



Risk evaluation of traffic standstills on winter roads using a state space model



Shogo Umeda ^{a,*}, Yosuke. Kawasaki ^b, Masao. Kuwahara ^a, Akira Iihoshi ^c

^a New Industry Creation Hatchery Center (NICHe), Tohoku University, 6-6-10 Aoba, Aramaki-aza Aoba-ku, Sendai 980-8579, Japan

^b Graduate School of Information Sciences, Tohoku University, 6-6-06 Aoba, Aramaki-aza Aoba-ku, Sendai 980-8579, Japan

^c Honda Motor Co., 8-1-1, Tsukaguchi-hommachi, Amagasaki City 661-8661, Japan

ARTICLE INFO

Keywords:

Risk evaluation
State space model
Anomaly event
Probe data
Road performance
Kalman filter

ABSTRACT

A method that evaluates the risk of traffic standstills on winter roads in real time using a state space model is proposed herein. In Japan, large-scale anomaly events such as traffic standstills that cause serious road disturbances occur frequently every year because of heavy snowfall. However, if the risk of anomaly events is known in advance, appropriate preparation and management can be undertaken to prevent such events and/or alleviate their impacts on road traffic. Therefore, this study attempts to evaluate the risk of standstills based on the degraded road performance estimated from probe vehicle speeds using sequential Bayesian filtering in a state space model (SSM). The SSM comprises a system model constructed by learning historical data and a measurement model using several exogenous variables such as snowfall amounts and temperature. The risk of anomaly events is then determined as the deviation of the filtered vehicle speed by the SSM from the statistically feasible speed distribution. The validation is performed by applying the proposed model to 58 traffic-standstill cases in northern Japan, and we confirm that the model successfully evaluates risks at a reasonable level that permits the practical use.

1. Introduction

Japan witnesses large-scale traffic anomaly events that cause serious traffic disturbances in winter every year. This study focuses on the evaluation of the risk of traffic standstills owing to heavy snowfalls. Fig. 1 shows a photograph of a severe standstill that occurred in Fukui in 2018. Once the traffic comes to a standstill, the lane is likely closed for at least a couple of hours and the two-way traffic may be completely shut down or forced to move alternately using a single available lane. According to Japanese government statistics from 2015, more than 500 instances of standstill causing road disturbances over 1640 km and for 1100 h in total were reported.

However, if the risk is known in advance, sufficient preparation and management measures can be undertaken to prevent severe traffic jams and/or alleviate their impacts on road traffic. Therefore, this study attempts to evaluate the risk of standstills based on the degradation of road performance estimated using probe vehicle speeds based on the sequential Bayesian filtering in a state space model (SSM).

The scope of this study is limited to the estimation of the risk and not the event occurrences. This is because the occurrence of a standstill depends not only on the environment (weather and road-surface conditions) but also on individual vehicle and driver

* Corresponding author.

E-mail addresses: umeda@plan.civil.tohoku.ac.jp (S. Umeda), kawasaki-y@plan.civil.tohoku.ac.jp (Yosuke. Kawasaki), kuwahara@plan.civil.tohoku.ac.jp (Masao. Kuwahara), akira_iihoshi@jp.honda (A. Iihoshi).



Fig. 1. Photograph showing vehicles come to a standstill in Fukui, Japan, in 2018.

Source: <https://www.fukuishimbun.co.jp/articles/gallery/415791?ph=1>.

characteristics, such as the tire quality, loading weight, and driving skill. Hence, the evaluation of a standstill occurrence is not feasible without knowledge of such individual characteristics.

In practice, under snowy conditions, road managers have to make decisions regarding implementing special treatments related to traffic regulation and snow removal mainly based on their experience and by referring to weather forecasts and limited sensing data from detectors and industrial television cameras. These sensing data cannot cover the entire road section that the managers are concerned with but are available only at the locations at which the sensors are installed. In addition, the continuous monitoring of sensing data such as camera images is a tedious task for traffic personnel.

Considering these facts, we decide to use probe vehicle data that provide information of individual vehicle speeds every few seconds continuously over a wide area. As probe vehicle data can be acquired without any roadside infrastructure, the area coverage of the risk assessment can be expanded. Further, an automatic real-time analysis of probe data along with weather data will reduce the tedious but simple work.

For the practical application of this method, it should satisfy the following requirements:

1. Data processing and analysis must be performed in real time so that road managers are informed of the risk without any delays.
2. The risk, which is estimated based on the sensing data, should be probabilistically evaluated to consider the uncertainty in sensing data. For example, the probe vehicle speed may depend both on the driver or vehicle characteristics.
3. The model can be used even with missing data. The model parameters have to be calibrated using their historical data, which may not be observed regularly over time and have frequent missing data. For example, probe speed measurements are likely to have data missing, particularly under low-traffic conditions in mountainous regions.

Finally, we clarify that this study does not intend to replace the current practice of risk assessment conducted by road managers; rather, the objective herein is to assist the managers in conducting assessments by providing them with additional useful traffic information.

2. Related work

Among a number of previous studies on anomaly events, we review studies on the impacts of winter weather on both traffic and the risk evaluation. Although considerable studies deal with the detection of anomaly events subsequent to the occurrences of these events (Horiguchi et al. (2013), Kinoshita et al. (2015), Asakura et al. (2015), Takenouchi et al. (2019), Zhao et al. (2019), Boquet et al. (2020), Huang et al. (2020)), these studies are excluded from our review because our study aims to evaluate the risks that are known prior to the occurrence of these events.

2.1. Impacts of winter weather on traffic

Several studies have assessed the impacts of winter weather on road traffic and reveal that snowfall and freeze conditions reduce vehicle speed and traffic capacity and even increase the risk of accidents as briefly explained below. Datla et al. (2013) analyzed the impact of snow and low temperature on the annual average daily traffic based on the multiple regression analysis considering road types and location characteristics. Yasanthi and Mehran (2020) examined the effects of speed on freeways during winters and report that the fluctuation in free-flow speeds in bad weather conditions is significantly smaller than those reported in HCM (2016) and other studies. Lu et al. (2019) found that the saturated flow rate and free-flow speed decreased by approximately 20% and 17%, respectively, during snowfalls. Yu et al. (2015) analyzed the impact of meteorological conditions on the collision risks on mountainous highways using a correlated random parameter Tobit model and found that the visibility and precipitation variables are correlated with the characteristics of geometric design such as steep downslopes and curvature. Seeherman and Liu (2015) performed a regression analysis to show that the interval of snowfall events greatly influenced the frequency of accidents. El-Basyouny et al. (2014) analyzed the effects

of heavy snow or rainfall on accident types using a multivariate model and revealed that the number of crashes increases at lower temperatures and higher snowfall intensity. [Heqimi et al. \(2018\)](#) adopted the spatial interpolation kriging method to analyze the impact of annual snowfall on highway accidents.

As seen above, all previous commonly reviewed studies have reported that snow and low temperatures decrease the speed on highways, reduce traffic-flow rates, and increase the accident risk. Therefore, risk evaluations must be conducted in relation to such variables on winter weather.

2.2. Risk evaluation of traffic anomaly events

In general, two types of methodologies have been proposed for real-time risk evaluation: statistical and machine-learning methods. Previous studies have mostly evaluated traffic accident risks on expressways because local streets bring complications of pedestrian / bicycle traffic, traffic lights, and on-street parking.

Regarding the statistical methods, [Yu and Abdel-Aty \(2013\)](#) and [Wang et al. \(2019\)](#) proposed a method based on support vector machine. However, [Yang et al. \(2018\)](#) argued that the Bayesian dynamic logistic regression model exhibits better performance than the standard logistic regression and the support vector machine. [Sun et al. \(2015\)](#), [Theofilatos \(2017\)](#), and [Zhai et al. \(2020\)](#) also employed similar methods to a Bayesian logistic regression. [Wu et al. \(2018\)](#) analyzed highway accidents from the perspective of weather conditions and estimated the risk of rear-end collisions under foggy conditions using a random parameter logistic model and a negative binomial model.

Machine-learning methods have been also used for the accident risk evaluation. [Ahmed and Abdel-Aty \(2013\)](#) proposed a method based on the stochastic gradient boosting using traffic, weather, and road-geometry data. [Jiang et al. \(2020\)](#) proposed a Long Short Term Memory (LSTM)-based framework using traffic data of different temporal resolutions. Their model exhibited good predictive performance in terms of crash detection and transferability. [Li et al. \(2020\)](#) applied the LSTM-based model to local streets and reported that an LSTM can learn the complex street functions such as traffic flow characteristics, signal timing, and weather conditions.

As briefly summarized above, the Bayesian dynamic logistic regression has been frequently used in the category of the statistical method. On the other hand in the category of machine learning, the LSTM-based models are actively used these days. The LSTM-based model is usually applied to time-series data and can explicitly capture the temporal effects such as the temporal correlation and the daily/weekly periodic sequences. In this regard, the LSTM-based model would be advantageous over the statistical model for the risk evaluation of standstills which clearly depends on the time sequence. However, the machine-learning method is, in general, predictive but not descriptive; the method including the LSTM-based may well predict future events but may not well explain the logical risk mechanism associated with the environment as the statistical method does.

Regarding the applications, although several studies report the impacts of winter weather on traffic, to the best of our knowledge, traffic standstills are severe and rare anomaly events that have not been dealt with in any study as yet. Moreover, both machine learning and statistical methods may have difficulty in being applied owing to insufficient data on such rare events. Also, in previous studies, attempts were made to evaluate the risk of an accident a few minutes prior to its occurrence. However, for traffic management against standstills, the risk evaluation should be performed much earlier because the countermeasures such as snow removal and lane closure over a fairly wide area would require some time.

3. Outline of the study

A risk is generally defined as the product of the probability of the occurrence of an event and the consequence of the event. However, the consequence of the standstill event is assumed constant herein as the evaluation of the exact consequence is difficult depending on individual local characteristics of the vehicle and the driver and also as the exposure level of the environment in mountainous areas with standstills would be similar in terms of the amount of traffic and residence. Therefore, the risk here is evaluated only based on the probability of the occurrence.

The probability of the occurrence of a standstill is evaluated based on the deviation of the current vehicle speed from the statistically feasible speed. The current vehicle speed is the speed at which a typical driver can safely travel in the current surrounding environment, which is defined by several factors such as road geometry, weather, brightness (daytime/night), and traffic condition (flow and density). As these factors principally depend on time and location, the current vehicle speed is estimated for every space and time via sequential Bayesian filtering in the SSM using several sensing data. The SSM is a data assimilation technique for estimating traffic states, which cannot be directly observed, by combining the model estimate with the measurement. The SSM is used because of the following four causes:

(1) The SSM can be used even with missing data, whereas other methods such as autoregressive models dealing with time-series data have difficulties in handling missing data, and (2) the SSM can stochastically estimate the current vehicle speed considering the noises in the model as well as those in the measurements. (3) The SSM does not require training data. Basically, the number of standstill events is not very many compared with other anomaly events such as traffic accidents, and hence models requesting training data may not be suitable for the standstill analysis. (4) The SSM explains the logical mechanism of anomaly event occurrences by explicitly revealing impacts of environmental factors. In order to prevent and/or mitigate the standstill risk in practice, we need not only to predict the risk but also to understand the mechanism of the risk. These properties of the SSM satisfy requirements 2 and 3 stated in Section 1.

