# Modeling the concept of a Safety Tolerance Zone: State-of-the-art and proposed alternatives

**Eva Michelaraki[1*], Christos Katrakazas[1], Amir Pooyan Afghari[2], Eleonora Papadimitriou[2], Christelle Al Haddad[3], Kui Yang[3], Constantinos Antoniou[3], Tom Brijs[4], George Yannis[1]**

[1]*National Technical University of Athens, Department of Transportation Planning and Engineering, 5 Heroon Polytechniou str., 15773, Athens, Greece*
[2]*Delft University of Technology, Faculty of Technology, Policy and Management, Jaffalaan 5, 2628 BX, Delft, the Netherlands*
[3]*Technical University of Munich, Chair of Transportation Systems Engineering, Arcisstrasse 21, 80333, Munich, Germany*
[4]*UHasselt, School for Transportation Sciences, Transportation Research Institute (IMOB), Agoralaan, 3590 - Diepenbeek, Belgium*

## Abstract

Within a transport system, a driver can be viewed as a (technology assisted) human operator, self-regulating control over

a vehicle in the context of crash avoidance. Based on the Fuller's task capability model, a Safety Tolerance Zone (STZ) concept has been recently developed in the i-DREAMS naturalistic driving study attempting to describe the point at which self-regulated control is considered safe. This paper aims to explicitly present the practical conceptualization of the STZ in order to transition from a theoretical framework to a practical implementation and a fully functional methodology. A thorough literature review of analytical models dealing with driver behavior and collision risk, both in real-time and post-trip, is first conducted and the most suitable modelling approaches for the STZ are selected. Specific machine learning algorithms and statistical models are then examined in order to relate driving performance with the probability of a rare event and the crash severityamong which, the most prominent approaches were initially found to be Dynamic Bayesian Networks (DBNs; a probabilistic graphical time-series model) and Long Short-Term Memory networks (LSTMs; a deep neural network formulation). Furthermore, Structural Equation Models (SEMs) and Discrete Choice Models (DCMs) were also deemed suitable for the i-DREAMS concept, providing 'static' or post-trip predictions, in contrast with DBNs and LSTMs which work dynamically (i.e. in real-time). For each of the aforementioned methods or techniques, a brief description of their underpinning procedure was presented, followed by their application for the identification of the STZ levels. The testing, calibration and enhancement of the mathematical models during the i-DREAMS simulation and on-road experiments can assure a sufficient and efficient data analysis, as well as timely initiation of the safety interventions.

Keywords: i-DREAMS project; Safety Tolerance Zone; real-time; post-trip; modeling.

[1] Corresponding author. Tel.: +30-210-772-1265;
E-mail address: evamich@mail.ntua.gr

## 1. Introduction

Within a transport system, a driver can be viewed as a (technology assisted) human operator, self-regulating control over a vehicle in the context of crash avoidance. Based on the Fuller's task capability model [2, 3], driving difficulty is inversely related to the difference between driving task demand and the driver's task capability. As a result, safe driving can be regarded as the practice of using driving strategies that minimize the risk on the road and thus help avoiding critical events (e.g. crashes) by predicting hazardous situations on the road. Conversely, dangerous driving is found when an individual's driving falls below the expected level of a careful and competent driver [1]. It can also be classed as dangerous driving scenarios where the vehicle being driven is in a dangerous condition and not suitable to be on public roads. It is worth noting that traffic safety conditions involve the quality of the road defined by the level of crashes and reflecting the degree of safety of traffic participants from road traffic crashes as well as their consequences. At the same time, road traffic safety can be understood as the result of the safe interaction of participants between themselves and the environment. Thus, when assessing the traffic safety on the road environment, driver's physiological and psychological capabilities should be taken into consideration.

The i-DREAMS naturalistic driving study project aims to establish a framework for the definition, development, testing and validation of a context-aware safety envelope for driving in a Safety Tolerance Zone (STZ), within a smart Driver, Vehicle and Environment Assessment and Monitoring System (i-DREAMS). Taking into account driver background factors and risk indicators associated with the driving performance as well as the driver state and driving task complexity parameters, a continuous real-time assessment will be made in order to monitor and determine if drivers are within acceptable boundaries of safe operation. Furthermore, delayed safety-oriented interventions and post-trip feedback aimed at enhancing the knowledge, attitudes and perceptions will be provided.

