 vehicles passed through the road segment with the road friction distribution described in Fig. 1c. In this simulated scenario, each vehicle is assumed to operate a rapid road friction estimator that is integrated into the vehicle dynamic model based on the work by (Beal 2019). Specifically, the surface friction is estimated through direct model inversion using the independent measurements of the left- and right-side front steering torques. White Gaussian noise with a 30 dB signal-to-noise ratio (standard deviation of the noise is around 0.011) was added to the simulated friction coefficient measurements and white Gaussian noise with a 35 dB signal-to-noise ratio (standard deviation of the noise is around 0.025 m) (Abbous & Samama 2017; Diggelen 2007; Karlsson and Mohammadiha 2018) was added to position measurements to realistically reproduce a real application.

The measured data from each vehicle is obtained at 100 Hz and includes the friction coefficient and the corresponding road-tire contact coordinates; these are pushed into a "raw data" database. In this study, all the data are managed by a PostgreSQL 12.09 database server. The data flow is shown in Fig. 1. Even within this very limited time range and highly limited section of the road, the simulation produces about 10 million raw friction measurement data points, which take up 2.7 GB of database storage space. The challenge to generate a compact friction map through the aggregation of such large-size measurement data is addressed in the next section.

Friction Data Aggregation

The goal of this study is to generate a road friction map by aggregating the measured road-tire friction coefficients along the individual trajectories of CAVs. The data flow of this step is also shown in Fig. 1. To present the process, this section begins by introducing the grid-based road surface representation. This section then details the method to aggregate the raw friction data based on the road surface grid.

Grid-Based Road Surface Representation

Before any friction data aggregation steps, the local coordinate system of the grid is carefully chosen as it plays an essential role in defining data similarity and the success of the data aggregation process. A path-aligned curved regular grid is a common way to describe road surface data in high spatial resolution (ASAM 2020). As illustrated in Fig. 3, a road surface space as shown in Fig. 3a can be divided into curved grid structures based on the road reference line which is generally the curved road centerline. Longitudinal

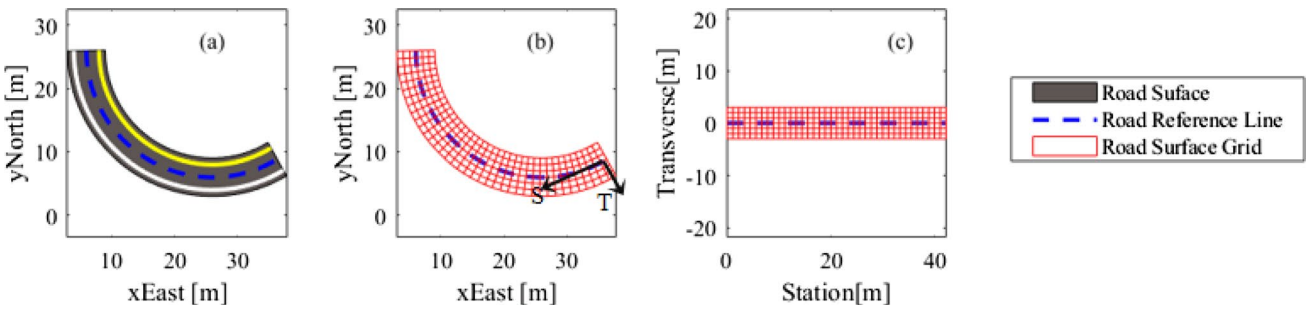

Fig. 3 Grid-based road surface representation example: **a** true road sample, **b** curved road grids in EN coordinates, **c** uncurved road grids in ST coordinates

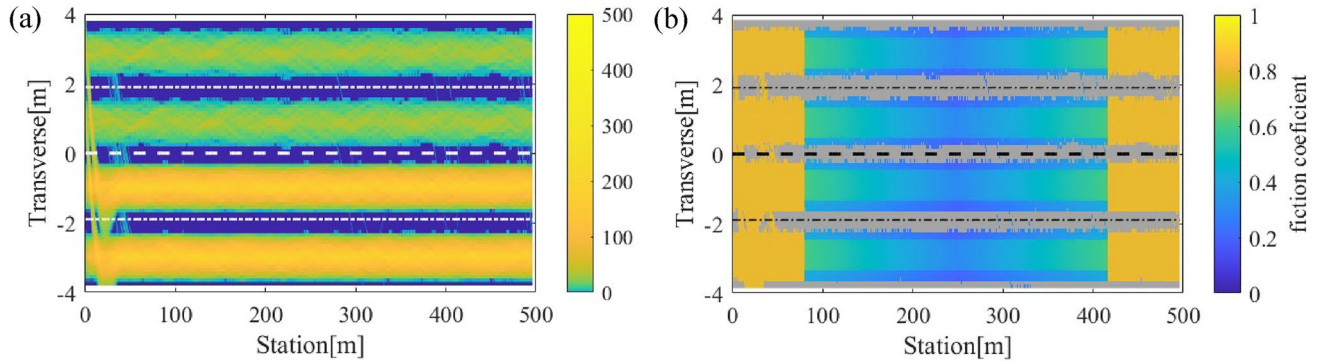

Fig. 4 Association of raw data into a road grid: **a** the number of raw data in each cell is indicated by the colormap in each grid cell, **b** the mean friction coefficient value is indicated by the colormap in each grid cell—the location of the icy bridge segment is visible clearly (no measurement data in the grey area)

cuts are parallel to the reference line, and the transverse cuts are orthogonal to the road reference line. The curved regular grid can be placed in an East–North (EN) Cartesian coordinate shown in Fig. 3b. Based on the grid, any road surface data, e.g., elevation and friction coefficients, is given in the U-direction orthogonal to the E/N-plane. In this way, a microscopic view of the road surface can be described in a three-dimensional(3D) data structure.