On the other hand, the statistically feasible speed is determined based on the sampling statistics. We consider the measurements of probe vehicle speeds as samples extracted from the population speed distribution. The probability distribution of the sample speeds

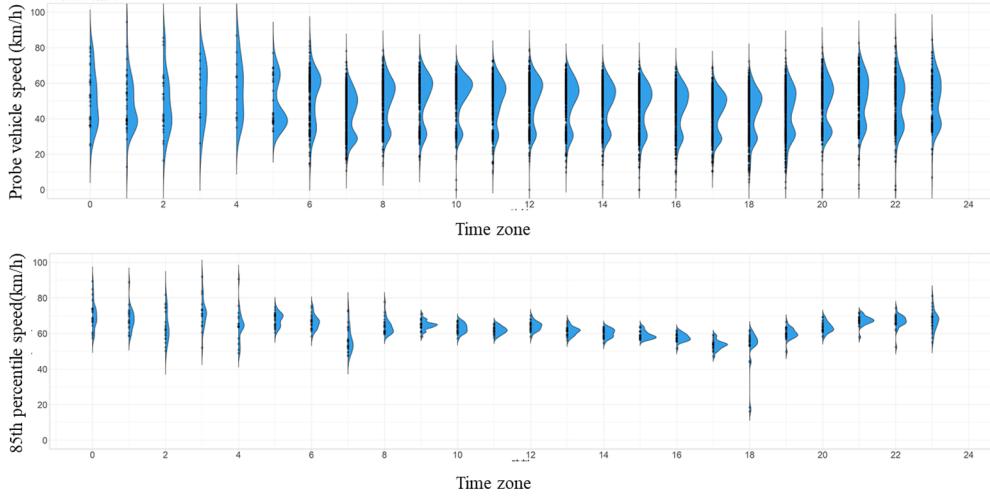


Fig. 2. Distributions of simple speeds and the 85th percentile speeds at a section with traffic lights in Fukui.

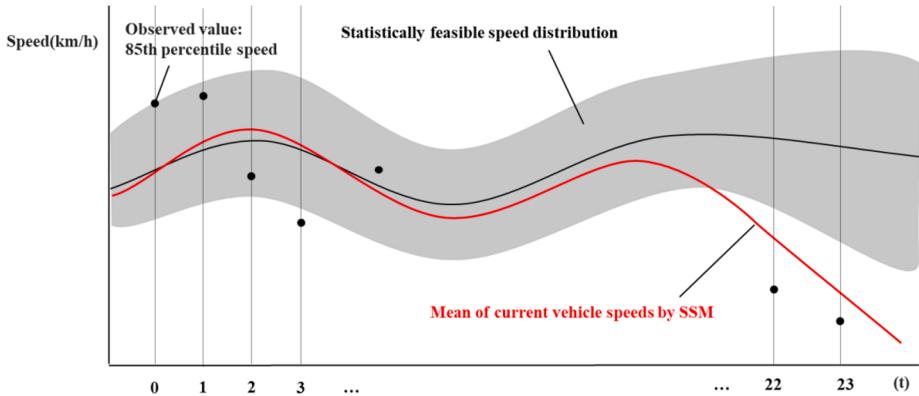


Fig. 3. Conceptual illustration of the proposed method.

can be statistically evaluated.

For both the current vehicle speed and statistically feasible speed, their 85th percentile speeds are used to evaluate their deviation for the risk evaluation as we would like to evaluate the speed reduction purely owing to winter weather environment. However, driving speeds are normally affected by artificial controls (e.g., signalized intersections, roadside facilities, and street parking) and these influences should be eliminated to extract the influences of winter weather only. As the 85th percentile speed is distributed mostly considering vehicles less influenced by the artificial controls, the 85th percentile speed is suitable for the risk evaluation. As seen in the upper half of Fig. 2, which illustrates the speed distributions for different time zones at a road section with traffic lights, the speed distributions tend to exhibit two peaks representing vehicles that are and are not being stopped by traffic lights. On the other hand, the 85th percentile speeds in the lower half of the figure exhibit single peaks less influenced by traffic lights. The distribution of the 85th percentile speed seems hence similar to the Normal distribution and this property justifies to use the Kalman filter algorithm in the SSM which assumes Gaussian noises.

Fig. 3 shows a conceptual illustration of the deviation in the current vehicle speed from the statistically feasible speed at a road section. The gray band represents a range of the statistically feasible speed and the red line represents the mean of the current vehicle speeds estimated using the SSM. When the red line sufficiently deviates from the gray band, the risk alert is activated.

4. Formulation

This section provides the formulations of two speeds: the statistically feasible speed and the current vehicle speed. As explained in Section 3, these two speeds are crucial for evaluating the risk of standstills, and they are defined for a particular location and time. The space and time are discretized into small lengths Δs and Δt , respectively, and s and t represent the discretized intervals: $s \equiv [s, s + \Delta s)$ and $t \equiv [t, t + \Delta t)$, respectively. Although the two speeds are defined for space s and time t , the formulation below focuses on one particular space s , and the symbol s is eliminated for simplicity.

4.1. Statistically feasible speed distribution

Vehicle speed at time t $X(t)$ is a continuous random variable with the probability density function $f_t(x)$ and the cumulative distribution function $F_t(x)$. Although the true functions of $f_t(x)$ and $F_t(x)$ cannot be exactly known, they can be estimated using historical speed data as shown in next Section 5.

As mentioned earlier, measurements of probe vehicle speeds are considered as samples extracted from the population speed distribution, $f_t(x)$. Let us estimate the q -percentile speed at time t , $X_q(t)$, using $n(t)$ samples from the population. Note that for the 85th percentile speed, $q = 0.85$. The $n(t)$ samples are sorted in the ascending order: $x_1, x_2, \dots, x_{n(t)}$, and if the m -th sample speed x_m is the q -percentile speed, $m = n(t)q$. In this case, one sample must be equal to x_m , $n(t)q - 1$ samples must have speeds lower than x_m , and $(n(t) - 1) - (n(t)q - 1) = n(t)(1 - q)$ samples must have speeds higher than x_m . Therefore, the probability density of the q -percentile speed, $f_{X_q(t)}(x_m)$, is written as follows:

$$f_{X_q(t)}(x_m) = n(t)f_t(x_m) \cdot {}_{n(t)-1}C_{n(t)q-1}F_t(x_m)^{n(t)q-1}(1 - F_t(x_m))^{n(t)(1-q)} \quad (1)$$

The integration of the probability density over x_m is equal to 1 as shown below; therefore, Eq. (1) is appropriate for the probability density function of the q -percentile speed, $f_{X_q(t)}(x_m)$.

$$\begin{aligned} \int_0^\infty f_{X_q(t)}(x_m) dx &= n(t) {}_{n(t)-1}C_{m-1} \int_0^\infty F_t(x)^{m-1}(1 - F_t(x))^{n(t)-m} f_t(x) dx \\ &= n(t) {}_{n(t)-1}C_{m-1} \int_0^1 F_t(x)^{m-1}(1 - F_t(x))^{n(t)-m} f_t(x) \frac{dx}{dF_t(x)} dF_t(x) \\ &= \frac{n(t)(n(t) - 1)!}{(m - 1)!(n(t) - m)!} \cdot \frac{(m - 1)!(n(t) - m)!}{n(t)!} = 1, \quad \text{where } m = n(t)q \end{aligned} \quad (2)$$

For more general cases wherein $n(t)q$ is not an integer, the probability density function is derived as explained in the Appendix. Appendix also validates $f_{X_q(t)}(x)$ using the random sampling from one of the sections in Fukui and shows $f_{X_q(t)}(x)$ can be approximated as the Normal distribution. As a whole, given the population speed distributions $f_t(x)$ and $F_t(x)$ at time t , the probability density of the q -percentile speed, $f_{X_q(t)}(x)$, can be evaluated in relation to the number of samples $n(t)$. Finally, from the evaluated probability density $f_{X_q(t)}(x)$, the mean $\mu_{X_q}(t)$ and variance $\sigma_{X_q}^2(t)$ are determined.

4.2. Current vehicle speed using the SSM

The current vehicle speed represents the degraded road performance attributed to heavy snowfalls. It is estimated using sequential Bayesian filtering in the SSM. To meet the requirement of real-time calculations, we employed a Gaussian SSM with a linear system and measurement models and normally distributed noises (Kalman, 1960). The efficient Kalman filter algorithm can be used to determine the filter distribution of the current vehicle speed.

4.2.1. Measurement model

The measurement model of the SSM is basically a linear regression model that explains the current vehicle speed considering several influential factors as shown below.

$$\hat{v}_{85}(t) = \beta_{\text{snow}}(t) \cdot \text{snow}(t) + \beta_{\text{temp}}(t) \cdot \text{temp}(t) + \beta_n(t) \cdot n(t) + \mu(t) + \gamma_1(t) + w(t) \quad (3)$$

where

- $\hat{v}_{85}(t)$ = the observed current 85th percentile speed at time t ,
- $\text{snow}(t)$ = the amount of snowfall at time t ,
- $\text{temp}(t)$ = the temperature at time t ,
- $n(t)$ = the number of probe vehicles at time t ,
- $\beta_{\text{snow}}(t)$, $\beta_{\text{temp}}(t)$ and $\beta_n(t)$ = regression coefficients of $\text{snow}(t)$, $\text{temp}(t)$, and $n(t)$,
- $\mu(t)$ = the level component,
- $\gamma_1(t)$ = the periodic component, and
- $w(t)$ = white noise $\sim (0, \sigma_w^2)$

The observed 85th percentile speed, $\hat{v}_{85}(t)$, is explained not only through measurable exogenous factors such as $\text{snow}(t)$, $\text{temp}(t)$, and $n(t)$ but also through the level and periodic components that represent the intercept of the regression and the periodic variation. In addition, $\hat{v}_{85}(t)$ includes white noise $w(t)$ to consider the stochastic fluctuations. In a conventional linear regression model, parameters of $\beta_{\text{snow}}(t)$, $\beta_{\text{temp}}(t)$, $\beta_n(t)$, $\mu(t)$, and $\gamma_1(t)$ are constants however, they vary over time in this model, and their temporal transitions are formulated in the system model in the next subsection.

For time-series data, four components are generally considered to describe the temporal transition: the level, trend, periodic, and

regression components. However, as $\hat{v}_{85}(t)$ does not explicitly decrease or increase over time, the trend component is eliminated, and the rest of the three components are included in the model. Furthermore, previous studies suggest that road-surface conditions, brightness, and the highway geometric design significantly affect the speeds. To consider the road-surface condition, we used $snow(t)$ and $temp(t)$ as the explanatory variables because they are clearly the factors that determine the road-surface condition and are measurable. Ideally, the snow depth on the road rather than the amount of snowfall would more directly influence the speed. However, the snow depth on the road cannot be accurately observed because the weather stations may not be always sufficiently close to analyze a section and as the snow depth depends on the snow-removal activities conducted by the road managers. The brightness, which depends on time of the day and varies periodically, can be described through the periodic component. Although the geometric design must be significantly influence $\hat{v}_{85}(t)$ in general, the geometric design is not included in the model because the model parameters are calibrated for each of the section. In addition, $\hat{v}_{85}(t)$ clearly depends on the number of samples $n(t)$, as explained later using Fig. 7; that is, the mean value of $\hat{v}_{85}(t)$ increases as $n(t)$ increases.

