Modelling the Safety Tolerance Zone: Recommendations from the i-DREAMS project

Eva Michelaraki[1]*, Christos Katrakazas[2], Tom Brijs[3], George Yannis[4]

[1,2,4]National Technical University of Athens, Department of Transportation Planning and Engineering, 5 Heroon Polytechniou str., GR-15773, Athens, Greece
[3]UHasselt, School of Transportation Sciences, Transportation Research Institute (IMOB), Agoralaan, 3590 - Diepenbeek, Belgium

E-mail[1] *corresponding author: evamich@mail.ntua.gr
E-mail[2]: ckatrakazas@mail.ntua.gr
E-mail[3]: tom.brijs@uhasselt.be
E-mail[4]: geyannis@central.ntua.gr

Abstract

The i-DREAMS project aims at defining, developing, testing and validating a "Safety Tolerance Zone" (STZ) in order to prevent drivers from getting too close to the boundaries of unsafe operation by mitigating risk in both real-time and post-trip. The aim of the current study is to provide guidelines for mapping the concept of the STZ using continuous variables of risk and the most reliable indicators (e.g. time headway, speed, harsh acceleration, distraction) are going to be investigated in real-time. For the purpose of the analysis, a variety of analytical methods and potential modelling approaches are proposed. According to the research question made, a mapping exercise of machine learning algorithms (e.g. Long Short-Term Memory or Dynamic Bayesian Networks) is implemented for real-time data prediction. The key output will be the correlation of the aforementioned explanatory variables and various indicators of task complexity and coping capacity with the dependent variable risk.

Keywords: *i-DREAMS project, safety tolerance zone, mapping methodology, continuous risk indicators, real-time prediction.*

Περίληψη

Το έργο i-DREAMS στοχεύει στον ορισμό, την ανάπτυξη, τη δοκιμή και την επικύρωση του "εύρους ανοχής ασφαλείας" (STZ), ώστε να αποφευχθεί η υπερβολική απόκλιση των οδηγών από τα όρια της ασφαλούς λειτουργίας, μειώνοντας τον κίνδυνο τόσο σε πραγματικό χρόνο όσο και μετά το ταξίδι . Στόχος της παρούσας μελέτης είναι να παρέχει οδηγίες για τη χαρτογράφηση της έννοιας της STZ χρησιμοποιώντας συνεχείς μεταβλητές κινδύνου και οι πιο αξιόπιστοι δείκτες (π.χ. απόσταση προπορευόμενου οχήματος, ταχύτητα, απότομη επιτάχυνση, απόσπαση προσοχής) θα διερευνηθούν σε πραγματικό χρόνο. Για τον σκοπό αυτό, προτείνονται αρκετές μέθοδοι ανάλυσης και προσεγγίσεις μοντελοποίησης. Σύμφωνα με τις ερευνητικές ερωτήσεις που πραγματοποιήθηκαν, διάφοροι αλγόριθμοι μηχανικής μάθησης (π.χ. Long Short-Term Memory ή Dynamic Bayesian Networks) εφαρμόζονται για την πρόβλεψη δεδομένων σε πραγματικό χρόνο. Το βασικό αποτέλεσμα θα είναι η συσχέτιση των προαναφερόμενων επεξηγηματικών μεταβλητών και των δεικτών πολυπλοκότητας που σχετίζονται με την οδήγηση με τον εξαρτώμενο κίνδυνο.

Λέξεις κλειδιά: *Έργο i-DREAMS, εύρος ανοχής ασφαλείας, μεθοδολογία χαρτογράφησης, δείκτες κινδύνου, πρόβλεψη σε πραγματικό χρόνο.*

1. Introduction

The overall objective of the European H2020 i-DREAMS[1] project is to define, develop, test and validate a context-aware safety envelope for driving in a 'Safety Tolerance Zone' (STZ), with a smart Driver, Vehicle & Environment Assessment and Monitoring System. Taking into account, on the one hand, driver background factors and real-time risk indicators, and on the other hand, driving task complexity indicators, a continuous real-time assessment will be created to monitor and determine if a driver is within acceptable boundaries of safe operation (i.e. Safety Tolerance Zone). Testing and validation will be applied to car, bus and truck drivers as well as to tram and train drivers.

Within a transport system, a driver can be regarded as a human operator (technology assisted) self-regulating control over transportation vehicles in the context of crash avoidance. The concept of the STZ within the i-DREAMS platform attempts to describe short of the range at which self-regulated control is considered safe. It is based on Fuller's Task Capability Interface Model (Ray Fuller, 2000, 2005, 2011) which states that loss of control occurs when the demand of a driving task outweighs the operator's capability.

The STZ contains three phases: normal driving phase, danger phase and avoidable accident phase. Firstly, the normal driving refers to the phase where conditions at that point in time suggest that a crash is unlikely to occur and therefore the crash risk is low and the operator is successfully adjusting their behavior to meet task demands. Fundamental goal of the i-DREAMS platform is to keep drivers within this normal phase. Secondly, the danger phase is characterized by changes to the normal driving that suggest a cash may occur and therefore, there is an increased crash risk. At this stage a crash is not inevitable but becomes more likely. The STZ switches to the danger phase whenever instantaneous measurements detect changes that imply an increased crash risk. Lastly, the switch to avoidable accident phase occurs when a collision scenario is developing but there is still time for the operator to intervene in order to avoid the crash. In this phase, the need for action is more urgent as if there are no changes or corrections in the road or rail traffic system or an evasive manoeuvre is performed by the operator a crash is very likely to occur.