The concept of the STZ is the core concept of the i-DREAMS project[2] and attempts to describe the point at which self-regulated control is considered safe. It is based on Fuller's Task Capability Interface Model [2, 3] which states that loss of control occurs when the demand of a driving task outweighs the operator's capability. The STZ is subdivided in three levels of safety, namely: the 'Normal Driving phase', the 'Danger phase' and the 'Avoidable Accident phase'. The Normal Driving phase represents the conditions in which a crash is unlikely to occur, i.e. the crash risk is low. During this phase, drivers can successfully adapt their behaviour in order to meet the task demand. The Danger phase is characterized by changes in normal driving that indicate that a crash may occur, therefore, the crash risk is increased. Finally, the Avoidable Accident phase occurs when a collision scenario develops but there is still time for the driver to intervene and avoid the crash. The need for action is more urgent than in the Danger phase and if the driver does not adapt their behaviour to the current circumstances, a crash is very likely to occur. The different driving phases of the STZ along with their definition are summarized in Table 1.

**Table 1: Different driving phases of the Safety Tolerance Zone (STZ)**

| Phases of STZ | Description |
|---|---|
| Normal Driving phase | Crash risk is minimal |
| Danger phase | Risk of crash increases as internal / external events occur |
| Avoidable Accident phase | Crash is very likely to occur if no preventative action taken by driver |

The fundamental goal of the i-DREAMS platform is to keep the driver in the Normal Driving phase for as long as possible and, where this is not possible, to prevent the transition from the danger to the Avoidable Accident phase. To this end, the platform combines both real-time and post-trip interventions which, respectively, aim to nudge and coach the driver. It is worth mentioning that the platform is a warning based driver assistance system which does not actively intervene physically in any way with the driving task. In order to estimate in which STZ phase the driver is in and which interventions should be provided, the i-DREAMS platform uses two modules. Firstly, it uses the monitoring module, which takes measurements related to the context, the operator and the vehicle in order to derive the demands of the driving task and the driver's ability to cope with these demands. This module estimates at which stage of the STZ the driver is operating at any given time. Secondly, the in-vehicle intervention module is responsible for keeping the driver within the Normal phase of the STZ, either by providing a warning or alert during the trip (i.e. real-time intervention) or by providing feedback about the journey after the completion of the driving task (i.e. post-trip intervention).

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

It should be noted that data analysis consists a pivotal part of this project for achieving its objectives as well as the methods for data analysis highly depend on the collected data. In order to model the STZ, the available data and the potential outcomes 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 not only in real-time but also post-trip and the measurements of task complexity (e.g. weather, road layout, time of the day) and coping capacity (e.g. distraction, fatigue, drowsiness) are going to be sequential. Moreover, as the STZ is the "trigger" for real-time and post-trip interventions, both dynamic and static modeling approaches need to be examined.

This study aims to explicitly describe the practical conceptualization of the STZ in order to transition from a theoretical framework to a practical implementation and a fully functional methodology. In order to fulfil the purpose of this research, the most suitable mathematical models to realize the STZ, applied in the i-DREAMS risk analyses, both from a real-time and post trip perspective are provided. The paper is structured as follows. In the beginning, the overall objective of the i-DREAMS project and the aim of the current research is presented. Then, the most prominent approaches are detailed and a brief description of their underpinning procedure is given. Subsequently, initial insights into analyses are highlighted. Lastly, overall conclusions and practical considerations concerning the modelling of the STZ are provided in order to assist practitioners and researchers.

## 2. Methodology

To date, predicting driving behavior by employing mathematical driver models, obtained directly from the observed driving-behavior data, has gained much attention [4]. A few models have been used to address road safety and the estimation of driving behavior, many of which in the context of experimental studies, including naturalistic driving or field operational trials and driving simulator studies. A review of safety models can be found in [5], where the authors noted inconsistency in the language of safety models and emphasized that additional factors should be investigated, such as the effect of organizational culture, emergency responses, the health system and economic influences on-road safety. In their opinions, there are models with potential to improve road safety, but yet to be applied.

In order to obtain the most suitable modeling approaches for the STZ, a thorough literature review of models dealing with driver behavior and collision risk, both in real-time and post-trip, was implemented. Several state-of-the-art methodological approaches that enable the modeling of crash risk were evaluated. In addition, specialized machine learning and statistical algorithms were examined in order to relate driving performance with the probability and the severity of a crash, among which four methods have been selected to be used in i-DREAMS: Dynamic Bayesian Networks (DBNs), Long Short-Term Memory (LSTMs) deep neural networks, Discrete Choice Models (DCMs) and Structural Equation Models (SEMs). Each of the aforementioned methods had strengths and limitations, making them 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. 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. Figure 1 provides the flowchart of the proposed approach.