4.2.2. System model

The parameters in the measurement model are classified into three categories: level, regression, and periodic components. Their time-to-time transitions can be explained using the system model as follows.

Level components:

$$\mu(t) = \mu(t-1) + \varepsilon(t), \quad \varepsilon(t) \sim N(0, \sigma_\varepsilon^2) \quad (4)$$

Regression components:

$$\beta_{snow}(t) = \beta_{snow}(t-1) + \xi(t), \quad \xi(t) \sim N(0, \sigma_\xi^2) \quad (5)$$

$$\beta_{temp.}(t) = \beta_{temp.}(t-1) + \psi(t), \quad \psi(t) \sim N(0, \sigma_\psi^2) \quad (6)$$

$$\beta_n(t) = \beta_n(t-1) + \zeta(t), \quad \zeta(t) \sim N(0, \sigma_\zeta^2) \quad (7)$$

Periodic components:

$$\begin{aligned} \gamma_1(t) &= -\sum_{i=1}^C \gamma_i(t-1) + \omega(t), \quad \omega(t) \sim N(0, \sigma_\omega^2) \\ \gamma_2(t) &= \gamma_1(t-1), \\ &\vdots \\ \gamma_{C-1}(t) &= \gamma_{C-1}(t-1) \\ \gamma_C(t) &= \gamma_{C-1}(t-1) \end{aligned} \quad (8)$$

The level component is described using the parameter $\mu(t)$, which corresponds to the intercept of linear regression and varies over time while receiving white noise $\varepsilon(t)$ every time. The regression component includes three parameters: $\beta_{snow}(t)$, $\beta_{temp.}(t)$, and $\beta_n(t)$, which are the coefficients of the regression: amount of snow $snow(t)$, temperature $temp.(t)$, and the number of probe vehicles $n(t)$. These coefficients also vary over time similar to the level component while receiving white noises $\xi(t)$, $\psi(t)$, and $\zeta(t)$. The periodic component includes $\gamma_i(t)$, $i = 1, 2, \dots, C$ that express the cyclic variation in $\hat{v}_{85}(t)$ with a cyclic period of $C + 1$. For example, to describe the daily cyclic pattern of $\hat{v}_{85}(t)$, $C + 1 = 24 \text{ h} / \Delta t$. Here, $\gamma_i(t)$ represents the periodic factor i intervals before current time t ; that is, $\gamma_i(t)$ represents the influence of $\hat{v}_{85}(t-i)$ on $\hat{v}_{85}(t)$. For instance, with $\Delta t = 1 \text{ h}$, when the current time $t = 7:00$, $\gamma_2(t)$ is the influence of $\hat{v}_{85}(t-2) = \hat{v}_{85}(5:00)$ on $\hat{v}_{85}(t)$. Therefore, when the current time shifts from $t-1$ to t , $\gamma_i(t-1)$ at $t-1$ should shift to $\gamma_{i+1}(t)$, as shown in the second and subsequent lines in Eq. (8). (In the example, the influence of $v_{85}(5:00)$ must be $\gamma_3(8:00)$ because $t-i = 8:00-3:00 = 5:00$.) For the 24-h cycle, basically 23 periodic parameters are required to represent the influences on $\hat{v}_{85}(t)$ in the last 23 h. However, because of the constraint that the sum of the periodic influences is zero, $\gamma_1(t)$ on the first line of Eq. (8) is obtained with white noise $\omega(t)$.

From the system model, Eqs. (4)–(8), $\hat{v}_{85}(t)$ clearly has the temporal dependency. However, $\hat{v}_{85}(t)$ may also have spatial correlation such that $\hat{v}_{85}(t)$ depends on the speeds in adjacent sections. However, owing to light traffic in mountainous road sections and due to the long section length of 500 m, the spatial correlations in $\hat{v}_{85}(t)$ may be negligible in a low-density environment.

The state space model explained above can be concisely described as shown below using vectors and matrices. Here, $x(t)$ is the state vector comprising parameters of the system model. $F(t)$ is the state transition matrix describing the system equations Eqs. (4)–(8). Furthermore, $v(t)$ is the system noise vector having a mean of 0 and covariance matrix Σ . Although Σ is time independent, noise $v(t)$ is time dependent because the noises change with time according to the fixed elements of Σ . Also, $y(t)$ is the measurable 85th percentile speed $\hat{v}_{85}(t)$, and $H(t)$ is the observation matrix describing the measurement model with noise $w(t)$.

System model:

$$x(t) = F(t)x(t-1) + v(t), \quad v(t) \sim N(\mathbf{0}, \Sigma) \quad (9)$$

$$x(0) \sim N(\mathbf{0}, P(0))$$

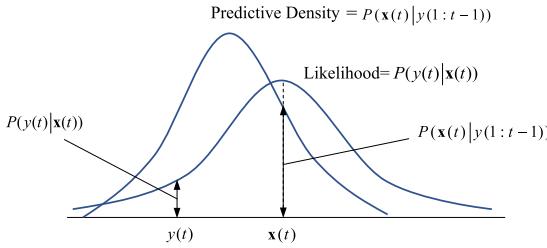


Fig. 4. Schematic illustration of the filter density.

Measurement model:

$$y(t) = \mathbf{H}(t)\mathbf{x}(t) + w(t), \quad w(t) \sim N(0, \sigma_w^2) \quad (10)$$

where

$$\mathbf{x}(t) = \text{state vector} = \begin{bmatrix} \mu(t) \\ \beta_{\text{snow}}(t) \\ \beta_{\text{temp.}}(t) \\ \beta_n(t) \\ \gamma_1(t) \\ \gamma_2(t) \\ \gamma_3(t) \\ \vdots \\ \gamma_C(t) \end{bmatrix}, \quad \mathbf{F}(t) = \begin{bmatrix} 1 & \mathbf{0} & \mathbf{0} \\ 1 & 0 & 0 \\ \mathbf{0} & 0 & 1 \\ 0 & 0 & 1 \\ \hline \mathbf{0} & -1 & -1 & -1 & \cdots & -1 \\ \mathbf{0} & 1 & 0 & 0 & \cdots & 0 \\ \mathbf{0} & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ \mathbf{0} & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, \quad \mathbf{v}(t) = \begin{bmatrix} \varepsilon(t) \\ \xi(t) \\ \psi(t) \\ \zeta(t) \\ \omega(t) \end{bmatrix}$$

$$y(t) = \hat{v}_{85}(t)$$

$$\mathbf{H}(t) = [1 \quad | \quad \text{snow}(t) \quad \text{temp}(t) \quad n(t) \quad | \quad 1 \quad 0 \quad 0 \quad \cdots \quad 0], \quad w(t) \sim N(0, \sigma_w^2)$$

$$\Sigma = \begin{bmatrix} \sigma_\varepsilon^2 & 0 & 0 & 0 & 0 \\ 0 & \sigma_\xi^2 & 0 & 0 & 0 \\ 0 & 0 & \sigma_\psi^2 & 0 & 0 \\ 0 & 0 & 0 & \sigma_\zeta^2 & 0 \\ 0 & 0 & 0 & 0 & \sigma_\omega^2 \end{bmatrix}$$

is the variance-covariance matrix of system noise Σ .

The SSM generally has parameters of initial state $\mathbf{x}(0)$ and a variance-covariance matrix $\mathbf{P}(0)$ as well as the variances of system noise Σ and measurement noise σ_w^2 . These parameters must be estimated using the standard maximum likelihood method in advance. However, it is known that the outputs of the Kalman filter are independent of the initial state $\mathbf{x}(0)$ if the elements of variance-covariance matrix $\mathbf{P}(0)$ are sufficiently large (Jong (1991)). Therefore, we set a large $\mathbf{P}(0)$ as recommended by Jong (1991). The Kalman filter estimates the filter distribution of the state vector comprising the variables in the system model Eqs. (4)–(8). The estimate of the mean current vehicle speed using the SSM, $\mu_{v_{85}}(t)$, is then evaluated by introducing the filtered state vector into the measurement model.

4.2.3. Predictive density and filter density

Predictive density at time t is defined as $P(\mathbf{x}(t)|y(1:t-1))$, which is the probability of $\mathbf{x}(t)$ given measurements $y(1:t-1)$; that is, measurements from time 1 to $t-1$. This density is called predictive because it uses only the measurements until $t-1$ (does not include the measurement at time t). It is written in terms of $\mathbf{x}(t-1)$ as shown below.

$$P(\mathbf{x}(t)|y(1:t-1)) = \int P(\mathbf{x}(t)|\mathbf{x}(t-1)) \cdot P(\mathbf{x}(t-1)|y(1:t-1)) \cdot d\mathbf{x}(t-1) \quad (11)$$

The first term on the right side of Eq. (11), $P(\mathbf{x}(t)|\mathbf{x}(t-1))$, is evaluated based on system model (9). For instance, if a particular value of $\mathbf{x}(t-1)$ is plugged into Eq. (9), $\mathbf{x}(t)$ can be determined. The second term, $P(\mathbf{x}(t-1)|y(1:t-1))$, is called the filter density at time $t-1$, and is explained in the next paragraph. However, if the filter density is known, the left hand side of Eq. (11) is evaluated by integrating $P(\mathbf{x}(t)|\mathbf{x}(t-1)) \cdot P(\mathbf{x}(t-1)|y(1:t-1))$ over $\mathbf{x}(t-1)$.

The filter density at time t is the posterior probability of $\mathbf{x}(t)$ considering measurements y until time t , $y(1:t)$, and is defined as follows:



Fig. 5. Standstills considered for model validation.

$$P(\mathbf{x}(t)|y(1:t)) = P(\mathbf{x}(t)|y(t), y(1:t-1)) = \frac{P(\mathbf{x}(t), y(t)|y(1:t-1))}{P(y(t)|y(1:t-1))} = \frac{P(y(t)|\mathbf{x}(t), y(1:t-1))P(\mathbf{x}(t)|y(1:t-1))}{P(y(t)|y(1:t-1))} \quad (12)$$

As $y(t)$ depends only on $\mathbf{x}(t)$ and is independent of $y(1:t-1)$ based on the Markov property (Eq. (10)), $P(y(t)|\mathbf{x}(t), y(1:t-1))$ is equal to $P(y(t)|\mathbf{x}(t))$. Moreover, the denominator does not depend on $\mathbf{x}(t)$ but only on the measurements. Therefore, the filter density is proportional to the nominator:

$$P(\mathbf{x}(t)|y(1:t)) \propto P(y(t)|\mathbf{x}(t)) \cdot P(\mathbf{x}(t)|y(1:t-1)) \quad (13)$$

The first part on the right side of Eq. (13), $P(y(t)|\mathbf{x}(t))$, is called the likelihood, which is the probability of measurement $y(t)$ given $\mathbf{x}(t)$ and which is evaluated from the measurement model Eq. (10) with the known noise density $w(t)$.