It should be mentioned that the i-DREAMS platform is composed of two modules. The first is the monitoring module that takes measurements relating to the "context" (i.e. environment including infrastructure), "operator" (i.e. driver state and demographic characteristics) and "vehicle" (i.e. technical specifications and current state). These "Context - Operator - Vehicle - COV" measurements are used to infer the demands of the driving task (i.e. task complexity) and the driver's capability to cope with these demands (i.e. coping capacity). These inferences on its turn are used to estimate in which phase within the STZ the driver is operating within at each moment in time. The second module is the in-vehicle intervention module, that is responsible for keeping the driver within the normal phase of the STZ all the times, either by providing a warning or instruction during driving (real-time intervention) or providing

[1]Further general project information can be found on the website: https://idreamsproject.eu

information with detailed feedback on the trip as well as improving their performance once the driving task has ended (post-trip intervention). The STZ phase, within which the driver is operating, dictates the type of real-time intervention that is delivered. In the normal driving phase, no intervention is needed. If it is detected that a driver has entered the danger phase, a warning or advice should be given. Entering the avoidable accident phase also requires an intervention, but this may need to be more specific and provide an instruction signal, which impels the operator to take a decisive action.

The conceptual state of the STZ changes dynamically depending upon changes in the driving conditions or system of which the operator is an integral part. The drivers' self-regulated control has many influences, one of which is the driver's own perception of the driving conditions. Drivers seek to maintain a level of risk that they are comfortable with and continuously adapt their behavior to achieve the subject to a complex network of underlying motivations, not all of which relate to safety. This implies that drivers may choose to intentionally behave in a way that objectively would be considered unsafe (i.e. travelling close to a vehicle ahead). A driver's subjective appraisal of risk does not necessarily therefore correspond to that calculated with objective measures, nevertheless the driver would still be classed as operating within the danger or avoidable accident phases of the STZ (Ray Fuller, 2011).

It is worth mentioning that data analysis consists a pivotal part of this project for achieving its objectives and the methods for data analysis highly depend on the collected data. In order to model the STZ, the available data as well as the potential outcome of the model need to be considered. For suggesting a positive outcome, the data to be used as input for the model will be available in real-time, as the measurements of task demand (e.g. road layout, weather conditions, time of the day) and coping capacity (e.g. fatigue, distraction, sleepiness) are going to be sequential. Furthermore, as the STZ is the "trigger" for real-time and post-trip interventions, the algorithm outputs are required also to be provided online as in real-time and hence both dynamic and static modelling approaches need to be considered. Distinguishing between the three levels of STZ (i.e. normal driving, danger and avoidable accident phases) in real-time, turns STZ modelling into a ternary classification problem, where raw measurements need to be classified as belonging to one of the three existing levels. This classification problem however implies that the feed to the training part of these algorithms needs to be conveniently labelled.

Within, the above framework, the aim of the current study is to provide guidelines for mapping the concept of the STZ using continuous outcome variables. The overall objective is to further elaborate on the analytical methods and modelling options, their strengths, limitations and applications in i-DREAMS. Particular focus will be given in continuous variables of risk and the most reliable indicators are going to be investigated in real-time. For instance, time headway, speed, acceleration, deceleration, distraction, fatigue or sleepiness consist some of the continuous variables in order to model the concept of the i-DREAMS project.

The paper is structured as follows. In the beginning, the overall objective of the i-DREAMS project as well as the aim of this research is provided. Subsequently, a thorough literature review of models dealing with driver behavior and collision risk modelling in real-time are provided. Then, the most prominent approaches are detailed analyzed and a brief description of their underpinning procedure is given. Additionally, initial insights into analyses and results for real-time purposes are presented. Lastly, overall conclusions as well as practical considerations concerning the modelling of the STZ are highlighted in order to assist researchers and practitioners.

2. Background

Predicting driving behavior by employing mathematical driver models, obtained directly from the observed driving-behavior data, has gained much attention in literature (Girma et al., 2019; Kanaan et al., 2019; McDonald et al., 2019; Xue et al., 2019; Zou et al., 2018). Several models have been used to address road safety and the estimation of driving behavior, many of which in the context of experimental studies, including driving simulator studies and field operational trials and/or naturalistic driving studies. The aim of this section is to examine different models as well as methodologies that include the relationship and interaction between task demand and coping capacity. both static and dynamic state-of-the-art approaches that could be employed to convert driving behavior data into meaningful STZ information are reviewed. The most suitable models, able to estimate driving behavior and crash risk will be employed for the scope of the i-DREAMS project. Literature was searched within popular scientific databases such as Scopus, ScienceDirect and Google Scholar. All studies were screened on the basis of their title and abstract in order to select the studies presented in the following review.

2.1 Bayesian Networks

In recent years, BNs have been quite popular in modelling massive amounts of data with the need for data aggregation and model flexibility (Li et al., 2014; Tandon et al., 2016). Lefèvre et al. (2012) proposed a DBN which focused on intersection accidents caused by driver errors. Their approach was formulated as an inference problem where intention and expectation were estimated jointly for the vehicles converging to the same intersection and the proposed solution was validated by field experiments using passenger vehicles. The results demonstrated the ability of the algorithm to issue a warning in dangerous situations, and the benefits of taking into account interactions between the vehicles when reasoning about situations and risk at road intersections. The use of the Bayesian formalism allowed to take into account uncertainties on the relationships between the variables. The intuitive formulation of risk provided the required flexibility for safety applications relevant to both ADAS and autonomous driving. However, information about drivers' actions, such as steering angle and pedal pressure were not taken into account.

In addition, Zhu et al. (2017) utilized a hierarchical BN model to investigate the relationship between observed vehicle motion and a driver's historical crash involvements through the hidden layers of driving behavior and crash risk. The results suggested that the contextual model performs significantly better than the non-contextual model. The method was also effective in handling massive trajectory data and flexible in the data aggregation process. However, the contextual indicators have been more comprehensive by including more variables beyond current roadway type, traffic and relative speeds.