But vehicle friction data typically exhibit similarities in the travel or lateral directions. To enforce this similarity, the grid was next placed in the curvilinear station-transverse (ST) coordinate shown in Fig. 3c. ST is a right-handed coordinate system. The station direction follows the tangent of the road reference line and magnitude is the distance traveled along the road reference line from the beginning of the reference line. The transverse direction is orthogonal to the direction of station S and the magnitude is the offset distance from the reference line. Road surface data can be given in height (H)-direction orthogonal to the ST-plane. Using the STH coordinates system, the road surface map can be represented with the uncurved grids and associated height values. Under the STH frame, some frequently used mathematical calculations, such as distance, projection, and transformation, can be highly simplified. It is also interesting to note that in STH coordinates each grid cell can be regarded as an image pixel, and thereby the road map can be represented through an image. Consequently, by converting road data into an STH frame, common digital image processing algorithms and hardware (GPUs, for example) can be used for road surface data processing.

In this work, to test the tire-based estimation of road varying friction, the grid size is 10 cm by 10 cm square to be comparable with a tire's contact patch. To represent the testing road segment surface shown in Fig. 2, about 382,000 grid cells are needed. In this study, the PostgreSQL database with a spatial database extender PostGIS (PostGIS Project Steering Committee 2021) is employed to manage the spatial grid data; the GIS features of the database include spatial query operations, nearest neighbor searches, and range searching.

Cluster the Friction Based on the Road Grids

The purpose of friction-data aggregation is to associate friction values with appropriate spatial context, which implies raw data clustering. Clustering large-size data directly is time-consuming (Aggarwal and Reddy 2014). Inspired by the grid-based clustering algorithm which has a great speed advantage when the data size is large compared to the grid size (Aggarwal and Reddy 2014). In this study, the database includes 10 million raw friction data points from 30 min of simulated measurements; these data are dense when compared to the 382,000 static grid cells. However, most of the 300 k cells are nearly identical to their neighbors. To produce a compact representation, the raw data are first clustered based on the pre-generated static road grid in the database to expedite the friction map generation. Specifically, the raw data are assigned to each road grid using nearest-neighbor searching, a fast process when using hierarchical tree data structures within the PostgreSQL spatial database. Next, the friction estimation in each grid is represented by the average friction coefficient value, the number of raw data, and the confidence interval. The confidence interval is calculated based on 95% probability intervals of t-distribution. A smaller interval magnitude indicates a more accurate estimation. The results are shown in Fig. 4.

From Fig. 4a, one can see that many measurements occur at the locations of snow ruts. Furthermore, one can observe that the right lane has more vehicle traffic and thus measurement data than the left lane. The results are consistent with the typical traffic pattern, wherein more vehicles prefer to drive on the right driving lanes, especially on snowy roads. In addition, one can also see, around the 20 m station mark, that many vehicles change lanes before entering the

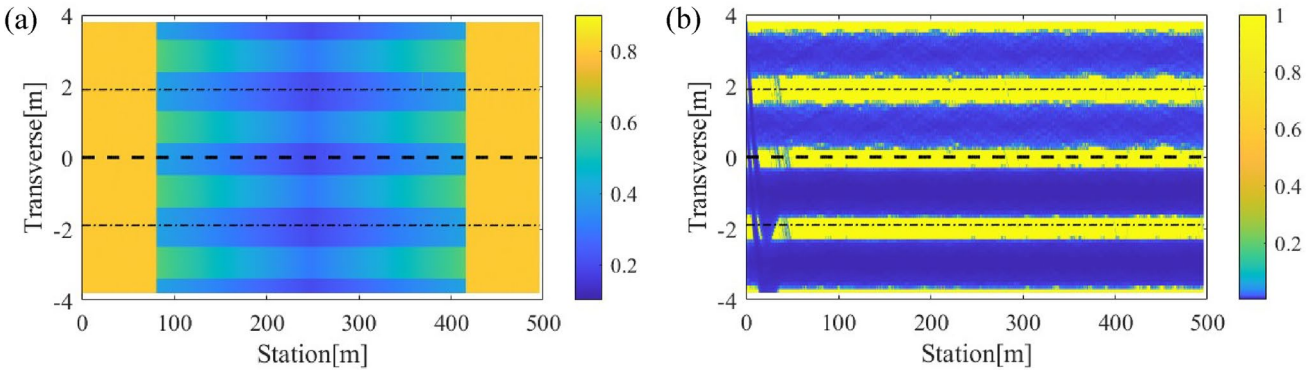

Fig. 5 Friction map after correcting for missing data: **a** the mean friction coefficient value of each grid cell (on a scale of 0 to 1), **b** the confidence interval magnitude of each grid cell (95% probability intervals of t-distribution)

snowy bridge. The plot of Fig. 4b illustrates that there are many grids, especially within the inter-lane area, that do not have any measurement as no vehicles pass through these areas—vehicles very rarely changed lanes on bridges in this simulation. To deal with the problem of missing data, this paper fills the friction mean value to the missing data grids using its nearest non-missing grid data, and as well fills the confidence values for missing data with the worst confidence interval to indicate high uncertainty. The mean and confidence interval are shown in Fig. 5. The camera-based friction estimation method (Roychowdhury et al. 2018) which can cover a larger road surface area can also be utilized as a supplement to the road-tire-based friction estimation in future work.

The results in Fig. 5a include the mean and represent a friction map of the road. Comparing the results with the true friction distribution shown in Fig. 2d, one can see that the boundaries between the snowy bridge and adjacent dry road and the edges between snow ruts and loose snow road area are mapped clearly. Quantitatively, the root-mean-square error (RMSE) between this friction map and the true friction distribution was found to be 0.00440, which indicates an accurate mapping. This map describes the entire road surface with a high spatial resolution of 0.1 m by 0.1 m. However, a database of about 131 MB in data size is required to store this less than 500 m road friction map information; this is not scalable to large road networks. Moreover, considering these data may be shared with each vehicle in the road segment for driving assistance, the large size limits the practical data sharing and transfer in practical applications of V2x systems.