Fig. 4 is a schematic illustration of the filter density. Suppose that predictive density $P(\mathbf{x}(t)|y(1:t-1))$ is known for all $\mathbf{x}(t)$, and measurement $y(t)$ is obtained at time t . For a particular value of $\mathbf{x}(t)$, the predictive density $P(\mathbf{x}(t)|y(1:t-1))$ is evaluated as shown. The likelihood is also drawn based on the given noise density of $w(t)$, and probability $P(y(t)|\mathbf{x}(t))$ is determined for $\mathbf{x}(t)$. The filter density is the product of likelihood $P(y(t)|\mathbf{x}(t))$ and predictive density $P(\mathbf{x}(t)|y(1:t-1))$. By repeatedly performing these calculations for different values of $\mathbf{x}(t)$, the filter density $P(\mathbf{x}(t)|y(1:t))$ is evaluated for all $\mathbf{x}(t)$.

Based on the above structure, the predictive and filter densities can be alternately evaluated. First, assuming the initial values of filter density at time $t=0$, $P(\mathbf{x}(0)|y(0))$, the predictive density at time $t=1$, $P(\mathbf{x}(1)|y(0))$, is evaluated from Eq. (11). Then, the filter density at time $t=1$, $P(\mathbf{x}(1)|y(1:1))$, is evaluated from Eq. (13), and so on. Thus, the predictive and filter densities are alternately determined from the initial condition at $t=0$.

The SSM repeats the prediction and the filtering for every time increment. Even if no measurement $y(t)$ is available at any instance, the same filtering procedure can be conducted using the likelihood $P(y(t)|\mathbf{x}(t))$ equal to 1.0. This is the main advantage of the SSM over other time-series analyses wherein missing data are not normally acceptable and some data must be created to fill these missing slots.

4.3. Standstill risk index

Thus far, we have explained the estimation of the statistically feasible speed distribution and current vehicle speed distribution. The risk of standstills is principally evaluated based on the difference between two probability distributions. The Kullback–Leibler divergence (KLD) is an index that evaluates the difference between distributions. However, it evaluates the difference between the shapes of two distributions and will likely have a large value (implying less similarity) when their variances are considerably different even if their means are closely equal. Hence, it is not suitable for our purpose as we aim to identify the extent to which the current vehicle speed $\mu_{v_{85}}(t)$ decreases relative to the statistically feasible speed distribution.

Therefore, the following $SRI(t)$ (Standstill Risk Index) is proposed instead.

$$SRI(t) = \int_0^\infty \frac{v - \mu_{X_{85}}(t)}{\sigma_{X_{85}}^2(t)} \cdot f_{v_{85}(t)}(v) dv = \frac{\mu_{v_{85}}(t) - \mu_{X_{85}}(t)}{\sigma_{X_{85}}^2(t)}, \quad (14)$$

where

- $\mu_{X_{85}}(t)$ = the mean of statistically feasible 85th percentile speed $X_{85}(t)$,
- $\sigma_{X_{85}}^2(t)$ = the variance of statistically feasible 85th percentile speed $X_{85}(t)$,
- $\mu_{v_{85}}(t)$ = the mean of current 85th percentile vehicle speed $v_{85}(t)$,
- $\sigma_{v_{85}}^2(t)$ = the variance of current 85th percentile vehicle speed $v_{85}(t)$,
- $f_{v_{85}(t)}(v)$ = the probability density of $v_{85}(t) \sim N(\mu_{v_{85}}(t), \sigma_{v_{85}}^2(t))$.

Table 1
Traffic regulations immediately after standstills.

	Regulations	Number of cases
1	Complete lane closure	14 cases
2	One-side alternate traffic	15 cases
3	Partial lane closure	29 cases

The $SRI(t)$ is the speed deviation between the current vehicle speed $v_{85}(t)$ and the mean of statistically feasible speed, $\mu_{X_{85}}(t)$, relative to its standard deviation $\sigma_{X_{85}}(t)$, but not the probability of deviation. The probability of deviation could be evaluated as $\int_0^\infty F_{X_{85}(t)}(v) \cdot f_{v_{85}(t)}(v) dv$, where $F_{X_{85}(t)}(v)$ is the cumulative distribution function of the statistically feasible speed: $F_{X_{85}(t)}(v) = \int_0^v f_{X_{85}(t)}(x) dx$. And, the risk alert could be activated when the probability of deviation exceeds the specified common threshold value. However, since the distributions of $F_{X_{85}(t)}(v)$ and $f_{v_{85}(t)}(v)$, determined based on the limited number of probe samples, include uncertainties, this study decides to use the speed deviation for the risk index $SRI(t)$ and the threshold value of $SRI(t)$ for the risk activation is determined separately in each application field.

5. Evaluation of model performance

5.1. Application fields and periods

The proposed model was applied to 58 standstill events that occurred at 53 sites in northern Japan during the winter of 2018 and 2019. The red symbols in Fig. 5 indicate the site locations. Table 1 lists the traffic regulations introduced immediately after the 58 standstills. Standstills often caused severe effects such as complete or partial lane closure for a fairly long time.

To calibrate model parameters, sensing data of snowfalls, temperature, number of probe vehicles, and probe speeds should be collected for sufficiently long periods. The collected sensing data are divided into two parts: one for parameter calibration and another for model validation. For the standstills that occurred in 2019, the model parameters were calibrated using the 25-day data in 2018 and the validation was conducted using the 40-day data in 2019. For the standstill that occurred in 2018, the calibration was conducted using the 40-day data in 2019 and the validation was conducted using the 25-day data in 2018.

As mentioned in section 4, the space, time, and speed are discretized into small lengths of $\Delta s = 500$ m, $\Delta t = 1$ h, and $\Delta x = 1$ km/h respectively. Statistically feasible speed $X_{85}(t)$ and the mean of current vehicle speed obtained by the SSM $\mu_{v_{85}}(t)$, $t = 1, 2, \dots, T$ are estimated for each of the discrete space and time. $\Delta s = 500$ m was selected for the following reasons:

- (1) The finest spatial resolution for weather data is 500 m.
- (2) Japanese regional mesh system has a mesh size of 500 m; therefore, this selection is convenient for combining several other related data.
- (3) As the number of probe vehicles being driven in a mountainous region is generally low, and the data uplink interval is a few seconds, a shorter Δs will make it difficult to acquire sufficient data.
- (4) Road managers tend to manage road maintenance and management with almost the same spatial resolution as $\Delta s = 500$ m as for larger values of Δs , a road section may not be homogeneous in terms of the geometric design and land use pattern.

$\Delta t = 1$ h was set owing to the following reasons:

- (1) Weather data such as snowfall and temperature data are normally aggregated in 1-h periods, and the forecast is provided with the same temporal resolution.
- (2) As few probe vehicles are operated in the region, $\Delta t \geq 1$ h is necessary to reasonably estimate the 85th percentile measurement, $\hat{v}_{85}(t)$, and also to reduce the number of time slots with missing data.

5.2. Parameter calibration using probe and weather data

5.2.1. Estimation of statistically feasible speed

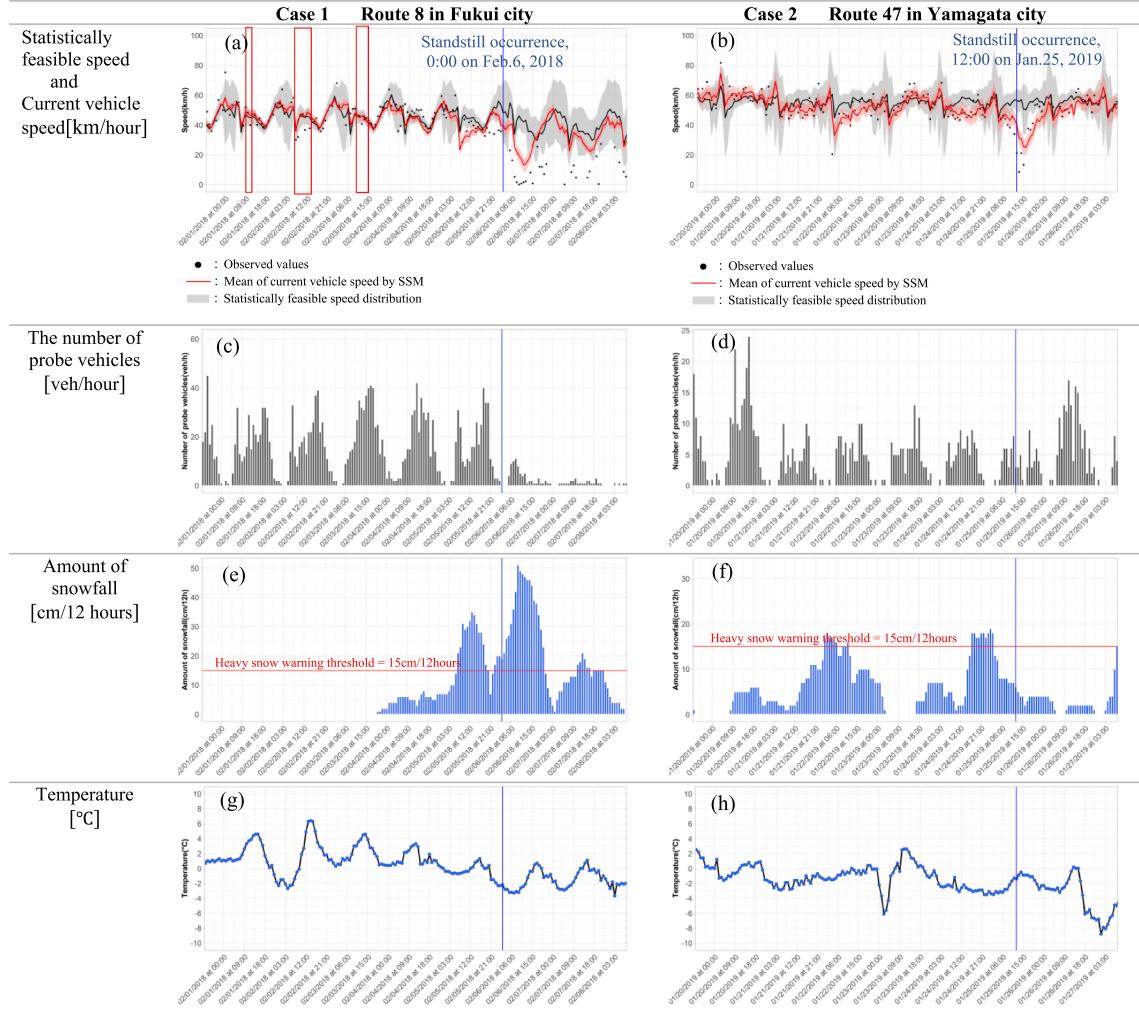
To estimate the probability distribution of $X_{85}(t)$, we must estimate the probability density function $f_t(x)$ of the population speeds. When calculating the speed at each time of day, we assumed the same population density. A set of times h in one day is denoted as Ω_h (the times of h are also discretized as $\Delta h = 1$ h). If time t belongs to set Ω_h , $f_t(x)$ is estimated from the historical speeds at all times in Ω_h . Note that at different times t and t' in Ω_h , although their population speed distributions $f_t(x)$ and $f_{t'}(x)$ are assumed to be identical, the number of speed samples $n(t)$ and $n(t')$ are different.