Katrakazas et al. (2019) developed a new risk assessment methodology that integrates a collision risk network-level (CRN) with collision risk vehicle-level (CRV) estimates in real-time under the joint framework of interaction-aware motion models and DBN. Results indicated an enhancement of the interaction-aware model by up to 10%, when traffic conditions were deemed as collision-prone. The network-level collision information could assist not only the identification of "dangerous" road users but also act as a safety net for all the motion planning levels and is suitable for Connected and Autonomous Vehicles (CAVs). It is however noteworthy that the extracted probabilities for all the scenarios were not sufficiently high and the scenarios were built on some assumptions and without highly detailed vehicle-level data.

The work by Shankar et al. (2008) pointed out that hierarchical DBN can be used to reflect how driver decisions are made: driver-level predictors, such as years of driving, can be used to parameterize the effects of event attributes and context. There were found some advantages related to parameter uncertainty, sample specificity and extensibility to large data sets, which can capture driver differences over time and space, but non automated storage of data through the DAS with a flag for potential risk was identified.

2.2 Clustering Models

Clustering techniques have been used by researchers to categorize drivers who are compliant and non-compliant. Xue et al. (2019) developed a driving style recognition method (safe, low-risk, high-risk, dangerous) based on vehicle trajectories from video recordings and k-means clustering, failing however to take into account road conditions and traffic flow levels. An individually-tailored, real-time feedback-reward system for in-vehicle interventions was installed in driver's own vehicles and its effect was researched in a field trial with 37 participants by Merrikhpour et al. (2014). Drivers were clustered by compliance rate, pre and post interventions, in more speed and headway compliant and less speed and headway compliant and the results showed that speed limit and headway compliance increased with post-intervention in the non-compliant group. However, it is not clear if the observed benefits were due to either feedback, or reward, or the unique combination of both. The study by Wang and Xu (2019) using SHRP2 data followed a two-level approach; first, a K-means algorithm was adopted to classify drivers into groups of high, moderate or low risk level and second, logistic regression models for each risk group indicated the probability of each driver getting involved in an incident. Drivers themselves participated in this study by validating any traffic event using an in-vehicle event button and by self-assessing their behavior due to inattention and inexperience errors.

2.3 Fuzzy Logic Models

Machine learning techniques have been used in primarily traffic flow modelling and second in road safety analysis. Imkamon et al. (2008) proposed a new Fuzzy Logic inference system which can record driving events, detect unsafe or risk driving behavior and classify levels of hazardous driving by employing data from sensors that measure three different perspectives (an ECU reader, an accelerometer, and a camera). The test results showed that the system can perform well compared with human opinions. However, the current system had a limitation of day-time operation due to constraints. In addition, Chong et al. (2013) trained a fuzzy rule-based neural network to model the acceleration of a car-following vehicle. Fuzzy logic was used to discretize traffic state and action variables and reinforcement learning was used for the neural network to learn driving behavior from naturalistic data. On the one hand this paper showcased the application of fuzzy rules on continuous variables with high R-squared values, but on the other hand the choice of model parameters and the number of car-following events were limited (ten in total). Fuzzy deep learning was also applied for traffic incident detection (El Hatri and Boumhidi, 2018), where the authors performed a comparison of machine learning models based on MSE with detection rate and mean time to detection as criteria. Their implementations showcased a high detection rate, low false alarm rate and a back-propagation feature to adjust the parameters in the deep network, although model validation was done on highly artificial street network and incident occurrences. The standard deviation of detection time was not given, indicating questionable potential for applying this algorithm.

2.4 Hybrid Input Output Automaton

According to Bouhoute et al. (2014), a Hybrid Input Output Automaton (HIOA) is a formal model that used to describe discrete and continuous behavior of a system. A driver-centric approach to model risky driving behavior in vehicular ad hoc networks was proposed. Their advantage consisted of providing a better analysis of hybrid systems. The constructed automaton corresponded to the supposed behavior of the driver in one trip, exploration of other states possible in next trips. The goal of the proposed example was to illustrate the idea of the modelling approach and how it can be applied. Consequently, despite the constructed model may be useful to predict the driver behavior in the future, prevent unsafe situations and provide more comfort to the driver, the implementation of the model and the learning process have been not implemented yet.

2.5 Long Short-Term Memory Models

Girma et al. (2019) proposed deep learning-based models, called LSTM models, to predict and identify drivers based on their individual's unique driving patterns based on vehicle telematics data. Results showed that the proposed model prediction accuracy remained satisfactory and outperformed other approaches despite the extent of anomalies induced in the data. Even under increasing noise and outliers' effect, the proposed approach maintained its accuracy above the acceptable value, 88%, while other models' accuracy fell below 40%. Neural network-based models such as LSTM performed better than Fully Connected Neural Network (FCNN), Decision Tree (DT) or RF, by avoiding over-fitting on the noise. Bao et al. (2019) trained a spatiotemporal convolutional LSTM network to determine a crash risk scale and to calibrate a crash risk alarm threshold. Their data comprised large-scale taxi GPS data, population data, weather and land use features, and they were used to compare econometric and machine learning models. Econometric models performed better than machine-learning models in weekly crash risk prediction tasks, while they exhibit worse results than machine-learning models in daily crash risk prediction tasks. However, because taxi trips were not representative of the general mobility patterns in a city, this study entails a significant sample bias problem.