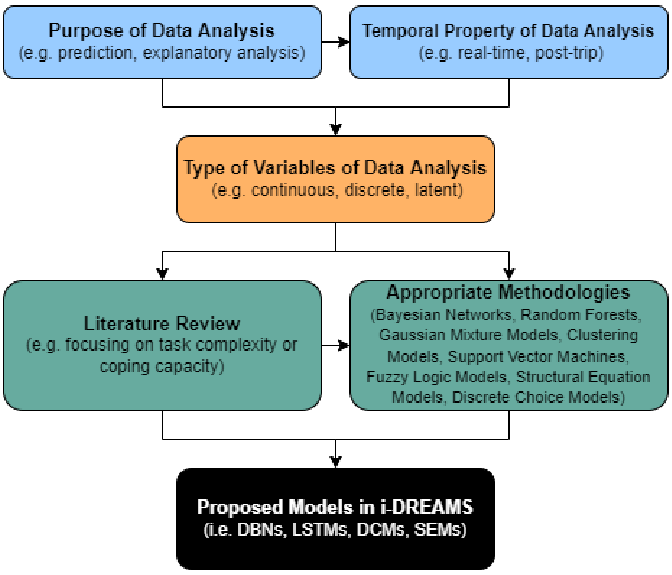


**Figure 1: Flowchart of the proposed approach**

## 3. Results

This section provides the mathematical formulation of the STZ model according to these popular methodologies, in order to provide flexibility in the practical implementation of the STZ estimation algorithm. To this end, a brief description of each algorithm is presented, followed by an explicit description of the proposed models.

### 3.1 Dynamic Bayesian Networks (DBNs)

Dynamic Bayesian Networks (DBNs) are the most appropriate method to model discrete indicators of risk. A DBN is a directed acyclic graphical model that can express a joint probability distribution of a large set of variables [6]. Usually, DBNs are utilized for learning causal relationships and hence are ideal for investigating the effect of interventions by combining new and prior knowledge data. The core of DBNs is the attempt to infer a "hidden" state based on a group of available observations.

The variables monitored by the i-DREAMS platform concerning task complexity and coping capacity (i.e. driver and vehicle state); thus the raw sensor measurements are observed. By filtering these raw measurements, the Context-Operator-Vehicle (COV) indicators become available, so they are used to determine the coping capacity and task complexity at each time moment. Hence, the two layers of task complexity and coping capacity 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.

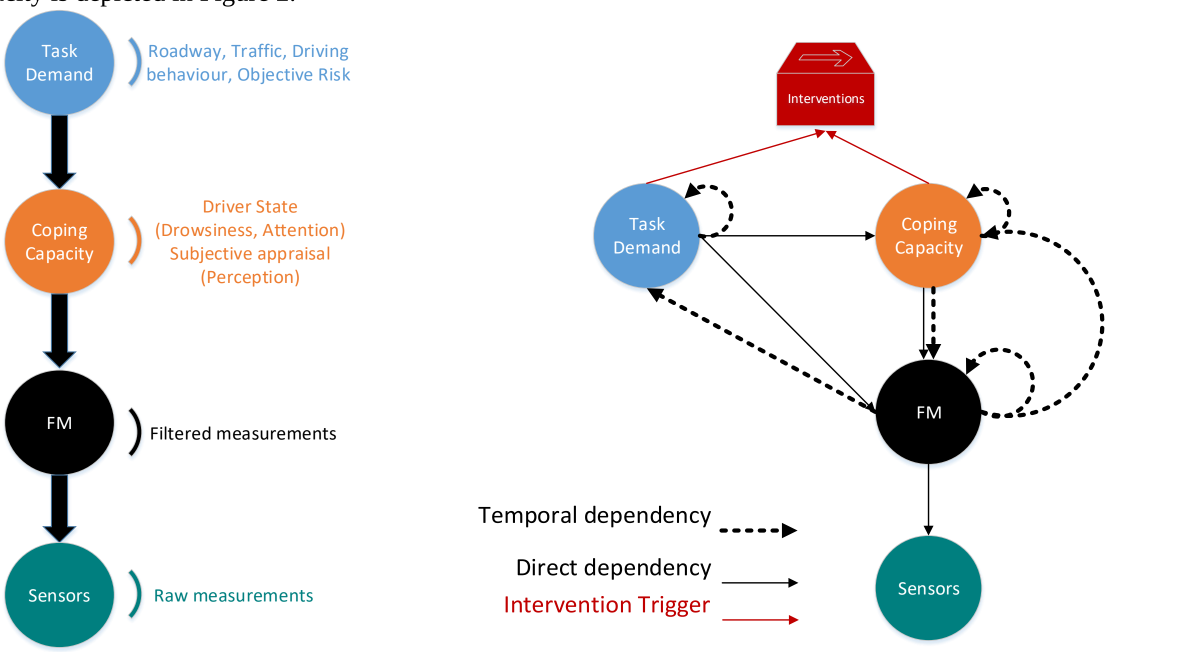


**Figure 2: The proposed DBN for STZ modeling**

The proposed DBN can be described by the joint distribution:

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

where TC refers to task complexity, CC refers to coping capacity, FM is filtered COV measurements, Z is raw measurements, t is current time step and T is total time of measurements.