Examining Fig. 5a data carefully, one observes that the friction grids in some areas are similar to each other. Thus, similar cells can be grouped into clustered regions to reduce the data size. To achieve this, this paper uses a spatial clustering method based on the K-Means clustering algorithm (Aggarwal and Reddy 2014) due to its easy application and high efficiency. K-Means clustering categorizes the N spatial friction grids into K clusters in which each grid belongs to the cluster with the nearest mean. Estimating the cluster number K is a major challenge in applying the K-Means algorithm. In this work, K is determined through the statistics results of friction values and expected cluster resolution. Specifically, the friction coefficient values at all grids are grouped into 10 bins at every 0.1 interval ranging from 0.1 to 0.9 (<0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, >0.9), which is shown in Fig. 6a. The number of clusters, K, was chosen to be the number of bins whose percentage is larger than 1%. From the plot, K is equal to 6 in this case. Next, the K-Means clustering algorithm is used to group the friction grids, i.e., partitioning the road surface. Because clustering seeks to maintain both constant spatial coordinates and constant friction values, it is important that the scaling of spatial information is similar to friction information. In particular, the ST location values of grids are normalized into the same scale (0–1) and a large weight (30) is assigned to the friction attribute (on a scale of 0–1) when creating each cluster. The clustering result is shown in Fig. 6b. It reveals that the friction grids are clustered into friction blocks and the friction value of the grid within each block is nearly identical.

Comparing the clustering results with the true friction distribution shown in Fig. 2d, one can see that the boundaries between the snowy bridge and adjacent dry road and the edges between the snow rut and loose snow are mapped clearly. Quantitatively, the root-mean-square error (RMSE) between this clustered friction map and the true friction distribution is 0.0180, the root-mean-square percentage error RMSPE is 5.36%, and the maximum absolute error (MAE) is 0.0453, which implies an accurate mapping. It should be noted that the interval and thereby the cluster number K can be tailored to get specific cluster results, which is further discussed in "Granularity, data size, and mapping accuracy".

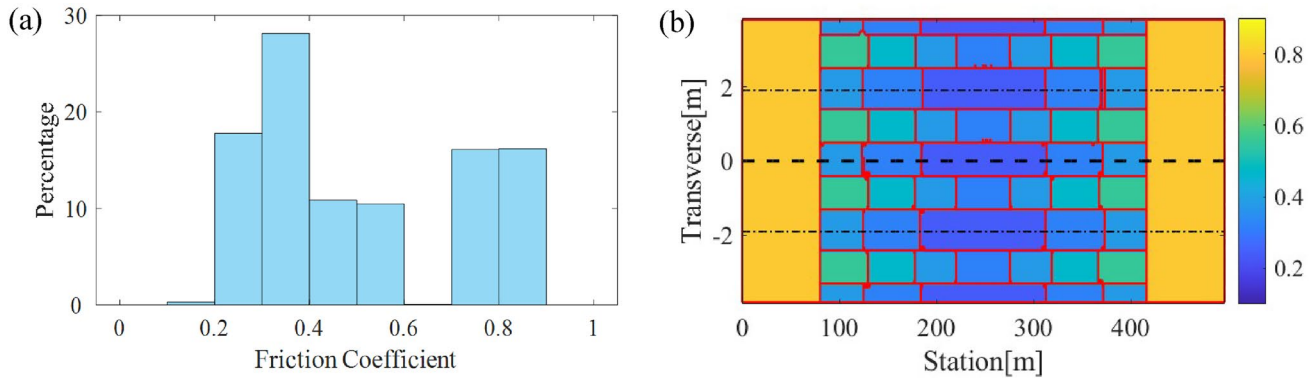

Fig. 6 Road friction grids clustering: **a** histograms of friction values in all grids, **b** clustering results with K = 6. A total of 55 clustering blocks are obtained, and the red lines are the cluster boundaries

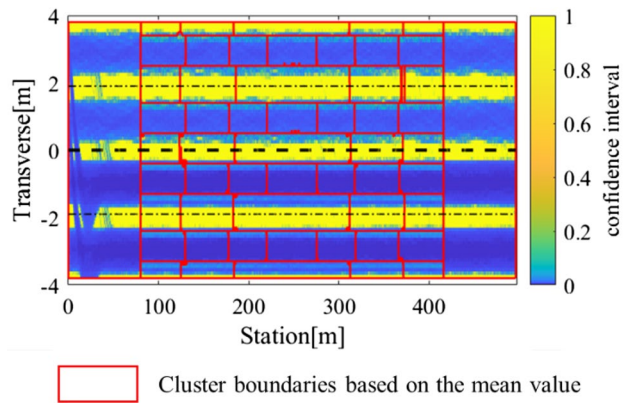

Fig. 7 Confidence interval magnitude of each grid cell and the cluster boundaries for the clustering results shown in Fig. 6b

Friction Grids Cluster Considering the Confidence Interval

Aside from the mean value, the level of uncertainty associated with the friction measurements is also important for friction representation (Beal and Brennan 2020; Walton 2018); this uncertainty can be captured by the confidence interval value on the friction within each cluster. However, the confidence intervals of each cell in a cluster may not be uniform. For example, Fig. 7 plots the confidence interval magnitude of each grid cell and the cluster boundaries for the clustering results shown in Fig. 6b.

As one can see, even in the same cluster region, the magnitude of each cell's confidence interval may be significantly different. One can create subclusters that match both the friction magnitude and the confidence by applying a clustering algorithm again within each friction cluster and then sub-clustering cells based on the confidence interval of the friction coefficient data.

The resulting subclusters, shown in Fig. 7, reveal friction data that approximate regions, where the friction has similar statistics as measured by the mean and variance. To determine the number of subclusters to use, one can use statistical analysis of the subcluster values.

Because these data are simulated and the noise properties are simulated with well-conditioned statistical properties (e.g., normally distributed and unbiased), the data show two very distinct confidence regions: (1) highly traveled regions, where the vast majority of tire contact points (e.g., the "wheel rut" area of the roadway), and (2) all the other areas that would only be traveled during excursions from the lane, such as lane changes or large lane deviations. The notion of two different areas of confidence is further seen in a histogram analysis. Figure 8a shows the histograms (20 bins) of the confidence interval magnitude over all grids. It can be seen that 2 clusters, one for tight confidence bounds (less than 0.05) and one for large uncertainty (greater than 0.9) account for 97.5% of all data. Figure 8b shows the clustering results of the confidence intervals imposing 2 confidence categories. The subclusters are then found by K-means clustering to produce subclusters and their boundaries. A total of 98 subclusters are obtained and the results are shown in Fig. 8b. Again, each friction subcluster now approximates regions with the same mean value and confidence interval on friction.