2.6 Recommendation

With regards to safety and risk level, several models and methodologies have been examined. From the aforementioned approaches examined, DBNs were found to be the most effective and extensible in handling massive trajectory data, as well as flexible for safety applications in the data aggregation process (Lefèvre et al., 2012; Shankar et al., 2008; Zhu et al., 2017). The use of the Bayesian formalism allowed to take into account uncertainties on the relationships between the variables (Lefèvre et al., 2012). However, variable selection, assumptions and non-highly detailed vehicle-level data were found to be some of the shortcomings of this approach (Katrakazas et al., 2019). Lastly, LSTMs revealed the highest accuracy, compared with other models examined. Specifically, the proposed model maintained its accuracy above the acceptable value 88%, while other models' accuracy fell below 40%. LSTM had an inherent ability to remember temporal information in data and conserve it for many time steps, unlike other conventional machine learning approaches (Girma et al., 2019).

3. Methodological Overview

According to the research question and hypothesis made, a mapping exercise of machine learning algorithms was implemented for real-time data prediction. It should be noted that the fundamental research question within the i-DREAMS project is how explanatory variables (i.e. various indicators of task complexity and coping capacity) are correlated with the dependent variable "risk".

A variety of analytical methods and potential modelling approaches has been reviewed, among which two methods have been selected to be used in i-DREAMS: Dynamic Bayesian Networks (DBN) and Long Short-Term Memory (LSTM) deep neural networks. Each of these two methods has strengths and limitations, making it suitable for a certain purpose in the project. Based on the methodological background, an attempt was made to transform the model approach into a suitable structure which will allow to response to the research question made. The key output is expected to be the correlation of the explanatory variables and various indicators of task complexity and coping capacity with the dependent variable risk.

3.1 Dynamic Bayesian Network analytical approach

It is hypothesized that a situation is risky if the level of task complexity is different from the level of coping capacity. For example, if the driving is task is difficult and the operator state is decreased, then risk is probable. In order to identify risk, the level of task complexity as well as the level of coping capacity need to be predicted and compared. As a result, the hypothesis forms a real-time multi-level classification problem, where the dependent variable takes the form of a category representing the difference of task complexity and coping capacity. Task

Complexity variables (X_1) and coping capacity variables (X_2) can be used to identify individual levels of coping capacity and task complexity, and can also be supplemented by other indicators to predict Y. The relationship between the variables and their causal relationship can be depicted in the following flowchart in Figure 1:

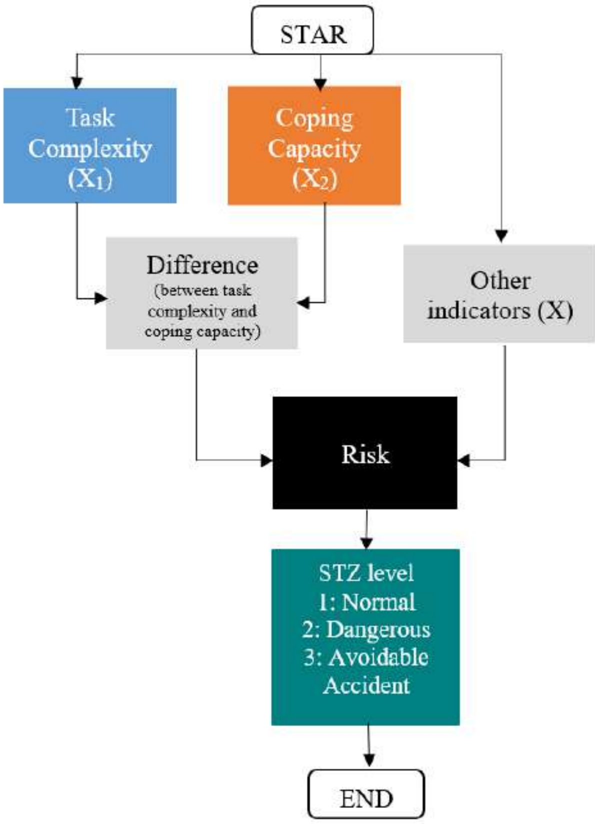

***Figure 1**: The causal relationship between the variables of task complexity and coping capacity*

With regards to the model specification, the raw sensor measurements will be observed. By filtering these raw measurements, the COV indicators will become available, so they will be used to determine the coping capacity and task complexity at each time moment. Hence, the two layers of coping capacity and task complexity depend on the COV indicators. Finally, as the operator's capacity indicates the ability of the driver to operate safely with regards to the task imposed, the operator's capacity depends on the complexity of the task. The proposed DBN structure along with the proposed characteristics to estimate task complexity and coping capacity is depicted in Figure 2.