**Task Complexity:** 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).

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

**Coping Capacity:** 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 [7] modified by a factor to take the previous coping capacity into account.

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

**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.

**Raw measurements:** For the probability of the raw measurements $P(Z_t|FM_t)$ a sensor model based on Agamennoni et al. [8], 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.

$$P(TC \neq normal \cup CC \neq normal\ |Sensors) \quad (4)$$

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

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

where $xi$ is the covariate vector, $\pi(xi)$ is the probability of the event for the $i^{th}$ subject which has covariate vector $xi$ and $yi$ 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:

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

where $\beta_0$ is the intercept and $\beta i$ is a coefficient for the explanatory variable $xi$

## 3.2 Long Short-Term Memory Networks (LSTMs)

Long Short-Term Memory Networks (LSTMs) are suitable for continuous indicators of risk. These models are a special kind of Recurrent Neural Network (RNN), capable of learning long-term dependencies [9]. They work tremendously well on a large variety of problems, and are now widely used. LSTMs are explicitly designed to avoid the long-term dependency problem. Remembering information for long periods of time is practically their default behavior and not something they struggle to learn. All recurrent LSTMs have the form of a chain of repeating modules of neural network.

LSTMs use "memory block" in the hidden unit to capture the long-term dependencies that may exist in the data [10]. This memorizing capability of LSTM has shown the best performance across many time-series tasks, such as activity recognition, video captioning and language translation. The cell state (memory block) of LSTM has one or more memory cells that are regulated by structures called gates, which control the addition of new sequential information and the removal of useless ones to and from memory, respectively. Gates are a combination of sigmoid activation functions and a dot (scalar) multiplication operation, and they are used to control information that passes through the network.

The problem of defining the STZ levels becomes more straightforward, since LSTMs as a sub-category of Deep Neural Networks act like "black-boxes" [11] 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 3.