As the speeds were probed every few seconds, the speed of one vehicle was observed several times along a given 500-m section during time t , providing multiple speeds. These speeds were averaged to obtain the speed of the vehicle in that section at time t . Then, the probability density function $f_t(x)$, $t \in \Omega_h$ was estimated based on the fraction of average vehicle speeds along the section at all times in Ω_h . From Eq. (1), the probability density of the 85th percentile speed, $f_{X_{85}(t)}(x)$, was evaluated using $f_t(x)$ and the number of speed samples $n(t)$, and the mean $\mu_{X_q}(t)$ and variance $\sigma_{X_q}^2(t)$ of $X_{85}(t)$ were also determined.

Table 2

Details of the standstill cases that occurred in Fukui and Yamagata.

Case	Standstill occurrence route	Standstill occurrence time	Regulation time	Regulation type
1	Route 8 in Fukui city	February 6, 2018 at 0:00	66 h	Road closure
2	Route 47 in Yamagata city	January 25, 2019 at 12:00	3.5 h	Road closure

**Fig. 6.** Results of analysis of cases 1 and 2 (standstill occurrence points).

5.2.2. Current vehicle speed

For the mean current vehicle speed obtained from the SSM $\mu_{85}(t)$, the SSM parameters should be calibrated to estimate the variances in the observation (σ_w^2) and system noises (σ_ξ^2 , σ_ψ^2 , σ_ζ^2 and σ_o^2) through the standard maximum likelihood method using the sensing data of $snow(t)$, $temp(t)$, $n(t)$, and $\hat{v}_{85}(t)$. For $snow(t)$ and $temp(t)$, the data pertaining to the amount of snowfall and temperature are downloaded from the website of Japan Meteorological Agency under the Ministry of Land, Infrastructure, Transport and Tourism (MLIT). For the number of probe samples $n(t)$ and $\hat{v}_{85}(t)$, probe vehicle data are obtained from Honda Motor Co. Ltd.; these data include vehicle ID, data acquisition time, vehicle position (latitude, longitude), and vehicle speed every few seconds. As our study areas are normally mountainous areas with heavy snowfalls, the number of probe vehicles may not be always sufficient, and there could be instances wherein no probe data are available. However, as mentioned earlier, the SSM can be implemented even with some missing data.

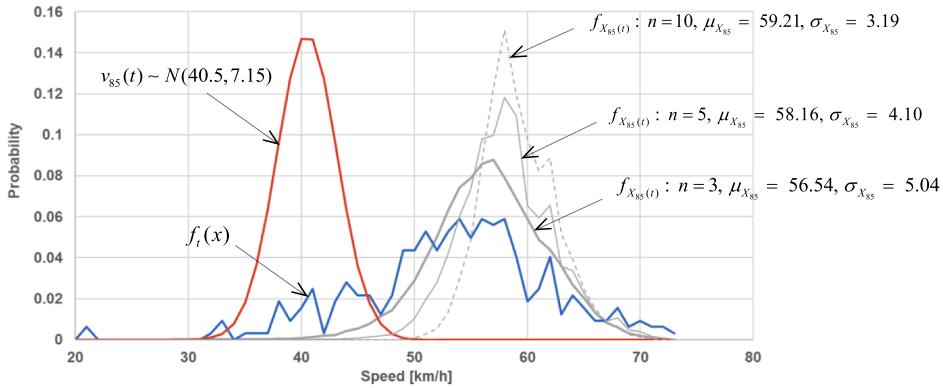


Fig. 7. Statistically feasible speed and current vehicle speed distributions at standstill section in Yamagata (Case 2) at 12:00 on January 25, 2019.

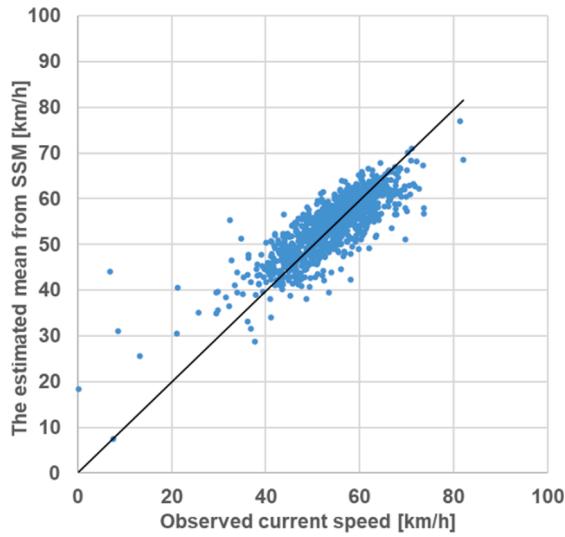


Fig. 8. Relationship between observed current speed $\hat{v}_{85}(t)$ and the estimated mean $\mu_{v_{85}}(t)$ from SSM at the standstill section.

5.3. Examples of speed distributions

In this section, the estimation of the statistically feasible and current vehicle speeds is explained in detail using the examples of two field applications in Fukui and Yamagata. Table 2 provides information regarding these field applications. Case 1 refers to a large-scale standstill that occurred at 12:00 on February 6, 2018, in Fukui, where 1500 vehicles remained stationary for 66 h on Route 8 in the center of Fukui city, and Case 2 refers to a standstill at 12:00 on January 25, 2019, in Yamagata, where the impact on traffic flow was less than that in Case 1; however, a number of cars were stuck and caused the approximately 2-km long traffic jam on Route 47 in Tazawa City, Yamagata Prefecture.

Focusing on the gridlocked road sections in Fukui (Case 1) and Yamagata (Case 2), the time-dependent profiles of the statistically feasible speeds were estimated using Eq. (1) and are shown in Fig. 6(a) and (b), respectively. Here, the black line represents the mean $\mu_{X_{85}}(t)$ and the gray band delineates the range, $\mu_{X_{85}}(t) \pm \sigma_{X_{85}}(t)$. The red line shows the mean of the current vehicle speed, $\mu_{v_{85}}(t)$, evaluated by the SSM using the observed 85th percentile speeds (black dots), and the red band delineates the range, $\mu_{v_{85}}(t) \pm \sigma_{v_{85}}(t)$. Panels (c), (e) and (g) of Fig. 6 display the number of probe samples, amount of snowfall, and temperature, respectively, in case (1); the corresponding results in case (2) are shown in panels (d), (f) and (h) of Fig. 6. During the standstill in Fukui, the amount of snowfall had drastically increased and the temperature decreased. In this circumstance, the current vehicle speed $\mu_{v_{85}}(t)$ sharply decreased and deviated from the statistically feasible speed $\mu_{X_{85}}(t)$ before the standstill. On the other hand, the standstill in Yamagata occurred when the snowfall amount was decreasing and the temperature was increasing while remaining below zero. However, the current vehicle speed $\mu_{v_{85}}(t)$ deviated increasingly from $\mu_{X_{85}}(t)$.

We now examine the distributions of $X_{85}(t)$ and $v_{85}(t)$ in more detail. Fig. 7 shows the statistically feasible speed distribution $f_{X_{85}(t)}(x)$ (gray lines) and the population speed distribution $f_t(x)$ (blue line) in the standstill in Yamagata (Case 2). The $f_{X_{85}(t)}(x)$ was

Table 3
Confusion matrix.

Predicted	Actual	
	Risky	Safe
Risky	True positive; TP False negative; FN	False positive; FP True negative; TN

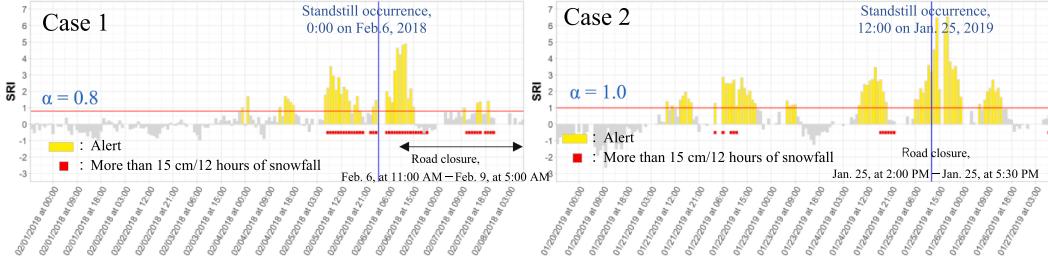


Fig. 9. Time variation of the standstill risk index (SRI) of the proposed method.

evaluated from $f_t(x)$ using Eq. (1) (or Eq. (A3)) with $n(t) = 3$ samples. The $f_{X_{85}(t)}(x)$ with mean $\mu_{X_{85}}(t) = 56.54 \text{ km/h}$ and variance $\sigma_{X_{85}}^2(t) = 5.04^2 [\text{km/h}]^2$ apparently remained in a higher speed range than $f_t(x)$. The gray bands in Fig. 6(a) and (b) correspond to $\mu_{X_{85}}(t) \pm \sigma_{X_{85}}(t)$. The fine and dashed gray lines in Fig. 7 denote $f_{X_{85}(t)}(x)$ with $n(t) = 5$ and $n(t) = 10$, respectively. Increasing the number of samples slightly increased the mean $\mu_{X_{85}}(t)$ and reduced the variance $\sigma_{X_{85}}^2(t)$.

The red line in Fig. 7 is the distribution of predicted speeds $v_{85}(t)$ based on the exogenous variables $\text{snow}(t)$, $\text{temp}(t)$, and $n(t)$ in the SSM is filtered by the observed 85th percentile speed. $\hat{v}_{85}(t) = 21.18 \text{ km/h}$, and the distribution was normal with mean $\mu_{v_{85}}(t) = 40.50 \text{ km/h}$ and variance $\sigma_{v_{85}}^2(t) = 2.67^2 [\text{km/h}]^2$. The red bands in Fig. 6(a) and (b) correspond to $\mu_{v_{85}}(t) \pm \sigma_{v_{85}}(t)$. In this example, $\hat{v}_{85}(t)$ was much lower than the estimated mean $\mu_{v_{85}}(t)$; however, this is not always the case. Fig. 8 plots the relationship between $\hat{v}_{85}(t)$ and $\mu_{v_{85}}(t)$ over 1560 h along the same road section. The observed $\hat{v}_{85}(t)$ was slightly lower than $\mu_{v_{85}}(t)$ and was more widely distributed.

From the estimated $\mu_{X_{85}}(t)$, $\sigma_{X_{85}}^2(t)$ and $\mu_{v_{85}}(t)$, the $SRI(t)$ was determined as $(56.54 - 40.50)/5.04 = 2.98$. Increasing the number of samples $n(t)$ would increase the $SRI(t)$ and the more likely activate the alert because $f_{X_{85}(t)}(x)$ deviates more from the distribution of $v_{85}(t)$ as seen in Fig. 7. Panels (a) and (b) of Fig. 6 were generated by repeating these calculations at each time t .