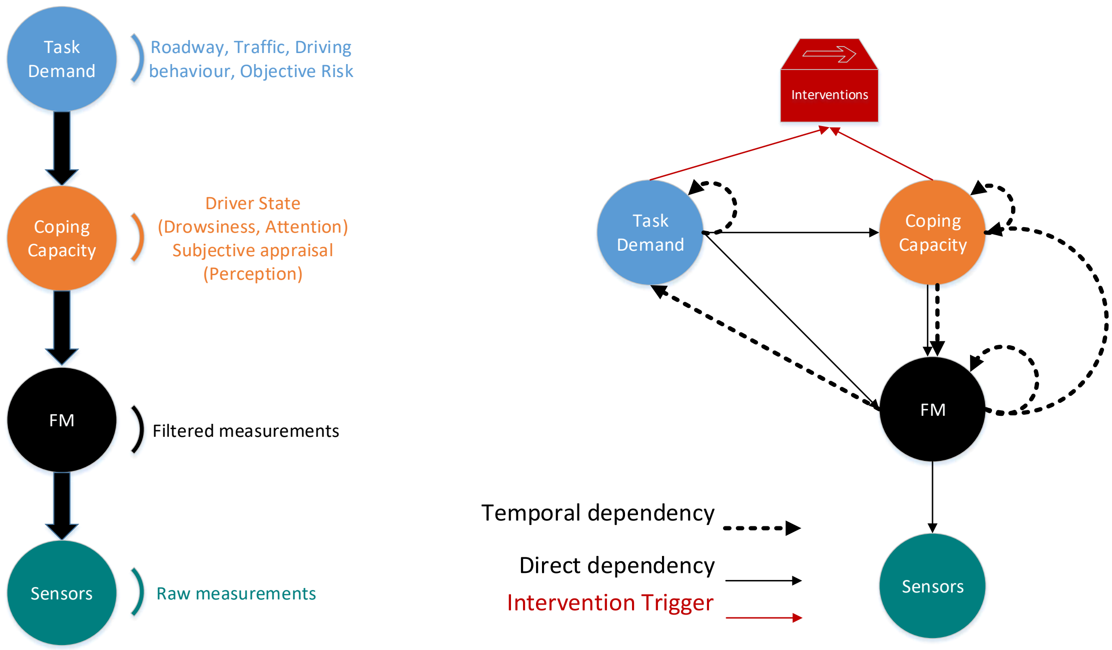

Figure 2:* *The proposed DBN for STZ modelling

The proposed DBN can be described by the joint distribution:

$$P(TC^{0:T}, CC^{0:T}, FM^{0:T}, Z^{0:T}) = P(TC_0, CC_0, FM_0, Z_0)$$
$$\prod_{t=1}^{T} P(TC_t|TC_{t-1}FM_{t-1})\ P(CC_t|TC_tCC_{t-1}FM_{t-1})\ P(FM_t|FM_{t-1}TC_tCC_tCC_{t-1})P(Z_t|FM_t)$$
$$t\ \in\ \mathbb{N}\ and\ t\ \leq T \qquad \textbf{(1)}$$

where:

- TC: Task Complexity
- CC: Coping Capacity
- FM: Filtered COV Measurements
- Z: Raw measurements
- t: current time step
- T: Total time of measurements

The expected task complexity $P(TC_t|TC_{t-1}FM_{t-1})$ is derived from the previous task complexity and the available indicators on environment variables (i.e. time of day, wipers on/off, low visibility indicator, road environment, road geometric configuration and traffic density).

$$P(TC_t|TC_{t-1}FM_{t-1}) = f(Environment, Vehicle\ variables, TC_{t-1}) \qquad \textbf{(2)}$$

Coping capacity $P(CC_t|TC_tCC_{t-1}FM_{t-1})$ can be estimated through functions that correlate the effect of task complexity on coping capacity (Faure et al., 2016) modified by a factor to take the previous coping capacity into account.

$$P(CC_t|TC_tCC_{t-1}FM_{t-1}) = f(Driver,\ TC_t, CC_{t-1}) \qquad \textbf{(3)}$$

The filtered measurements $P(FM_t|FM_{t-1}TC_tCC_tCC_{t-1})$ is the probability of the indicator values based on the current task complexity and coping capacity as well as their previous values and the previous coping capacity can be mapped based on the specific scenarios that will be tested in the simulators. In that way, specific ranges of values or task complexity - and coping capacity-specific measurements along with their corresponding probabilities of appearance can be identified.

For the probability of the raw measurements $P(Z_t|FM_t)$ a sensor model based on Agamennoni et al. (2011) and the Student t-distribution can be followed. In order to identify the different STZ levels, a comparison between the layers of task complexity and coping capacity will be made. The following probability is proposed to be inferred in order to identify avoidable accident or dangerous STZ levels. It should be mentioned that this probability refers to situations that task complexity and coping capacity are beyond normal operations (i.e. increased or high task complexity with decreased or low coping capacity) given the available sensor observations.

$$P(TC \neq normal \cup CC \neq normal \mid Sensors) \qquad \textbf{(4)}$$

Examples of the different STZ levels according to task complexity and coping capacity are highlighted in Table 1. It can be observed that low coping capacity leads to avoidable accident or dangerous phase, decreased coping capacity leads to dangerous or normal phase, while high coping capacity leads to Normal phase, regardless the other layers of task complexity.

***Table 1**: Different STZ levels according to task complexity and coping capacity*

Task Complexity	Coping Capacity	STZ Level
High	Low	Avoidable Accident
High	Decreased	Dangerous
High	High	Normal
Increased	Low	Avoidable Accident
Increased	Decreased	Dangerous
Increased	High	Normal
Low	Low	Dangerous
Low	Decreased	Normal
Low	High	Normal

The likelihood function for Bayesian Networks is the same as in the frequentist inference. More specifically,

$$likelihoodi = \pi(xi)^{yi}(1-\pi(xi))^{(1-yi)} \quad \textbf{(5)}$$

where:

- xi is the covariate vector
- $\pi(xi)$ is the probability of the event for the i^{th} subject which has covariate vector xi
- yi is the multiple dependent variable representing the risk probability which has the outcomes y=0 (STZ: Normal Phase), y=1 (STZ: Dangerous Phase) and y=2 (STZ: Avoidable Accident Phase)

The logistic regression equation is:

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n \quad \textbf{(6)}$$

where:

- β_0 is the intercept
- βi is a coefficient for the explanatory variable xi

In addition, similarly to the frequentist approach, taking the exp(β) provides the odds ratio for one unit change of that parameter.