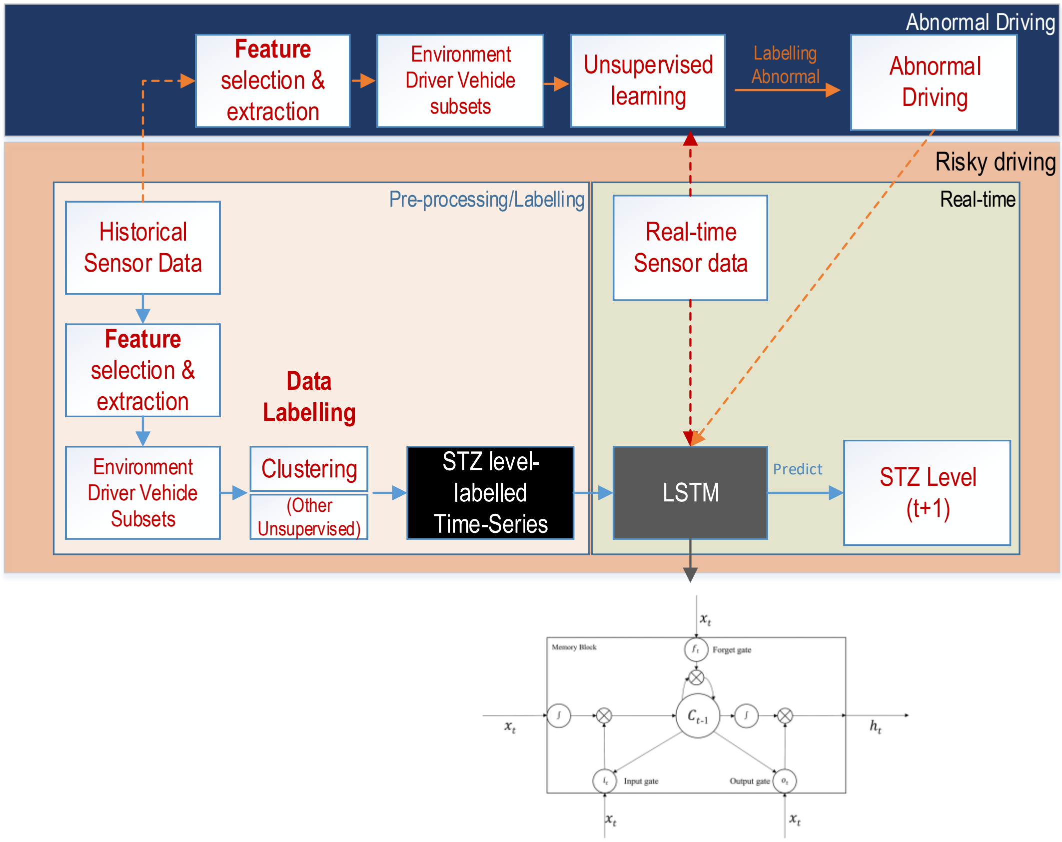


**Figure 3: The proposed LSTM for STZ modeling**

## 3.3 Discrete Choice Models (DCMs)

Discrete Choice Models (DCMs) are the most common statistical approaches to model discrete indicators of risk. These models rely on the maximum utilisation theory in economics [12] stating that among many alternatives, individuals select the alternative (i.e. discrete category) that maximises their utility. Thus, the first step in formulating DCMs is defining a utility for each discrete alternative. This utility will not have a physical meaning but is rather an auxiliary term to determine the probability of selecting an alternative over the other alternatives. Depending on the nature of the discrete variable being nominal (e.g. occurrence of a rare event/no rare event) or ordered (i.e. STZ levels), DCMs can take the form of either unordered or ordered.

### *3.3.1 Unordered Discrete Choice Models*

Let $Y$ be a discrete dependent variable with $s$ nominal multiple categories (e.g. s=0: high crash severity, s=1: medium crash severity, s=2: low crash severity). The utility of the $s^{th}$ category ($U_s$) is stated as:

$$U_s = \beta_s X_s + \varepsilon_s \qquad (7)$$

where $\beta_s$ are estimable parameters (including the intercept), $X_s$ are explanatory variables (e.g. sociodemographic factors, vehicle type, etc.) and $\varepsilon_s$ is the random error term assumed to be identically and independently distributed across observations and describing the random part of the utility. Assuming that $\varepsilon_s$ is generalised extreme value distributed [13], the probability of the s$^{th}$ category can be presented as:

$$P(\mathrm{Y} = s) = \frac{e^{(\beta_s X_s)}}{\sum_{j=1}^{S} e^{(\beta_j X_j)}} \qquad (8)$$

The likelihood of occurring the $s^{th}$ category across all individuals can then be determined by the product of the above equation over the entire observations. This model is referred to as the multinomial logit discrete choice model in the statistical and econometrics literature [12].

When the dependent variable has only two categories (s=2), the above model reduces to the binary logit model. This model can be used to determine the probability of a rare event (e.g. a near-miss). Additional variants of this model such as Integrated Choice and Latent Variable (ICLV) binary logit model may also be useful depending on the hypothesis between risk, task complexity and coping capacity.

### *3.3.2 Ordered Discrete Choice Models*

Let $Y$ be a discrete dependent variable with s ordered categories (e.g. $S = 1$ if Normal Driving phase, $S = 2$ if Dangerous phase, and $S = 3$ if Avoidable Accident phase). In ordered discrete choice models, the actual category of the dependent variable ($Y_s$) is associated with an underlying latent variable ($Y_s{}^*$). This latent variable is then mapped to the actual categories by thresholds ( $\tau$ ) and using the following linear function:

$$Y_s^* = \kappa X_s + \delta_i \quad \text{and} \quad Y_s = S \quad \text{if} \quad \tau_{s-1} < Y_s^* < \tau_s \qquad (9)$$

where $\kappa$ is the vector of parameters, $X_s$ is the vector of covariates for the $s^{th}$ category and $\delta_i$ is the random error term. To estimate the latent propensity of the dependent variable, it is assumed that:

$$E(Y_s|X_s) = H_s(.),\ 0 \leq H_s(.) \leq 1,\ \sum_{s=1}^{S} H_s = 1 \qquad (10)$$

where $H_s(.)$ is the probability density function for the discrete category $s$. Depending on the distributional assumption for the probability of error terms, $H_s(.)$ can take standard normal or standard logistic probability density functions for the ordered probit or ordered logit discrete choice models, respectively. Similar to the unordered models, additional variants of this model such as such as Integrated Choice and Latent Variable (ICLV) ordered probit (logit) model may also be useful depending on the hypothesis between risk, task complexity and coping capacity.

## 3.4 Structural Equation Models (SEMs)

Structural Equation Models (SEMs) are suitable for continuous indicators of risk. These models represent a natural extension of a measurement model and establish a mature statistical modelling framework [14]. In particular, they are designed to deal with several difficult modelling challenges, including cases in which some variables of interest to a researcher are unobservable or latent and are measured using one or more exogenous variables, endogeneity among variables, and complex underlying social phenomena. SEMs are widely used for modelling complex and multi-layered relationships between observed and unobserved variables. Observed variables are objectively measurable, whereas unobserved variables are latent constructs - analogous to components in a factor/principal component analysis. SEMs have two components: a measurement model and a structural model. The measurement model is used to determine how well various observable exogenous variables can measure (i.e. load on) the latent variables, as well as the related measurement errors. The structural model is used to explore how the model variables are inter-related, allowing for both direct and indirect relationships to be modelled. In this sense, SEMs differ from ordinary regression techniques in which relationships between variables are strictly.