Represent the Friction Cluster Block with Rectangular Boxes

Examining the clustering results shown in Fig. 6b, one can see that the road surface is partitioned into friction blocks which is a collection of road grids with identical friction coefficient values. Fig. ure 9a provides an example of a friction block which is the region within the bold red polygon. To represent friction blocks efficiently, they are further partitioned into a collection of rectangular boxes shown as the

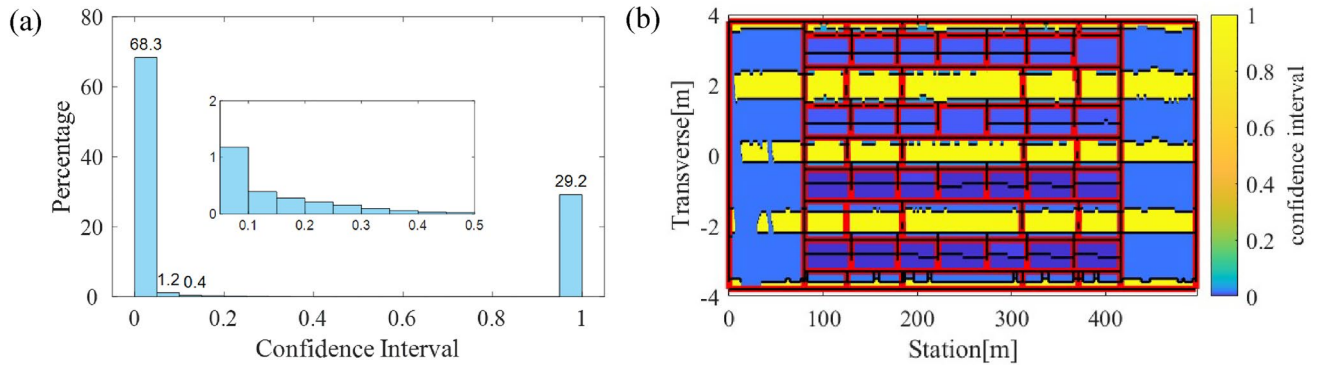

Fig. 8 Friction grids clustering results considering the confidence interval: **a** the histograms of confidence interval magnitude in all grids, **b** the red lines are the boundaries based on friction mean value and the black lines are the boundaries based on both mean value and confidence interval

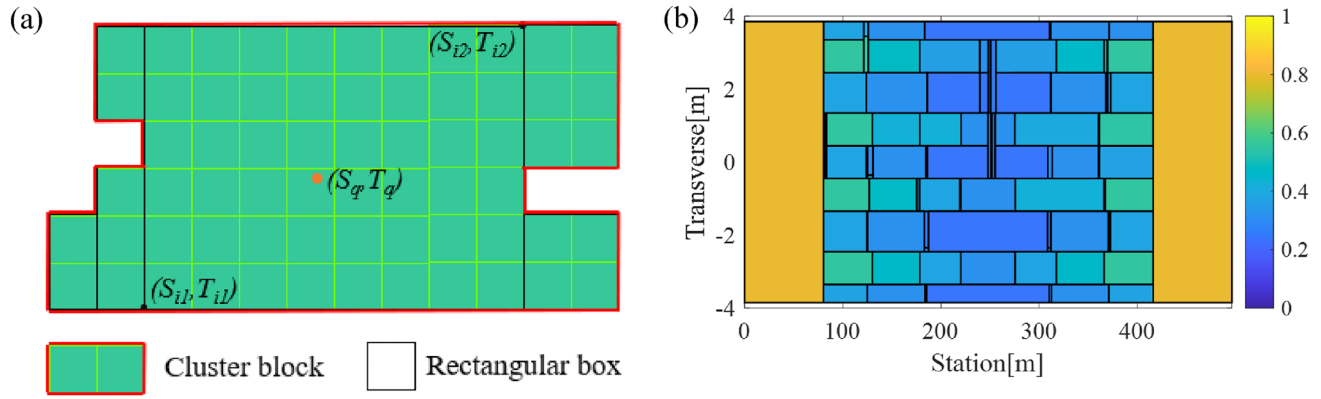

Fig. 9 Partition cluster blocks into rectangular boxes: **a** a partition example for one clustering block, **b** AABB partition results for the clustering shown in Fig. 6b

black boxes in Fig. 9a. The location and shape of the ith STH-coordinate axis-aligned bounding box (AABB) can be represented by the coordinates of the lower-left corner (S_{i1},T_{i1}) and upper-right corner (S_{i2},T_{i2}). Partitioning cluster blocks shown in Fig. 6b into boxes results in the friction map shown in Fig. 9b. In this simulation, 104 boxes are generated from the original 300 k friction cells. It is interesting to note that, under the ST coordinates, nearly all the friction boxes are intrinsically axis-aligned bounding boxes whose width is always aligned to the S axis, and the height to the T axis. Many spatial searching tasks thereby become trivial with such an axis alignment feature. For example, one can find which box a point (S_q,T_q) on the road belongs to via just comparing its coordinates with all boxes' corner locations, respectively. This is useful for vehicles to conduct spatial friction queries very rapidly even without high computational loads.

With the grouping of friction into partitions considering the mean value and confidence interval, the road friction map can be represented as a set of axis-aligned bounding boxes with associated friction coefficient values and confidence intervals, as shown in Fig. 10. Queries of friction boxes can be managed efficiently through the spatial database. In this study, the PostgreSQL database with a spatial database extender PostGIS (PostGIS Project Steering Committee 2021) is employed to manage data. The database can create a spatial index, a generalized search tree (GiST) (Hellerstein et al. 1995), for the box object to facilitate various spatial queries. A common spatial query is to find the nearest box given a point. Furthermore, the resulting database requires only about 260 KB of space to store all 880 friction boxes. With this compact representation, the friction map data can easily be shared with the connected vehicles, as shown in Fig. 1, with less data cost and time delay. With the friction map, the vehicle can plan the driving style according to the available friction (Gao et al. 2021a, b). For the friction scenario in Fig. 2b,

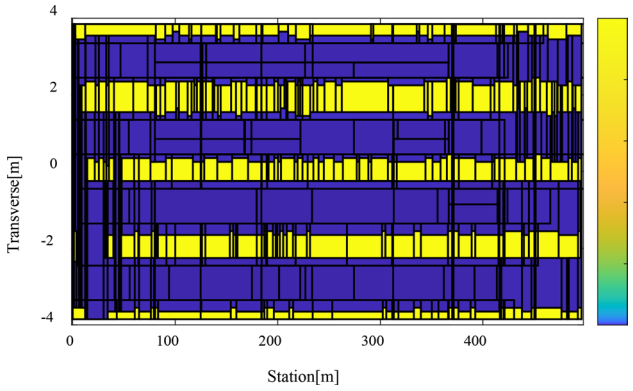

Fig. 10 AABB partition results for the clustering shown in Fig. 8b. Each friction box approximates regions with the same mean value and confidence interval on friction. Black lines are the edges of the friction 880 boxes

the vehicle may need to slow down when approaching the snowy bridge and avoid lane changing for driving stability.