3.2 Long Short-Term Memory analytical approach

With regards to the modelling approach, it is assumed that there is a specific risk factor (i.e. task complexity or coping capacity) along with the corresponding measurements and metrics for each variable. At each time, a specific risk factor (i.e. STZ levels of each risk factor are known) is targeted but other important variables (e.g. weather conditions, distraction, etc.) can also be used in the same model in order to make the prediction. The entire dataset will be split into train and test set. Based on these indicators, there is a need to predict the risk, and therefore, the time spent in each STZ level (i.e. normal, dangerous, avoidable accident). The problem is a real-time regression problem and can be solved by the LSTM formulation. In order to make sure that the risk calculated is reliable, a good level of forecast accuracy for all the STZ levels should be performed. For instance, if a good prediction for the "avoidable accident" phase can be produced, it should be made clear that a good prediction for the "normal" phase, can be produced as well. This implies that the level of the STZ should be known beforehand, otherwise this hypothesis needs to be supplemented by a classification problem or a clustering one. The flowchart associated with this hypothesis is shown in Figure 3.

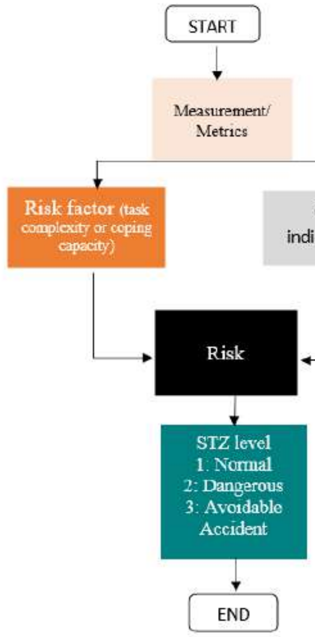

***Figure 3**: The proposed LSTM for STZ modelling*

With regards to the second proposed LSTM model, the problem of defining the STZ levels becomes more straightforward, since LSTMs as a sub-category of Deep Neural Networks act like "black-boxes" (Xu et al., 2013) and thus the only input that needs to be provided to the model are labelled time series data. The proposed approach using LSTMs is given in Figure 4.

In the proposed solution with LSTMs, historical sensor data will be used to extract and select features of the measurements to obtain the most important for STZ level detection. Afterwards, the most important measurements for monitoring the environment the vehicle and the driver become the input to an unsupervised learning algorithm that will group together measurements according to task demand and coping capacity, which, in turn, will act as input for training the LSTM model. After training the LSTM model with the labelled time-series data, the available real-time sensor data will be used as input for the model to predict the STZ level in the subsequent time.

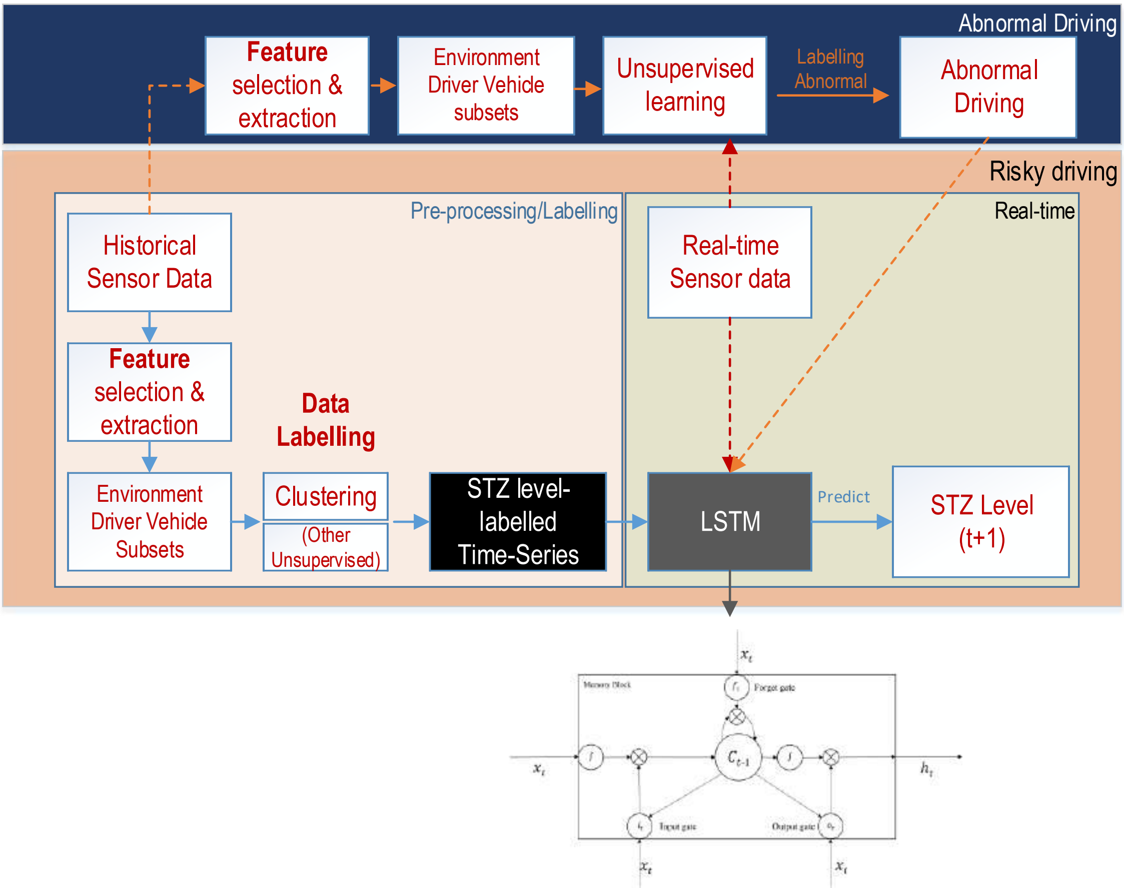

***Figure 4**: STZ modelling using LSTMs*

4. Results

In order to model the concept of the i-DREAMS project several parameters were examined. Particular emphasis was given to average speed and a new variable was created taking into account the different levels of STZ. Thus, the dependent variable was the STZ_speed which was divided into three levels (i.e. normal phase:0, dangerous phase: 1, avoidable accident phase: 2). Then, the correlation of the independent variables (i.e. TTC – time to collision with vehicle ahead, Headway - time headway to vehicle ahead in same lane, Distance travelled - distance driving, HandsOnEvent - whether hands are on the steering wheel, FatigueEvent - KSS score, ME_ForwardCollisionWarning - whether forward collision warning is active and ME_LaneDepartureWarningActive - whether lane departure warning is active) was investigated. No strong correlation among these indicators was identified, as shown in Table 2.