According to the i-DREAMS concept of the STZ, it is hypothesized that latent risk is measured by a composite variable consisting of all risk factors (e.g. Y:fatigue, loss of sleep, hands on wheel or mobile phone use, speeding, harsh acceleration, harsh deceleration, harsh cornering, lane departure warning, illegal overtaking warning, forward collision warning, vulnerable road user warning), and latent task complexity and latent coping capacity predict the latent risk. Latent task complexity and latent coping capacity are also measured by observed indicators.

The proposed path diagram using SEM is provided in Figure 4.

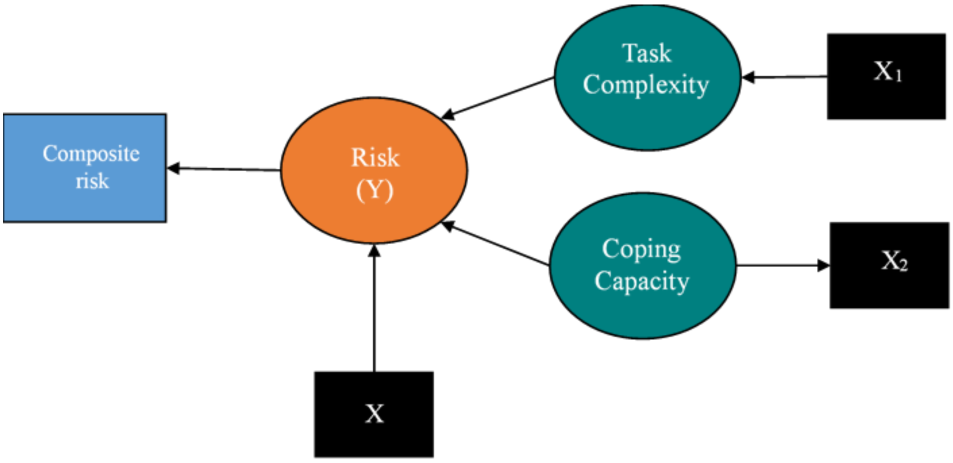


**Figure 4: The proposed SEM for STZ modeling**

SEMs are estimated using ordinary least squares (OLS) approach. Let $Y_i$ be a continuous indicator of risk. A structural equation modelling approach is used to correlate this dependent variable to the independent variables. As previously mentioned, the SEM consists of two components: a structural equation and measurement equations. The structural equation is a regression model capturing the relationship between variables:

$$Y_i = \beta X_i + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i \quad (11)$$

where $\beta_i$ are estimable parameters (including the intercept), $X_i$ are explanatory variables (e.g. demographics, coping capacity and task complexity) and $\varepsilon_i$ is the random error term assumed to be normally distributed across observations and describing the random part of the structural equation.

The measurement equations, on the other hand, are concerned with how well various measured exogenous indicators measure latent variables. In other words, and in estimating the above structural equation, the latent variables (e.g. latent risk, latent task complexity, latent coping capacity) can be measured (i.e. measurement equation) using a linear additive combination of certain observed indicators. However, many of these indicators often have high autocorrelation with one another.

To address this problem, the Principal Component Analysis (PCA) can be used to summarise the observed indicators into orthogonal variables (i.e. principal components) that are not correlated. The PCA creates a set of new variables, referred to as principal components (PC), each of which is a linear and orthogonal combination of the original variables in such a way that each orthogonal combination captures the maximum variability in the original set of variables and has the minimum autocorrelation with other linear combinations. For instance, further explanations with regards to the combinations are shown below:

$$PC_{risk} = w_{KSS} KSS + w_{LOS} Loss\ of\ Sleep + w_{HW} hands\ on\ wheel + \cdots \quad (12)$$
$$PC_{Cop} = w_{att} ATT + w_{agress} Agressions + w_{SN} Subjective\ Norms + \cdots \quad (13)$$
$$PC_{Task} = w_{WP} Wipers + w_{day} Day\ Time + w_{vis} Visibility + \cdots \quad (14)$$

where $w_n$ are weights (i.e. factor loadings) can be obtained by applying the orthogonal transformation and finding the Eigenvectors and Eigenvalues of the Spearman correlation matrix of the original set of explanatory variables.

The principal components are then arranged based on their decreasing contribution to the total variance of the original set of explanatory variables: the first principal component explains the highest variability in the explanatory variables; the second principal component explains the second-highest variability in the explanatory variables, and so forth (the cumulative contribution of all principal components is equal to one). These principal components can then be used in the analysis as indicators of the original latent variables. The number of principal components to be used in the model depends on the specific research objective, though the common practice is to use all principal components with Eigenvalues greater than one [15]. Assuming that $\varepsilon_i$ is normally distributed, the structural equation can be estimated using generalised least squares or maximum likelihood estimation approaches.