The results indicate that the road friction distribution can be measured effectively by collecting and aggregating the friction data from CAVs, with compact data sharing defined by the axis-aligned rectangular friction boxes. This road friction mapping strategy provides great potential for improving CAVs' control performance and stability via database-mediated feedback systems, with a cloud-based data sharing method that is suited for real-time deployment in actual networks of CAVs.

Analysis and Discussion

In this section, the location accuracy of mapped friction boundaries and several variations of the proposed map generation method are investigated.

Accuracy Analysis of the Location of Critical Friction Boundaries

In "Friction grids cluster considering the confidence interval", the accuracy of the friction value representation was investigated. However, it is equally important to investigate the accuracy of the location of friction boundaries, especially the critical friction boundaries. In this paper, the critical friction boundary is defined as the location, where the friction changes abruptly (a variation of friction coefficient by larger than 0.1), such as the edges of snow ruts and icy bridges. When maneuvering through these boundaries at high speed, vehicles are highly likely to spin out or skid unexpectedly (Beal 2017; Hebden et al. 2004). Thus, it is important that the critical friction boundaries are detected in the correct location to provide an accurate friction preview.

Figure 11a illustrates the ground truth critical boundaries for the true friction distribution shown in Fig. 2d. The detected critical boundaries are shown in Fig. 11b.

To evaluate the accuracy of the detected friction boundaries quantitively, three metrics were calculated based on the difference between the detected boundaries and the ground truth boundaries; RMSE, mean error, and MAE. The RMSE is 4.4 mm, the mean error is 0.19 mm, and the MAE is 0.2 m. This error level is comparable to the positioning error of a differential global positioning system (DGPS) (Karlsson and Mohammadiha 2018; Specht 2020) which provides improved location accuracy by enhancing the normal global navigation satellite system. In addition, this error level is smaller than the grid size used in this study (10 cm by 10 cm square), implying good detection accuracy. This error analysis also provides a way to evaluate the uncertainty of the friction distribution estimation. For the implementation of the estimated friction map, such as optimization-based path tracking control with friction preview (Gao et al. 2022) and path planning in response to the road-tire grip capability, both the uncertainty of the boundary location and the friction values could be considered to improve the system's robustness against that uncertainty. However, due to the non-differentiability and even discontinuity of this friction map at the boundaries, gradient-based solvers may be invalid, but derivative-free solvers can be utilized, such as mesh adaptive direct search (MADS) (Ravera et al. 2019) and Derivative-free Trust-region method (DFTRM) (Dæhlen et al. 2014).

Granularity, Data Size, and Mapping Accuracy

Recall "Cluster the friction based on the road grids", where we utilize the K-Means algorithm to partition the friction grids and the K is estimated via histogram statistics with 0.1 bin width. If one conducts the same clustering and friction boxes partition method but uses different bin widths and thereby different K values, the number of friction boxes and aggregation error between clustered friction map and the true friction distribution will change. This test was performed and the results are summarized in Table 1. The results reveal the trade-off between the partition granularity, i.e., the number of friction boxes, and mapping accuracy. This trade-off can be tuned through the choice of cluster number K. A relatively larger cluster number, K, results in a higher mapping accuracy but also increases the number of friction boxes, data aggregation times, and database size. However, if the cluster number is too large, the accuracy may deteriorate even with more clusters. As shown in Table 1, when the cluster friction interval is smaller than 0.03, which is approximately the three standard deviations of friction measurement noise 0.011, the errors of the generated friction map increase instead of decrease. This is because when the cluster interval is small, the cluster may be formed due to

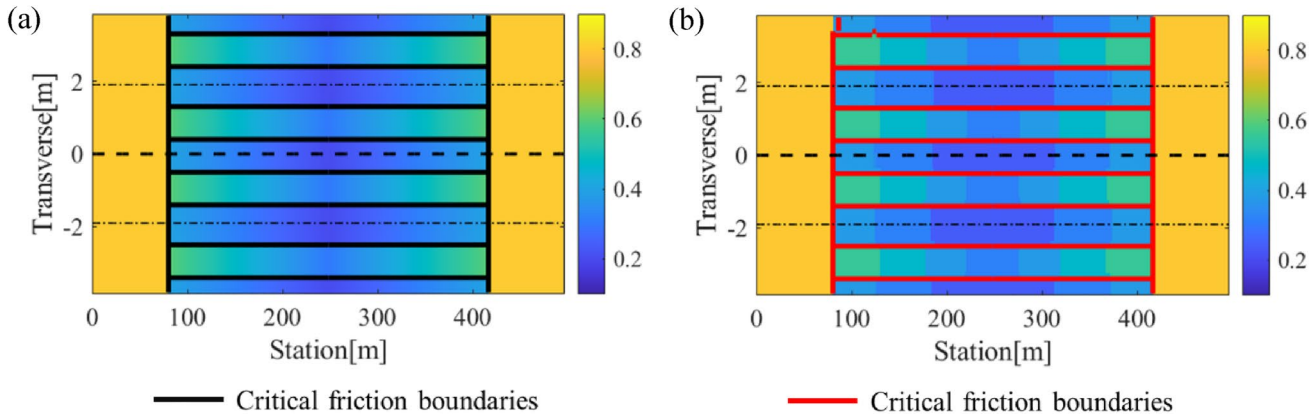

Fig. 11 Critical boundaries: **a** ground truth critical boundaries, **b** detected critical boundaries

the variation of noise value rather than the true value. Given that at least 99.7% of measurement noise values lie within three standard deviations for Gaussian noise, even with data averaging in measurement grids, the clustering of noise values into too many partitions can lead to inaccuracies in the generated friction map.