***Table 2**: Correlation of independent variables*

	Headway	TTC	Distance.travelled	HandsOnEvent	FatigueEvent	ME_ForwardCollision Warning	ME_LaneDeparture WarningActive
Headway	1,000	0,000	-0,081	-0,098	-0,124	-0,005	-0,024
TTC	0,000	1,000	0,000	-0,010	-0,078	0,001	-0,007
Distance.travelled	-0,081	0,000	1,000	0,024	0,097	0,004	0,000
HandsOnEvent	-0,098	-0,010	0,024	1,000			
FatigueEvent	-0,124	-0,078	0,097		1,000		
ME_ForwardCollision Warning	-0,005	0,001	0,004			1,000	-0,008
ME_LaneDeparture WarningActive	-0,024	-0,007	0,000			-0,008	1,000

A feature importance algorithm extracted from XGBoost was used in order to evaluate the significance of variables on forecasting STZ and select the most appropriate independent variables. Headway, distance travelled and TTC were the most important factors of all examined indicators, while the parameters of forward collision warning and lane departure warning active were less significant. Table 3 provides the feature importance of independent variables based on XGBoost algorithm implemented.

***Table 3**: Feature importance of independent variables*

	Variables	Importance
1	Headway	0.4628857977881
2	Distance.travelled	0.3689770387145
3	TTC	0.1681097785940
4	ME_ForwardCollisionWarning	0.0000203474815
5	ME_LaneDepartureWarningActive	0.0000004199277

A short dataset of 10,000 rows was used and a Neural Network model was implemented. As presented in Figure 5, in this study, there are three neurons in the input layer (i.e. headway, distance.travelled, TTC) and one neuron in the output layer (STZ).

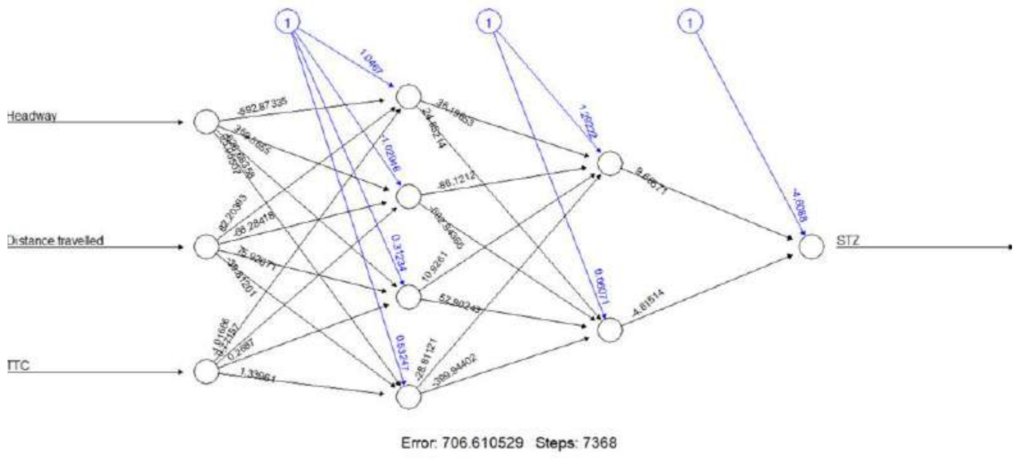

__Figure 5__: The multi-layer Neural Network model layout for STZ

Then, a table of confusion which contains two rows and two columns that reports the number of false positives, false negatives, true positives, and true negatives was created. This allows more detailed analysis than mere proportion of correct classifications (e.g. accuracy). In particular, negative class refers to the Normal phase, positive class refers to the Dangerous phase, while no instances for Avoidable Accident Phase were detected, as shown in Table 4.

__Table 4__: Confusion data matrix

	Predicted Class	
Actual (True) Class	0	1
0	True Negative (TN) **157**	False Positive (FP) **797**
1	False Negative (FN) **0**	True Positive (TP) **1546**

With regards to the assessment of classification model, average classification accuracy represents the proportion of correctly classified observations, while precision, recall and specificity, which are three major performance metrics, describe a predictive classification model. F1-score is calculated in order to find a balance between precision and recall. The geometric mean G-mean is a product of the prediction accuracies for both classes, i.e recall: accuracy on the positive examples, and specificity: accuracy on the negative examples. Lastly, the false positive rate refers to the expectancy of the false positive ratio (i.e. the probability of falsely rejecting the null hypothesis for a particular test). Table 5 provides the assessment of classification model.

***Table 5**: Assessment of classification model*

Accuracy	Precision	Recall	Specificity	f1-score	G-Means	FP Rate
0,68	0,66	1	0,16	0,80	0,81	0,84

5. Conclusions

Following a thorough literature review of models dealing with driver behavior and collision risk modelling in real-time, the most prominent approaches were found to be Dynamic Bayesian Networks or DBNs (a probabilistic graphical time-series model) and Long Short-Term Memory networks or LSTMs (a deep neural network formulation). These two dynamic approaches were chosen due to their efficiency and flexibility in real-time predictions and were found to be suitable for prediction of continuous indicators of risk (e.g. fatigue, speed, time headway, distraction, harsh acceleration). Such a continuous indicator of risk may be the result of combining discrete indicators of risk for different risk factors (which will help validate STZ) or may be the time that is spent in each phase of STZ (which will help tuning the frequency/pitch/presentation of warnings).