5.4. Model validation

5.4.1. Confusion matrix and F-score

The performance of the proposed model was evaluated using the confusion matrix (see Table 3), which comprises four elements, i.e., true positive (TP), false positive (FP), false negative (FN), and true negative (TN). For the risk alert herein, true (false) positive implies that our model predicts the actually risky (safe) condition as risky, and true (false) negative implies our model predicts actually safe (risky) condition as safe. Using these four elements, the recall, specificity, and precision are defined as shown in Eqs. (15), (16), and (17). The recall indicates the fraction of actually risky conditions predicted as risky. Thus, a small recall implies that the model misjudged many risky conditions as safe. The specificity is the fraction of actually safe conditions predicted as safe. Hence, specificity is the counterpart of recall. Conversely, precision shows the fraction of predicted risky conditions that are actually risky. Thus, a small precision implies that the model misjudged many safe conditions as risky. Although the model with a higher recall and precision is better, normally there exists a trade-off between recall and precision. Considering a balance of these indices, the F-score, i.e., the harmonic mean of recall and precision, was used as an index of model performance.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (15)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{FP} + \text{TN}} \quad (16)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (17)$$

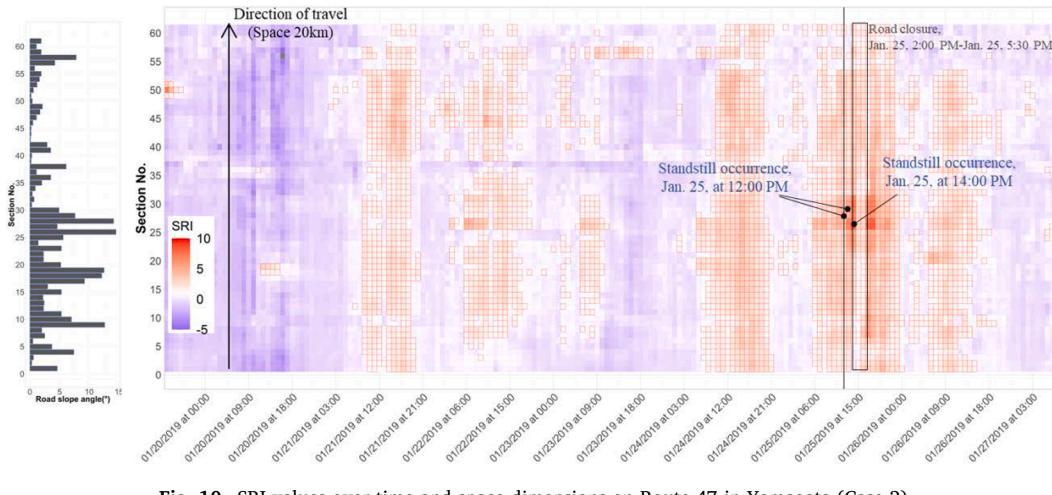
$$F\text{-score} = \frac{2}{1/\text{Recall} + 1/\text{Precision}} \quad (18)$$

To calculate the above performance indices, we defined actual risky conditions and threshold value α as shown in Eq. (19).

Table 4

Model performance for two cases.

Case	Standstill occurrence route	Model	F-score	Recall	Specificity	Precision
1	Route 8 in Fukui city	OCSVM	0.45	0.89	0.83	0.30
		Our Model ($\alpha = 0.8$)	0.51	0.64	0.93	0.42
2	Route 47 in Yamagata city	OCSVM	0.38	0.41	0.87	0.35
		Our Model ($\alpha = 1.0$)	0.46	0.70	0.76	0.34

**Fig. 10.** SRI values over time and space dimensions on Route 47 in Yamagata (Case 2).

Although it is tough to define the actually risky condition, we referred the condition defined by MLIT—a condition with more than 15 cm/12 h of snowfall because this condition was established based on practitioners' daily experiences. Fig. 9 illustrates the risky conditions defined using MLIT in red squares for cases in Fukui and Yamagata. We also include the conditions (12) h before the standstill events as the risky conditions.

On the other hand, the yellow bars in the figure indicate $SRI(t)$ values exceeding the threshold value $\alpha > 0$, implying that $\mu_{v_{85}}(t)$ is lower than $\mu_{X_{85}}(t) - \alpha \cdot \sigma_{X_{85}}(t)$. In these examples, α was determined as 0.8 and 1.0 for Cases 1 and 2, respectively, to maximize the F-score. The F-score, recall, specificity, and precision in each case at the adopted thresholds are shown in Table 4, in which our model performance is compared with that of One Class Support Vector Machine (OCSVM) that is a major method to detect anomaly events without training data. The OCSVM is applied to the same data sets in Fukui (Case 1) and Yamagata (Case 2). However, since the OCSVM cannot detect anomaly events without probe speed measurements, for the fair comparison, F-score is calculated excluding periods without probe data such as just after the standstill occurrences. Two parameters in OCSVM, the Gaussian-Kernel and parameter of an upper bound on the fraction of training errors, are calibrated so as to maximize the F-score. The F-scores of the proposed model are higher than those of OCSVM for both cases: 0.51 and 0.45 in Case 1 and 0.46 and 0.38 respectively.

$$\text{Alert activation : } SRI(t) > \alpha \quad (19)$$

The activated alerts seem reasonably agree with the red squares even before the standstill occurrences. The alert was not always activated during periods in which the snowfall exceeds the MLIT threshold of 15 cm/12 h (the period of red squares). However, the alert was successfully activated in advance to the standstill occurrence even if the snowfall did not exceed the MLIT threshold, particularly in Case 2. This possible advantage of our model can be attributed to its handling of both weather conditions and probe vehicle speeds.

The number of probe samples $n(t)$ typically decreases mostly after a standstill event; however, in some cases, like the one in Fukui, as seen in Fig. 6(c), the number of samples began decreasing before the event. Therefore, a decreasing number of probe samples could indicate a risky situation. Herein, the number of probe vehicles $n(t)$ has been included as one of the exogenous variables in the measurement model of SSM. The positive coefficient $\beta_n(t)$ implied that the 85th percentile speed decreased with decreasing $n(t)$. Although the effect of $n(t)$ has been considered in this way, the impact of $n(t)$ may not be sufficient compared with other exogenous explanatory variables. The low impact of $n(t)$ was attributed to the property of the statistically feasible speed; that is, if $n(t)$ decreases, the variance $\sigma_{X_{85}}(t)$ increases, as shown in Fig. 7, and the alert is unlikely to be activated.

The threshold α that maximizes the F-score can be determined in road sections wherein risky situations have already occurred. In northern part of Japan, there are several road sections wherein standstills frequently occur, and historical data have been acquired. Conversely, for a section lacking historical standstill data, the threshold value can be inferred based on the environment similarity to a section with already calibrated threshold values. The threshold value should also be updated using the newly accumulated data and adjusted based on the experiences of road managers.

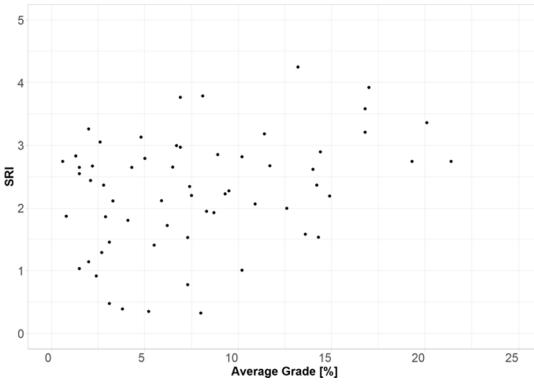


Fig. 11. Relationship between SRI and the grade.

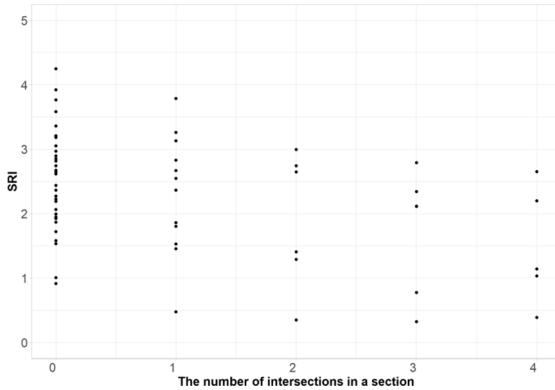


Fig. 12. Relationship between SRI and the number of intersections in a section.

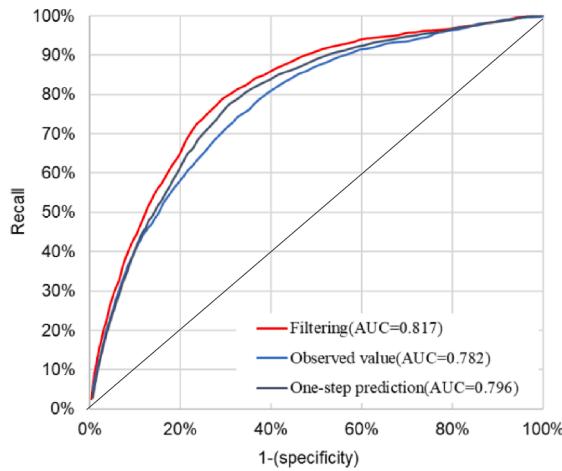


Fig. 13. Receiver operating characteristics (ROC) curves for each method of estimating current road performance.

Fig. 10 shows the temporal and spatial *SRI* variations on Route 47 in Yamagata (Case 2). The red (blue) areas denote higher (lower) *SRI* values and the scale to the left denotes the grade along the road section. The small red squares indicate the time and space where the *SRI* values exceeded the threshold $\alpha = 1.0$. At the time of the standstill, the *SRI* values were high in both the standstill section and in its vicinity. The *SRI* values gradually increased and decreased over the space. This gradual change is reasonable because the amount of snowfall and the highway geometric design do not abruptly change in the space. Figs. 11 and 12 relate the *SRI* values to the road grade and number of intersections in a section. A weak tendency that the steeper the grade the lower the *SRI* values was observed. A section with steeper grade seems to increases the speed reduction, yielding the higher *SRI* values. Furthermore, the number of intersections

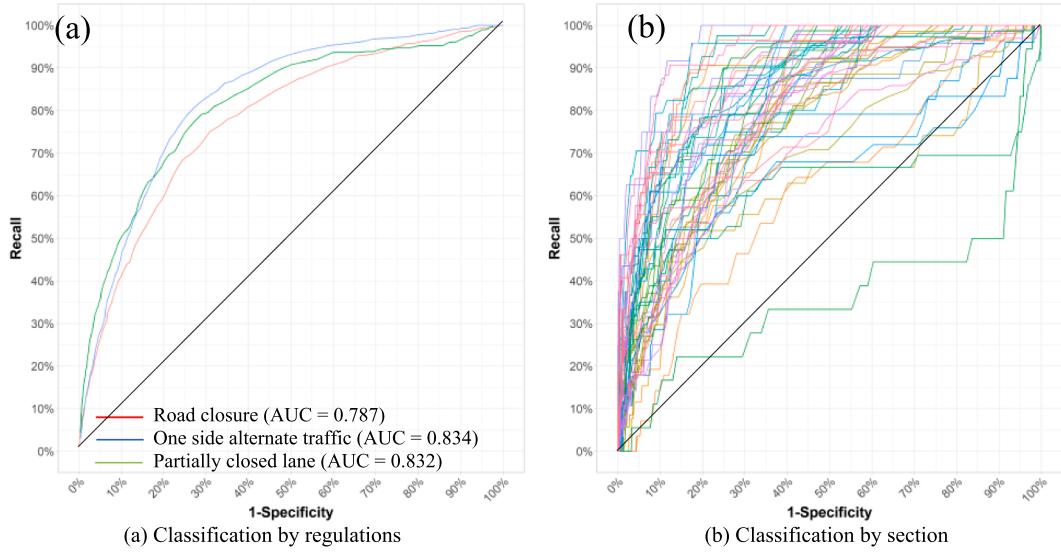


Fig. 14. Receiver operating characteristics (ROC) curves of the proposed method.