For all cases examined, the percentage reduction in data size exceeded 99.5%, indicating a higher level of compactness and efficiency in the boxes-based representation compared to grid-based road friction maps. Conducting the data aggregation process in "Friction Data Aggregation" to the true friction data shown in Fig. 2, where no measurement noise is involved, only 47 friction boxes are generated, whereas 104 friction boxes are generated when aggregating the noisy measured friction data from vehicles. Comparing the results, we find that the additional boxes are mainly located in the uneven transition boundaries resulting from the data noise between friction blocks.

To further evaluate the mapping performance, we conduct the friction mapping process for the same road section shown in Fig. 2 but with two extreme friction distribution scenarios. The first scenario assumes a trivial case, where the true friction coefficient at each friction grid is identically 0.65, meaning there is no spatial variation in the friction distribution, and thus only one friction partition would be needed in theory. The mapping results of this scenario are presented in Table 2. A comparison with the snowy bridge mapping results in Table 1 shows that the mapping accuracy is much higher, and the number of friction boxes is much lower for this trivial friction pattern. This implies a more compact and accurate data representation for simple road scenarios, such as uniform dry or wet roads.

In the second scenario, the true friction coefficient at each grid has a value drawn from the same normal distribution with 0.65 mean and 0.1 standard deviations, resulting in a friction distribution with a large value and spatial variance. The data aggregation outcomes are listed in Table 3. In comparison with the results in Table 1, the mapping accuracy is lower and the number of friction boxes is higher for this friction pattern. There are two main reasons: first, most of the measurements occur at the locations of the driving lane as indicated in Fig. 4a and as a result, the location without measurement but with large variance degrades the overall mapping accuracy. Second, the large spatial variance, namely, each grid, tends to have a significantly different friction coefficient value from the neighbor, and consequently, much more friction boxes are required to map the variation. However, even for this

Table 1 Friction mapping performance with various interval values for the snowy bridge

Interval	K	Number of friction boxes	Data size reduction percentage (%)	Number of friction blocks	RMSE	RMSPE (%)	MAE
0.01	40	751	99.6716	223	0.0164	4.24	0.0115
0.02	22	392	99.8317	175	0.0142	3.80	0.0092
0.03	15	285	99.9371	136	0.00761	2.24	0.0055
0.05	10	177	99.9614	91	0.0114	3.22	0.0080
0.1	6	104	99.9771	55	0.0180	5.36	0.0127
0.15	5	85	99.9812	47	0.0234	6.61	0.0165
0.25	4	67	99.9852	30	0.0302	8.90	0.0213
0.3	3	54	99.9881	15	0.0467	15.67	0.0330

Table 2 Friction mapping performance for a road segment with constant friction distribution

Interval	*K*	Number of friction boxes	Data size reduction percentage (%)	Number of friction blocks	RMSE*10^6	RMSPE*10^4 (%)	MAE*10^6
0.01	3	152	99.9665	3	25.3	34.1	24.99
0.02	2	30	99.9934	2	18.8	25.2	17.87
0.03	1	1	99.9998	1	5.84	7.79	5.841
0.05	1	1	99.9998	1	5.84	7.79	5.841
0.1	1	1	99.9998	1	5.84	7.79	5.841
0.3	1	1	99.9998	1	5.84	7.79	5.841

Table 3 Friction mapping performance for a road segment with normal friction distribution

Interval	K	Number of friction boxes	Data size reduction percentage (%)	Number of friction blocks	RMSE	RMSPE	MAE
0.01	26	204,321	55.0021	93,500	0.0763	12.7777	0.0487
0.02	15	171,120	62.3140	89,899	0.0760	12.7608	0.0483
0.03	11	149,662	67.0397	78,129	0.0759	12.7500	0.0482
0.05	8	130,552	71.2484	63,171	0.0762	12.8392	0.0495
0.1	4	93,640	79.3775	28,621	0.0776	13.1689	0.0536
0.3	2	44,955	90.0995	8564	0.0836	14.3829	0.0628

extreme case, which closely resembles a real-world gravel road surface but is rare for a highway, the friction map has considerable compactness and accuracy with an interval of 0.1. Moreover, the driving lane with high mapping estimation confidence as shown, for example, Fig. 5b could provide a drivable route with good friction knowledge on the road for vehicle navigation.

From all the mapping scenarios in this work, it is worth summarizing that the errors of the generated friction map increase instead of decrease when the cluster friction interval is smaller than 0.03, approximately three standard deviations of friction measurement noise (0.011). This finding provides valuable guidance for selecting an appropriate cluster interval in the data aggregation process: cluster intervals should not be smaller than approximately three times the measurement noise.

Mapping Accuracy and the Number of Measurements

In "Friction Data Acquisition Simulation", the friction measurement from the traffic volume with 1038 vehicles in 30 min is collected as the raw data for the road friction mapping. The same friction map generation process elaborated in "Friction Data Aggregation" for the snowy bridge was also conducted for various traffic volumes. The results, which are summarized in Table 4, reveal the trade-off between the traffic volume, an indicator of collected raw measurements

Table 4 Friction mapping performance with different traffic volume

Number of vehicles	*K*	Number of friction boxes	RMSE	RMSPE (%)	MAE
1038	6	104	0.0180	5.36	0.0127
779	6	159	0.0199	5.74	0.0129
519	6	217	0.0205	6.06	0.0130
256	6	441	0.0268	7.87	0.0146
128	6	587	0.0314	9.85	0.0166

data size, and mapping accuracy. Larger data sets provide more accuracy at the cost of longer data aggregation time and thereby mapping latency. In addition to the amount of measurement data, the confidence results in Fig. 8 emphasize the importance of data spatial diversity in improving the overall mapping accuracy, especially for the friction pattern with large spatial variance as the second extreme case examined in "Granularity, data size, and mapping accuracy". Thus, both the size and distribution of the raw data play critical roles in the accuracy of the generated friction map.