## 4. Discussion

A variety of analytical methods and potential modeling approaches has been reviewed, among which four methods have been selected to be used in i-DREAMS: Dynamic Bayesian Networks (DBNs), Long Short-Term Memory (LSTMs) deep neural networks, Discrete Choice Models (DCMs) and Structural Equation Models (SEMs). Each of the aforementioned 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. 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.

When the purpose of data analysis is the prediction of risk (e.g. prediction of the STZ phases), the data should be analyzed in real-time because the predictions in i-DREAMS aim to provide the basis for triggering (real-time) in-vehicle interventions. Prediction of risk after the trip has completed may not be useful in i-DREAMS. As such, the arrow of the post-trip analysis has been turned off for the prediction purpose.

Machine learning algorithms are found to be proper analytical methods for real-time data analysis. However, the type of these algorithms to be used certainly depends on the type of risk indicators being discrete or continuous. The Dynamic Bayesian Network models are suitable for prediction of discrete indicators of risk, while the Long Short-Term Memory and deep neural networks are suitable for prediction of continuous risk. 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). Although some types of statistical models such as the Dynamic Discrete Choice Models (DDCM) may be a good alternative for prediction of discrete indicators of risk, they are prone to big data (i.e. data in real-time) and so their applicability needs to be confirmed via empirical testing.

When the purpose of data analysis is explanatory analysis, the data should be analyzed after the trip has been completed, because the explanatory analysis in i-DREAMS is primarily done for identifying relationships between driving behavior (at an aggregate level) and risk. As such, the arrow of real-time is for now turned off for explanatory analysis. However, there may be sufficient motivation for investigating this turned-off arrow for scientific research. For example, investigating the inter-relationship between risk and coping capacity in real-time and finding whether such an inter-relationship can influence real-time predictions could be another research direction.

Statistical models are suitable for explaining the underlying mechanisms of risk and so are proper analytical methods for post-trip data analysis. However, the type of statistical models to be used depends on the type of risk indicators too. Structural Equation Models (SEMs) are only suitable for continuous dependent variables i.e. risk indicators. They are estimated using ordinary least squares (OLS) approach (the equivalent of SEM in maximum likelihood estimation approach is referred to the latent variable models). When the dependent variable is discrete, Discrete Choice Models (ordered or nominal) are needed.

While this literature review provided a good understanding of the potential modeling candidates in i-DREAMS and the selected models seem plausible, there are still some open issues that need to be considered for model selection. Specifically, the suggested models may be confronted with additional limitations considering the different types of data being collected in i-DREAMS. Additionally, several new limitations have been identified with additional deeper investigations into these models. For instance, it is noted that LSTM is not able to incorporate the inter-relationship between variables into real-time predictions (endogeneity) and SEM is not suitable for analyzing discrete dependent variables.

As a result, and prior to further applying the selected mathematical models, it seems necessary to map these models to the research questions in i-DREAMS. The mapping of the models to research questions depends on three dimensions for data analysis in i-DREAMS: (1) the purpose of data analysis –being prediction or explanatory analysis, (2) the time element of data analysis –being real-time or post-trip, and (3) the variable type of risk indicators –being discrete or continuous (as it may be necessary to test alternative definitions of risk in addition to the three-level STZ definition). The mathematical model to be used in i-DREAMS depends on a combination of these three dimensions.

All in all, considering risk as a dependent variable in i-DREAMS, the type of mathematical model to be used for data analysis highly depends on the definition of risk adopted in each case. A schematic overview of the proposed mathematical models (DBN, LSTM, DCM and SEM) to be considered for the analysis is given in Figure 5.