While the review of the analytical methods presented previously provided a good understanding of the potential modelling candidates in i-DREAMS and the final selected methods seem plausible, there are still some open issues that need to be considered for model selection. For example, the suggested methods may be confronted with additional limitations considering the ongoing discussions on the different types of data being collected in i-DREAMS. In addition, a number of new limitations have been identified with additional deeper investigations into these methods. For example, it is noted that LSTM is not able to incorporate the inter-relationship between variables into real-time predictions.

Preliminary results indicated a strong relationship between STZ and the independents variables of headway, distance.travelled and TTC. However, it should be highlighted that when more data are available, the most crucial risk indicators of task demand and coping capacity will be extracted and the initial hypothesis described in the paper will be confirmed.

6. References

Agamennoni, G., Nieto, J. I., & Nebot, E. M. (2011). A Bayesian approach for driving behavior inference. In 2011 IEEE Intelligent Vehicles Symposium (IV) (pp. 595-600). IEEE.

Bao, J., Liu, P., & Ukkusuri, S. V. (2019). A spatiotemporal deep learning approach for citywide short-term crash risk prediction with multi-source data. Accident Analysis & Prevention, 122, 239-254.

Bouhoute, A., Berrada, I., & El Kamili, M. (2014). A formal model of human driving behavior in vehicular networks. In 2014 International Wireless Communications and Mobile Computing Conference (IWCMC) (pp. 231-236). IEEE.

Chong, L., Abbas, M. M., Flintsch, A. M., & Higgs, B. (2013). A rule-based neural network approach to model driver naturalistic behavior in traffic. Transportation Research Part C: Emerging Technologies, 32, 207-223.

El Hatri, C., & Boumhidi, J. (2018). Fuzzy deep learning based urban traffic incident detection. Cognitive Systems Research, 50, 206-213.

Girma, A., Yan, X., & Homaifar, A. (2019). Driver identification based on vehicle telematics data using LSTM-recurrent neural network. In 2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI) (pp. 894-902). IEEE.

Imkamon, T., Saensom, P., Tangamchit, P., & Pongpaibool, P. (2008). Detection of hazardous driving behavior using fuzzy logic. In 2008 5th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (Vol. 2, pp. 657-660). IEEE.

Kanaan, D., Ayas, S., Donmez, B., Risteska, M., & Chakraborty, J. (2019). Using Naturalistic Vehicle-Based Data to Predict Distraction and Environmental Demand. International Journal of Mobile Human Computer Interaction (IJMHCI), 11(3), 59-70.

Katrakazas, C., Quddus, M., & Chen, W. H. (2019). A new integrated collision risk assessment methodology for autonomous vehicles. Accident Analysis & Prevention, 127, 61-79.

Lefèvre, S., Laugier, C., & Ibañez-Guzmán, J. (2012, June). Risk assessment at road intersections: Comparing intention and expectation. In 2012 IEEE Intelligent Vehicles Symposium (pp. 165-171). IEEE.

Li, M., Liu, J., Li, J., & Kim, B. U. (2014). Bayesian modeling of multi-state hierarchical systems with multi-level information aggregation. Reliability Engineering & System Safety, 124, 158-164.

McDonald, A. D., Ferris, T. K., & Wiener, T. A. (2020). Classification of driver distraction: A comprehensive analysis of feature generation, machine learning, and input measures. Human factors, 62(6), 1019-1035.

Merrikhpour, M., Donmez, B., & Battista, V. (2014). A field operational trial evaluating a feedback–reward system on speeding and tailgating behaviors. Transportation research part F: traffic psychology and behaviour, 27, 56-68.

Fuller, R. (2000). The task-capability interface model of the driving process. Recherche - Transports - Sécurité, 66, 47–57.

Fuller, R. (2005). Towards a general theory of driver behaviour. Accident analysis & prevention, 37(3), 461-472.

Fuller, R. (2011). Driver control theory: From task difficulty homeostasis to risk allostasis. In Handbook of traffic psychology (pp. 13-26). Academic Press.

Shankar, V., Jovanis, P. P., Aguero-Valverde, J., & Gross, F. (2008). Analysis of naturalistic driving data: prospective view on methodological paradigms. Transportation research record, 2061(1), 1-8.

Tandon, P., Huggins, P., Maclachlan, R., Dubrawski, A., Nelson, K., & Labov, S. (2016). Detection of radioactive sources in urban scenes using Bayesian Aggregation of data from mobile spectrometers. Information Systems, 57, 195-206.

Wang, X., & Xu, X. (2019). Assessing the relationship between self-reported driving behaviors and driver risk using a naturalistic driving study. Accident Analysis & Prevention, 128, 8-16.

Xu, C., Wang, W., & Liu, P. (2012). A genetic programming model for real-time crash prediction on freeways. IEEE Transactions on Intelligent Transportation Systems, 14(2), 574-586.

Xue, Q., Wang, K., Lu, J. J., & Liu, Y. (2019). Rapid driving style recognition in car-following using machine learning and vehicle trajectory data. Journal of advanced transportation, 2019.

Zhu, X., Yuan, Y., Hu, X., Chiu, Y. C., & Ma, Y. L. (2017). A Bayesian Network model for contextual versus non-contextual driving behavior assessment. Transportation research part C: emerging technologies, 81, 172-187.

Zou, Q., Li, H., & Zhang, R. (2018). Inverse Reinforcement Learning via Neural Network in Driver Behavior Modeling. In 2018 IEEE Intelligent Vehicles Symposium (IV) (pp. 1245-1250). IEEE.