Table 5
Model performance for sections.

AUC Values	Model performance	Number of sections
>0.9	Excellent	11 sections
0.8–0.9	Very good	24 sections
0.7–0.8	Good	14 sections
0.6–0.7	Satisfactory	3 sections
0.5–0.6	Unsatisfactory	0 section
<0.5	Worse than random	1 section

was weakly related to the *SRI* values. In a section with the more number of intersections, the vehicle speed could be lower even under the normal condition; therefore, the *SRI* value could be reduced owing to the smaller speed reduction compared with the section with less number of intersections. Conversely, a large number of intersections enhances the acceleration/deceleration opportunities and may increase the risk of standstills.

5.4.2. Receiver operating characteristics (ROC) curve and AUC

The four elements of the confusion matrix as well as the indices depend on the threshold value α . Fig. 13 shows the relationship between “1.0 – specificity” and “recall” for different values of threshold α . The curve is called the receiver operating characteristics (ROC) curve. Generally, the alert will be frequently activated with small α even if the mean of the current vehicle speed $\mu_{v_{85}}(t)$ is slightly smaller than $\mu_{X_{85}}(t)$, and hence, the recall may be close to 1. However, since most conditions will be judged risky, the specificity may be close to zero; that is, (1.0 – specificity) will be nearly equal to 1. Therefore, the ROC curve normally increases monotonically.

The area under the curve (AUC), which is the area below the ROC curve, was proposed as a reasonable index to represent the overall performance of the model (Metz (1978), Mandrekar (2010)). In general, an AUC of 0.5–0.6 is unsatisfactory, 0.6–0.7 is satisfactory, 0.7–0.8 is good, 0.8–0.9 is very good, and >0.9 is excellent Trifonova et al. (2013), Safari et al. (2016).

Fig. 14 shows the ROC curves for all 58 cases. The blue line in the figure is the ROC curve when the alert is activated using the mean of current vehicle speed $\mu_{v_{85}}(t)$ filtered from the observed speed and the speed predicted by the SSM. The black and orange lines are the ROC curves based on the alert activation using the predicted and observed speeds, respectively. Among the three curves, the filtered speed $\mu_{v_{85}}(t)$ achieved the largest AUC (0.817, implying very good predictive performance), confirming that the proposed SSM model outperformed the models using only the observed or predicted speed. Also, as seen the observed speeds (black dots) in the period of red squares in Fig. 6 (a), they sometime deviate from the grey band and cause unnecessary error alerts even during periods without heavy snowfall.

Fig. 14(a) shows the ROC curves for three traffic regulation types immediately after standstills (see Table 1). The AUC values are 0.787, 0.834, and 0.832 for complete lane closure, one-side alternate traffic, and partial lane closure, respectively. The AUC tends to decrease as the impact becomes more severe.

Fig. 14(b) shows the ROC curves for each of the 58 standstills, and Table 5 lists the AUC values for the standstills. Despite section-wise variations in the AUC, the performance of the proposed model was excellent in 11 sections, very good in 24 sections, good in 14 sections, satisfactory in 3 sections, and worse in 1 section.

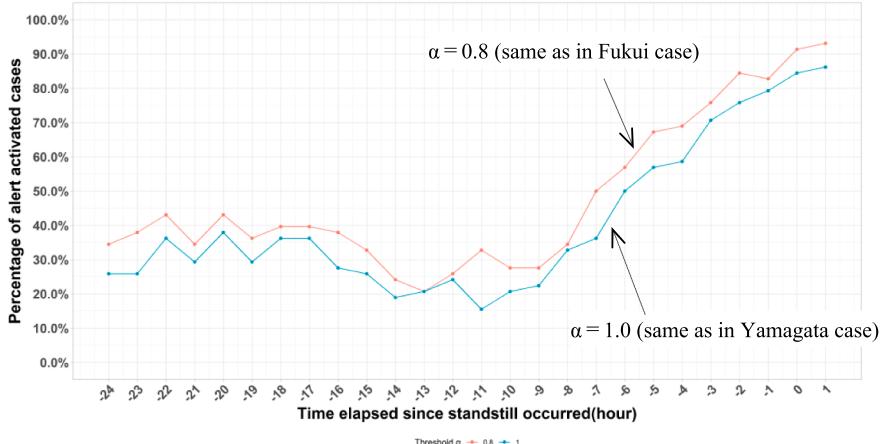


Fig. 15. Percentage of alert activation before and after the standstill occurrence (time 0 is the standstill occurrence time.)

Fig. 15 shows the earliness of the risk evaluation. The horizontal axis represents the time before (negative values) and after (positive values) the standstill occurrence and the vertical axis is the percentage of alert activations among 58 applications at two different threshold values of $\alpha = 0.8$ and 1.0. As the lower threshold α activates the alert even at smaller SRI value than at the larger threshold, the percentage of the alert activation when $\alpha = 0.8$ is higher than that when $\alpha = 1.0$. However, both these instances showed similar tendencies; for instance, about 20% cases activated the alert 12 h in advance, 50% cases activated 4 h in advance, and 70% cases activated the alert 2 h in advance.

As discussed above, our model outperformed the MLIT assessment based solely on the snowfall and a model that used the observed speeds alone. The AUC value exceeded 0.8, indicating reasonable model performance. However, it is difficult to verify whether the performance of our proposed model is sufficient for the practical use. This is mainly because the truly risky situations are not exactly known, even if actual occurrences of standstills are known. The occurrences of standstills depend not only on risky environments but also on individual driver/vehicle conditions; thus, a standstill may not occur even under a truly risky situation and vice versa. In this analysis, we tentatively assume a true risky condition based on the MLIT snowfall threshold as well as actual standstill occurrences and propose the optimal threshold α that maximizes the F-score. However, this tentative assumption may not be always true and road managers may not always prefer to maximize the F-score, which balances the recall and precision. For instance, some road managers may instead prefer a higher recall rate to reduce the chance of missing risks or they may prefer a higher precision to reduce the number of false alerts. Our model is not intended to replace the current practice of the risk evaluation; rather it can assist in their judgement. Therefore, the threshold α value as well as the model parameters should be customized based on the requirements of practitioners.

6. Conclusion and future work

This study proposes a method that evaluates the risk of traffic standstills, quite frequently happening in winter in Japan, in real time by using a state space model. If the risk of the standstills is known in advance, appropriate preparation and management can be undertaken to prevent such events and/or alleviate the impacts on road traffic. (Note that this study does not aim to estimate the occurrences of standstill events but the risk, since the event occurrences greatly depend on individual vehicle and driver characteristics.) Also, this study does not intend to replace the current practice of risk judgment by road managers with the proposed model; instead, the objective is to assist the managers in making judgment by providing them additional useful traffic information.

Using probe vehicle and weather data, the risk of traffic standstills was evaluated as the deviation of the current 85th percentile speed estimated by the SSM from the statistically feasible 85th percentile speed. The SSM is advantageous because it evaluates the risk probabilistically without training data considering the uncertainty in probe data. Also, it can handle missing data situations, which typify snowy mountainous areas. Based on the intensive model validation, we confirm that the model successfully evaluates the risks prior to the event occurrence at the reasonable level.

The major remarks are summarized as follows:

- (1) The 85th percentile speed is specifically used rather than the simple speed for the standstill analysis to eliminate minor influential factors on vehicle speeds such as speed reductions due to traffic lights, on-street parking, pedestrian crossings and so on. The similarity of the 85th percentile speed to the Normal distribution also justifies the use of the Kalman Filter which assumes the Gaussian noises.
- (2) The measurement model in the SSM is specifically constructed to evaluate the current 85th percentile speed using the amount of snow, the temperature, the number of probe vehicles and the periodical factor as the explanatory variables.
- (3) The statistically feasible 85th percentile speed is mathematically derived from the population speed distribution and the number of probe samples.

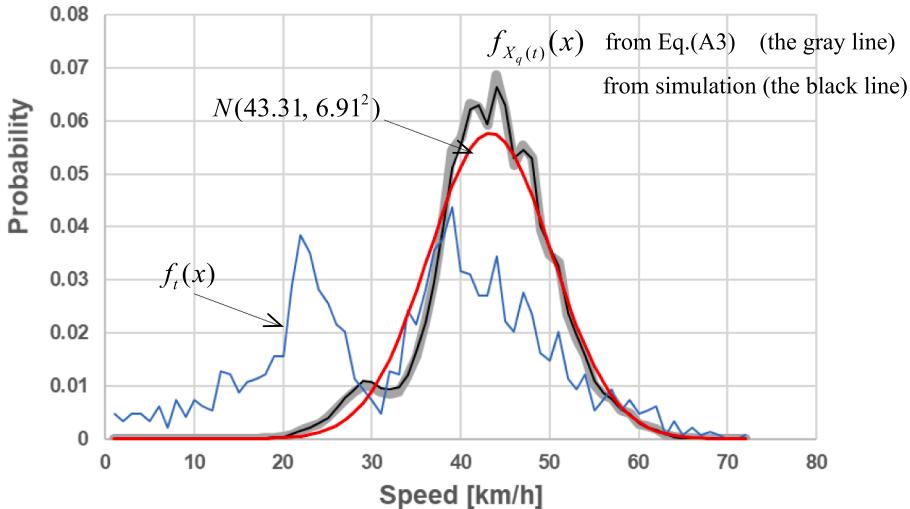


Fig. A1. Estimated 85th percentile speed distribution from Eq. (A3) and by the random sampling. blue line: the population speed distribution estimated from 1504 probe samples in Fukui. gray line: the estimated $f_{X_q(t)}(x)$ from Eq. (A3). orange line: the simulated $f_{X_q(t)}(x)$ by the random sampling. red line: Normal distribution with mean $\mu_{X_q}(t)$ and variance $\sigma_{X_q}^2(t)$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

- (4) The model is validated by applying it to 58 standstill cases in the northern part of Japan. Namely, the model performance is checked using the receiver operating characteristics (ROC) curve and the area under the curve (AUC) value of the proposed method is ‘very good’ level of 0.817. The proposed model is also compared with One Class Support Vector Machine, a major method of the anomaly detection without training data, and shows the better performance over the OCSVM.