Polygon Representation of Friction Blocks

In "Represent the friction cluster block with rectangular boxes", the axis-aligned bounding boxes (AABBs) were used to represent the friction cluster block. In fact, it is also possible to define the cell boundaries simply using the cluster

block boundary. Figure 7a provides an example of a friction block which is the region within the bold red polygon boundaries. In this paper, the boundaries of each block are detected using image dilate and erode operation with a disk-shaped morphological structuring element of radius 1. The polygon boundaries can be depicted using the vertices coordinates values in ST coordinates. With this method, fewer polygons are required to represent the friction map, which is shown in Table 1. The drawback of representing friction maps using boundary definitions rather than AABBs is that managing an arbitrary polygon object, including data storage and the spatial query, is much less efficient in a database, especially for polygons with holes. Because most databases are designed to support AABB queries (for example, R-tree structures), rectangular partitions have clear implementation advantages.

Conclusions and Future Work

This paper presents a road friction map generation strategy by aggregating the measured road-tire friction coefficients from a fleet of CAVs. To demonstrate the strategy, a simulation was developed to collect friction measurement data from CAVs. The simulation integrates traffic simulations, vehicle chassis dynamics and on-vehicle friction estimators, and a highway road surface with varying friction, particularly over bridge segments that have strong friction variation due to weather events. Due to the large size of collected raw data, a road-grid-based data aggregation process is introduced to generate a friction map that associates friction values with appropriate spatial context. The results indicate that the road friction distribution can be measured effectively by collecting and aggregating the friction data from CAVs. Moreover, an insight of work is that the friction map can be represented compactly by rectangular axis-aligned bounding boxes defined by a pair of corner coordinates in space and a friction value within the box, as long as STH (curvilinear) coordinates are used.

This road friction mapping strategy provides great potential for improving CAVs' control performance and stability via database-mediated feedback systems, with a cloud-based data sharing method that is suited for real-time deployment in actual networks of CAVs. In addition, the friction information could potentially support road maintenance strategy development, especially in adverse weather conditions, thus mitigating the impact of inclement road conditions on traffic mobility and safety. Such friction information may also benefit human drivers with more assistance.

Future work can readily improve the approach of this work toward data aggregation. Specifically, the identical fine grid size is not necessary for all road sections, and a coarse grid may be suited for non-bridge areas. Thus, an algorithm can be developed that adaptively estimates grid sizes for different road sections and road conditions. Similarly, this study ignored time-varying friction effects (as only 30 min of data is studied), simply because typical road friction variation occurs at much longer time scales (hours). Given time-varying data, there are methods that can be used to determine the appropriate time window scale or forgetting factor to represent the time variation of friction, including, for example, Allan Variance (AVAR) (Sinanaj et al. 2021). One could even develop a friction prediction model by considering the time dynamics of road and weather conditions. Moreover, in this work, we assume that all vehicles measure the same road-tire friction. However, it is worth noting that the measured value of road-tire friction may also depend on the tire's characteristics in addition to the road texture. Therefore, it would be interesting to explore the mapping performance with the data obtained from a fleet of vehicles equipped with different tires. Such exploration can provide insights into the robustness and generalizability of the proposed friction mapping approach in real-world scenarios, where different types of vehicles and tires are used. This can also help to understand the impact of tire characteristics on road-tire friction measurement and improve the accuracy and reliability of the generated friction map. Finally, the integration of on-vehicle data with roadside sensors (Hippi 2010) or in-vehicle camera friction estimation methods (Roychowdhury et al. 2018) holds promise for future work.

Author Contributions LG wrote most of the manuscript and provided all the figures. JM contributed to the entire manuscript and specifically wrote sections 4.3 and 5.1. JM also revised the manuscript and coordinated the submission process. CB and SB conceived many of the ideas and demo code for the project. SB and CB also supervised the project and assisted with the outline and editing. SPM and WB provided the traffic simulation demo used. CC, KJ, HH, and LS provided help with database design and building, data communication, and data classification method. All authors reviewed and provided feedback for the final manuscript.

Funding This material is based upon work supported by the National Science Foundation under grant numbers CNS-1932509, CNS-1931927, CNS-1932138 “CPS: Medium: Collaborative Research: Automated Discovery of Data Validity for Safety–Critical Feedback Control in a Population of Connected Vehicles”.

Availability of Data and Materials Access to the GitHub repository hosting the code used for this publication is available on request to Sean Brennan, sbrennan@psu.edu.

Declarations

Conflict of Interest There are no competing interests.

Ethical Approval Not applicable.

References

Abbous AEl, Samama N (2017) A modeling of GPS error distributions. Eur Navig Conf (ENC). https://doi.org/10.1109/EURONAV.2017.7954200

Acosta M, Kanarachos S, Blundell M, Acosta M, Kanarachos S, Blundell M (2017) Road friction virtual sensing: a review of estimation techniques with emphasis on low excitation approaches. Appl Sci 7(12):1230. https://doi.org/10.3390/app7121230

Aggarwal CC, Reddy CK (2014) DATA clustering: algorithms and applications, 1st edn. Chapman and Hall/CRC

Alhasan A, Nlenanya I, Smadi O, MacKenzie CA (2018) Impact of pavement surface condition on roadway departure crash risk in Iowa. Infrastructures. https://doi.org/10.3390/infrastructures3020014

ASAM (2020) ASAM OpenCRG Homepage. http://www.opencrg.org/

Beal CE (2017) Stabilization of a vehicle traversing a short low-friction road segment. IEEE Conf Control Technol Appl (CCTA) 2017:1898–1903. https://doi.org/10.1109/CCTA.2017.8062733

Beal CE (2019) Rapid road friction estimation using independent left/right steering torque measurements. Veh Syst Dyn. https://doi.org/10.1080/00423114.2019.1580377

Beal CE, Brennan S (2020) Friction detection from stationary steering manoeuvres. Veh Syst Dyn 58(11):1736–1765

Cao H, Song X, Zhao S, Bao S, Huang Z (2017) An optimal model-based trajectory following architecture synthesising the lateral adaptive preview strategy and longitudinal velocity planning for highly automated vehicle. Veh Syst Dyn 55(8):1143–1188. https://doi.org/10.1080/00423114.2017.1305114

Chen S, Saeed TU, Labi S (2017) Impact of road-surface condition on rural highway safety: a multivariate random parameters negative binomial approach. Anal Methods Accid Res 16:75–89. https://doi.org/10.1016/j.amar.2017.09.001