For the practical use, the following studies must be further conducted.

- (1) Some model parameters, such as the warning threshold value α , must be customized in relation to the road type and the road management system with the help of road managers.
- (2) The road alignment such as the gradient and the curvature may be included in the SSM to improve the model performance.
- (3) The risk is evaluated based on vehicle speeds, which largely depend on the degree of road disturbance caused by a standstill. Therefore, to validate the model performance more accurately, we have to exactly know the degree of disturbance that has the large variability: a short time vehicle stopping, a partial lane closure, and a complete lane closure.
- (4) Sizes of the discretized space and time have to be examined. Also, the introduction of variable discretization lengths may be a future extension.

CRediT authorship contribution statement

Shogo Umeda: Conceptualization, Methodology, Software, Writing - original draft. **Yosuke Kawasaki:** Validation, Writing - review & editing, Formal analysis. **Masao Kuwahara:** Supervision, Project administration. **Akira Iihoshi:** Resources, Data curation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was supported by Commissioned Research (201) of the National Institute of Information and Communications Technology of Japan, CART (Committee on Advanced Road Technology), Ministry of Land, and JSPS KAKENHI grant number JP19H00777.

Appendix. Probability density of the q -percentile speed, $f_{X_q(t)}(x)$

Eq. (1) below is the probability density function of the q -percentile speed, when $m = n(t)q$ is an integer.

$$f_{X_q(t)}(x) = n(t)f_t(x) \cdot {}_{n(t)-1}C_{n(t)q-1}F_t(x)^{n(t)q-1}(1-F_t(x))^{n(t)(1-q)} \quad (\text{A1})$$

Since this can be reinterpreted as the probability density that m -th value out of $n(t)$ samples is x , the expression is slightly modified as shown below:

$$f_{X_q(t)}(x) = g_{n(t)}^m(x) = n(t)f_t(x) \cdot {}_{n(t)-1}C_{m-1}F_t(x)^{m-1}(1-F_t(x))^{n(t)-m} \quad (\text{A2})$$

In general, $n(t)q$ may not always be an integer, especially if the number of samples is small. In this case, the probability density function of the q -percentile speed can be evaluated as follows.

1. Obtain integer m so that $\frac{m-1}{n(t)} < q \leq \frac{m}{n(t)}$. Then, the q -percentile speed x must exist between x_{m-1} and x_m .
2. The q -percentile speed x is estimated by the linear interpolation between x_{m-1} and x_m :

$$x = \alpha x_m + (1 - \alpha)x_{m-1}, \text{ where } \alpha = n(t)q - (m - 1)$$

In other words, if the $(m-1)$ -th speed is x_{m-1} , the m -th speed x_m must be as shown below so that the q -percentile speed is x . Hence, x_m is written as a function of x_{m-1} : $x_m(x_{m-1})$.

$$x_m(x_{m-1}) = \frac{1}{\alpha}x + \left(1 - \frac{1}{\alpha}\right)x_{m-1}$$

3. Therefore, $f_{X_q(t)}(x)$ is evaluated by integrating the conditional probability shown below.

$$\begin{aligned} f_{X_q(t)}(x) &= \int_0^x g_{n(t)}^m(x_m(x_{m-1})|x_{m-1})g_{n(t)}^{m-1}(x_{m-1})dx_{m-1} \\ &= \int_0^x \frac{(n(t) - m + 1) \cdot f_t(x_m(x_{m-1})) (1 - F_t(x_m(x_{m-1})))^{n(t)-m}}{(1 - F_t(x_{m-1}))^{n(t)-m+1}} g_{n(t)}^{m-1}(x_{m-1}) dx_{m-1} \end{aligned} \quad (\text{A3})$$

The $g_{n(t)}^m(x_m|x_{m-1})$ is the probability density that the m -th speed is x_m when the $(m-1)$ -th speed is x_{m-1} . Under this condition, since $m-1$ samples must be smaller than x_m , only $n(t)-m+1$ samples can be x_m . The nominator on the second line is therefore the probability that one out of $n(t)-m+1$ samples takes x_m and the rest of $n(t)-m$ samples are larger than x_m , and the denominator is the probability that all of $n(t)-m+1$ samples are larger than x_{m-1} . Since $g_{n(t)}^{m-1}(x_{m-1})$ is given from (A2) and both $f_t(\cdot)$ as well as $F_t(\cdot)$ are known, the integration in (A3) can be carried out with respect to x_{m-1} for $0 \leq x_{m-1} \leq x$.

Fig. A1 compares the estimated 85th percentile speed distribution from Eq. (A3) with the simulated distribution by the random sampling using 1504 observed speed samples in a segment in Fukui at time of day $h = 12:00$. The blue line is the estimated population speed distribution, $f_t(x)$, which has two peaks because of the traffic light and the gray line is the estimated 85th percentile speed distribution from Eq. (A3), $f_{X_q(t)}(x)$, with $n(t) = 5$. On the other hand, the black line, mostly overlapped with the gray line, is drawn from the random sampling of 1 million times; that is, 5 samples are randomly extracted from 1504 observed speed and the 85th percentile speed is estimated from the extracted 5 samples. From this validation, it is confirmed that Eq. (A3) accurately estimates $f_{X_q(t)}(x)$ with mean $\mu_{X_q}(t) = 43.31$ [km/h] and variance $\sigma_{X_q}^2(t) = 6.91^2$ [km/h]², and the distribution may be approximated as the Normal distribution as shown in the red line even if the population speed distribution $f_t(x)$ has two peaks.

References

- Ahmed, M., Abdel-Aty, M., 2013. A data fusion framework for real-time risk assessment on freeways. *Transp. Res. Part C Emerg. Technol.* 26, 203–213.
- Asakura, Y., Kusakabe, T., Long, N.X., Ushiki, T., 2015. Incident detection methods using probe vehicles with on-board GPS Equipment. *Transp. Res. Procedia* 6, 17–27.
- Boquet, G., Morell, A., Serrano, J., Vicario, J.L., 2020. A variational autoencoder solution for road traffic forecasting systems: Missing data imputation, dimension reduction, model selection and anomaly detection. *Transp. Res. Part C Emerg. Technol.* 115, 102622.
- Metz, Charles E., 1978. Basic principles of ROC analysis. *Semin. Nucl. Med.* VIII. 4, 283–298.
- Datla, S., Sahu, P., Roh, H.-J., Sharma, S., 2013. A comprehensive analysis of the association of highway traffic with winter weather conditions. *Procedia – Soc. Behav. Sci.* 104, 497–506.
- El-Basyouny, K., Barua, S., Islam, M.T., 2014. Investigation of time and weather effects on crash types using full bayesian multivariate poisson lognormal models. *Accid. Anal. Prev.* 73, 91–99.
- Heqimi, G., Gates, T.J., Kay, J.J., 2018. Using spatial interpolation to determine impacts of annual snowfall on traffic crashes for limited access freeway segments. *Accid. Anal. Prev.* 121, 202–212.
- Horiguchi, R., Iijima, M., Kobayashi, M., Hanabusa, H., 2013. Traffic anomaly detection for surface street networks with the mesh-wised traffic indices on macroscopic fundamental diagram. *Optim. 2013 – Int. Symp. Recent Adv. Transp. Model.*, pp. 1–7.
- Huang, T., Wang, S., Sharma, A., 2020. Highway crash detection and risk estimation using deep learning. *Accid. Anal. Prev.* 135.
- Jiang, F., Yuen, K.K.R., Lee, E.W.M., 2020. A long short-term memory-based framework for crash detection on freeways with traffic data of different temporal resolutions. *Accid. Anal. Prev.* 141, 105520.
- Jong, P., De, 1991. The diffuse kalman filter. *R. Ann. Stat.* 19, 1073–1083.
- Kalman, R.E., 1960. A new approach to linear filtering and prediction problems. *J. Fluids Eng. Trans. ASME* 82, 35–45.

- Kinoshita, A., Takasu, A., Adachi, J., 2015. Real-time traffic incident detection using a probabilistic topic model. *Inf. Syst.* 54, 169–188.
- Li, P., Abdel-Aty, M., Yuan, J., 2020. Real-time crash risk prediction on arterials based on LSTM-CNN. *Accid. Anal. Prev.* 135, 105371.
- Lu, Z., Kwon, T.J., Fu, L., 2019. Effects of winter weather on traffic operations and optimization of signalized intersections. *J. Traffic Transp. Eng. English Ed.* 6, 196–208.
- Mandrekar, J.N., 2010. Receiver operating characteristic curve in diagnostic test assessment. *J. Thorac. Oncol.* 5, 1315–1316.
- Safari, S., Baratloo, A., Elfil, M., Negida, A., 2016. Evidence based emergency medicine; Part 5 Receiver Operating Curve and Area under the Curve. *Emerg* 4, 111–113.
- Seeherman, J., Liu, Y., 2015. Effects of extraordinary snowfall on traffic safety. *Accid. Anal. Prev.* 81, 194–203.
- Takenouchi, A., Kawai, K., Kuwahara, M., 2019. Traffic state estimation and its sensitivity utilizing measurements from the opposite lane. *Transp. Res. Part C Emerg. Technol.* 104, 95–109.
- Theofilatos, A., 2017. Incorporating real-time traffic and weather data to explore road accident likelihood and severity in urban arterials. *J. Saf. Res.* 61, 9–21.
- Trifonova, O.P., Lokhov, P.G., Archakov, A.I., 2013. Metabolic profiling of human blood. *Biochem. Suppl. Ser. B Biomed. Chem.* 7, 179–186.
- Wang, J., Luo, T., Fu, T., 2019. Crash prediction based on traffic platoon characteristics using floating car trajectory data and the machine learning approach. *Accid. Anal. Prev.* 133, 105320.
- Wu, Y., Abdel-Aty, M., Cai, Q., Lee, J., Park, J., 2018. Developing an algorithm to assess the rear-end collision risk under fog conditions using real-time data. *Transp. Res. Part C Emerg. Technol.* 87, 11–25.
- Yang, K., Wang, X., Yu, R., 2018. A bayesian dynamic updating approach for urban expressway real-time crash risk evaluation. *Transp. Res. Part C Emerg. Technol.* 96, 192–207.
- Yasanthi, R.G.N., Mehran, B., 2020. Modeling free-flow speed variations under adverse road-weather conditions: Case of cold region highways. *Case Stud. Transp. Policy* 8, 22–30.
- Yu, R., Abdel-Aty, M., 2013. Utilizing support vector machine in real-time crash risk evaluation. *Accid. Anal. Prev.* 51, 252–259.
- Yu, R., Xiong, Y., Abdel-Aty, M., 2015. A correlated random parameter approach to investigate the effects of weather conditions on crash risk for a mountainous freeway. *Transp. Res. Part C Emerg. Technol.* 50, 68–77.
- Zhai, B., Lu, J., Wang, Y., Wu, B., 2020. Real-time prediction of crash risk on freeways under fog conditions. *Int. J. Transp. Sci. Technol.*
- Zhao, J., Yi, Z., Pan, S., Zhao, Y., Zhuang, B., 2019. Unsupervised traffic anomaly detection using trajectories. *CVPR Work.*