Dæhlen JS, Eikrem GO, Johansen TA (2014) Nonlinear model predictive control using trust-region derivative-free optimization. J Process Control 24(7):1106–1120. https://doi.org/10.1016/j.jprocont.2014.04.011

Falcone P, Borrelli F, Asgari J, Tseng HE, Hrovat D (2007) Predictive active steering control for autonomous vehicle systems. IEEE Trans Control Syst Technol 15(3):566–580. https://doi.org/10.1109/TCST.2007.894653

Gao L, Beal C, Fescenmyer D, Brennan S (2021a) Analytical longitudinal speed planning for CAVs with previewed road geometry and friction constraints. IEEE Conf Intell Transp Syst Proc. https://doi.org/10.1109/ITSC48978.2021.9564602

Gao L, Maddipatla SSP, Beal C, Jerath K, Chen C, Sinanaj L, Haeri H, Brennan S (2021b) A micro-simulation framework for studying CAVs behavior and control utilizing a traffic simulator, chassis simulation, and a shared roadway friction database. Proc Am Control Conf. https://doi.org/10.23919/ACC50511.2021.9483221

Gao L, Beal C, Mitrovich J, Brennan S (2022) Vehicle model predictive trajectory tracking control with curvature and friction preview. Annu IFAC Adv Automot Control Symp. https://doi.org/10.1016/j.ifacol.2022.10.288

Hebden RG, Edwards C, Spurgeon SK (2004) Automotive steering control in a split-μ manoeuvre using an observer-based sliding mode controller. Veh Syst Dyn 41(3):181–202. https://doi.org/10.1076/vesd.41.3.181.26511

Hellerstein JM, Naughton JF, Pfeffer A (1995) Generalized search trees for database systems. In: Proceedings of the 21st International Conference of Very Large Databases VLDB, pp 562–573

Highway AF (2020) How do weather events impact roads? https://ops.fhwa.dot.gov/weather/q1_roadimpact.htm

Hippi M (2010) A statistical forecast model for road surface friction. In: 15th International Road Weather Conference Quebec City, Canada, 5–7 February 2010, February, 5–7

Hussein NA (2016) Assessing the impact of pavement surface condition and geometric characteristics on the performance of signalised intersections (Issue December) [SWINBURNE UNIVERSITY OF TECHNOLOGY]. https://pdfs.semanticscholar.org/02a5/749df14155fbcdfdff68095eec4cba657d7e.pdf

Jain NK, Saini RK, Mittal P (2019) A review on traffic monitoring system techniques. In: Ray K, Sharma TK, Rawat S, Saini RK, Bandyopadhyay A (eds) Soft computing: theories and applications. Springer, pp 569–577

Karlsson E, Mohammadiha N (2018) A statistical GPS error model for autonomous driving. IEEE Intell Veh Symp Proc. https://doi.org/10.1109/IVS.2018.8500422

Martin DP, Schaefer GF (1996) Tire-road friction in winter conditions for accident reconstruction. SAE Tech Pap. https://doi.org/10.4271/960657

Merritt DK, Lyon C, Persaud B (2015) Evaluation of Pavement Safety Performance NO. FHWA-HRT-14-065. In: Federal Highway Administration

Mobile Road Condition Map (2021). https://roadweather.online/

Panahandeh G, Ek E, Mohammadiha N (2017) Road friction estimation for connected vehicles using supervised machine learning. IEEE Intell Veh Symp Proc. https://doi.org/10.1109/IVS.2017.7995885

Pennsylvania Department of Transportation (2021) Traffic data report 2020. https://www.penndot.gov/ProjectAndPrograms/Planning/TrafficInformation/Pages/2020-Pennsylvania-Traffic-Data.aspx

PostGIS Project Steering Committee (2021) Introduction to PostGIS. http://postgis.net/

Pu Z, Liu C, Shi X, Cui Z, Wang Y (2020) Road surface friction prediction using long short-term memory neural network based on historical data. J Intell Transp Syst. https://doi.org/10.1080/15472450.2020.1780922

Ravera A, Oliveri A, Lodi M, Storace M (2019) Embedded linear model predictive control through mesh adaptive direct search algorithm. IEEE Int Conf Electr Circ Syst. https://doi.org/10.1109/ICECS46596.2019.8964821

Robinson D (2021). Icy bridges: taking drivers by surprise. http://icyroadsafety.com/icybridges.shtml

Roychowdhury S, Zhao M, Wallin A, Ohlsson N, Jonasson M (2018) Machine learning models for road surface and friction estimation using front-camera images. Int J Conf Neural Netw (IJCNN). https://doi.org/10.1109/IJCNN.2018.8489188

Santini S, Albarella N, Arricale VM, Brancati R, Sakhnevych A (2021) On-board road friction estimation technique for autonomous driving vehicle-following maneuvers. Appl Sci (switz) 11(5):1–27. https://doi.org/10.3390/app11052197

Sinanaj L, Haeri H, Gao L, Maddipatla SP, Chen C, Jerath K, Beal C, Brennan S (2021) Allan variance-based granulation technique for large temporal databases. Int Jt Conf Knowl Discov Knowl Eng Knowl Manag. https://doi.org/10.5220/0010651500003064

Specht M (2020) Statistical distribution analysis of navigation positioning system errors—issue of the empirical sample size. Sensors (switzerland) 20(24):1–25. https://doi.org/10.3390/s20247144

Teconer (2020) Mobile road and runway weather monitoring with Road Condition Monitor RCM511. https://www.teconer.fi/en/surface-condition-friction-measurements/

Vaisala (2021) Remote Surface State Sensor DSC211. https://www.vaisala.com/en/products/weather-environmental-sensors/remote-surface-state-sensor-dsc211

van Diggelen F (2007) System design & test-gnss accuracy-lies, damn lies, and statistics-this update to a seminal article first published here in 1998 explains how statistical methods can create many different. GPS World 18(1):26–33

Walton RJ (2018) Characterization of road surfaces using high resolution 3D surface scans to develop parameters for predicting tire-surface friction. The Ohio State University

Zhan Y, Li JQ, Yang G, Wang KCP, Yu W (2020) Friction-ResNets: deep residual network architecture for pavement skid resistance evaluation. J Transp Eng Part B 146(3):04020027. https://doi.org/10.1061/jpeodx.0000187