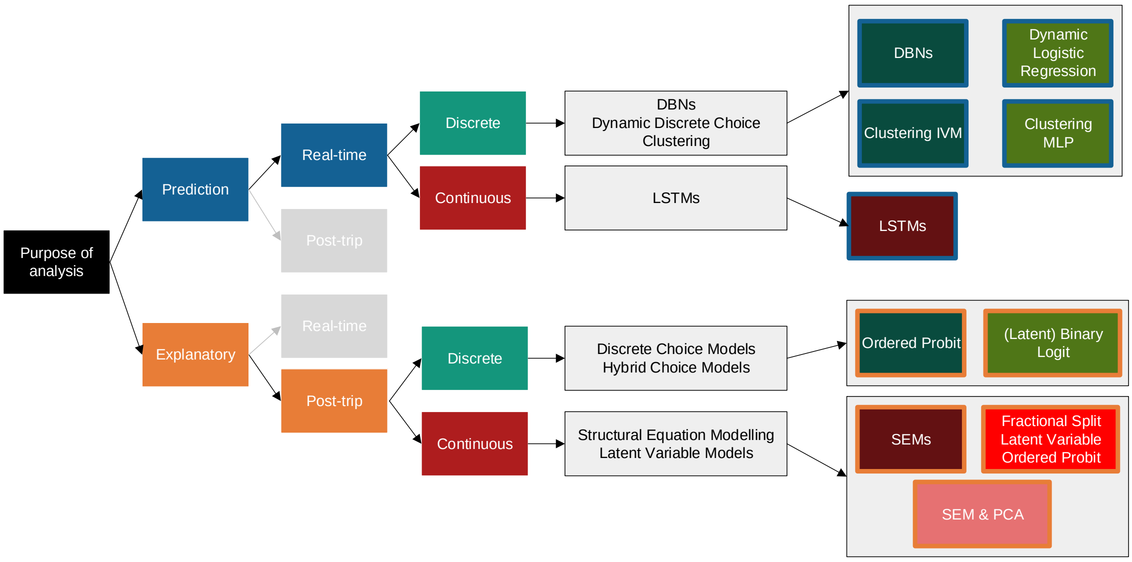


**Figure 5: Schematic overview of modeling approaches considered for the analysis of risk factors**

## 5. Conclusions

The aim of the current research is to present the practical conceptualization of the STZ in order to transition from a theoretical framework for operational design into a practical implementation and a fully functional methodology of the STZ concept. Four different methodological formulations were proposed to turn the available measurements into meaningful information on the level of driving safety. The most prominent approaches that can model driving behavior and recognize the three phases of the STZ were initially found to be Dynamic Bayesian Networks (DBNs; a probabilistic graphical time-series model) and Long Short-Term Memory networks (LSTMs; a deep neural network formulation), due to their efficiency and flexibility in real-time predictions. Furthermore, Discrete Choice Models (DCMs) and Structural Equation Models (SEMs) were also deemed suitable for the i-DREAMS concept, providing 'static' or post-trip predictions, in contrast with DBNs and LSTMs which work dynamically (i.e. in real-time). For each of the aforementioned methods or techniques, a brief description of their underpinning procedure was presented, followed by their application for the identification of the STZ levels.

Undoubtedly, it is not possible to specify the exact hypotheses behind the variables of interest without looking at the data. As a result, for all the proposed approaches, a labelled dataset is needed for training and this should be taken into consideration for the data collection. The effort, however, is that the details of all potential analytical models to be used (e.g. dependent and independent variables and the hypothesis about their relationship as well as the unit of analysis and model specifications) are documented so that the actual data analysis can start as soon as the data become available. Thus, the testing, calibration and enhancement of the mathematical models during the i-DREAMS simulation and on-road experiments can assure a sufficient and efficient data analysis, as well as timely initiation of the safety interventions. When preliminary results are available, the most crucial risk indicators of task complexity and coping capacity will be extracted, the proposed models will be tested and the suitable models will be selected for data analysis.

## Acknowledgments

The research was funded by the European Union's Horizon 2020 i-DREAMS project (Project Number: 814761) funded by European Commission under the MG-2-1-2018 Research and Innovation Action (RIA).

## References

1. Dula, C. S., & Geller, E. S. (2003). Risky, aggressive, or emotional driving: Addressing the need for consistent communication in research. Journal of safety research, 34(5), 559-566.
2. Fuller, R. (2005). Towards a general theory of driver behaviour. Accident analysis & prevention, 37(3), 461-472.
3. Fuller, R. (2011). Driver control theory: From task difficulty homeostasis to risk allostasis. In Handbook of traffic psychology (pp. 13-26). Academic Press.

4. 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.
5. Hughes, B. P., Newstead, S., Anund, A., Shu, C. C., & Falkmer, T. (2015). A review of models relevant to road safety. Accident Analysis & Prevention, 74, 250-270.
6. Sun, J., & Sun, J. (2015). A dynamic Bayesian network model for real-time crash prediction using traffic speed conditions data. Transportation Research Part C: Emerging Technologies, 54, 176-186.
7. Faure, V., Lobjois, R., Benguigui, N. (2016). The effects of driving environment complexity and dual tasking on drivers' mental workload and eye blink behavior. Transp. Res. Part F Traffic Psychol. Behav. 40, 78–90.
8. Agamennoni, G., Nieto, J.I., Nebot, E.M. (2011). A Bayesian approach for driving behavior inference. IEEE Intell. Veh. Symp. Proc. 595–600.
9. Gers, F. A., Schraudolph, N. N., & Schmidhuber, J. (2002). Learning precise timing with LSTM recurrent networks. Journal of machine learning research, 3(Aug), 115-143.
10. Girma, A., Yan, X., & Homaifar, A. (2019). Driver Identification Based on Vehicle Telematics Data using LSTM-Recurrent Neural Network. arXiv preprint arXiv:1911.08030.
11. Xu, C., Wang, W., Liu, P. (2013). A Genetic Programming Model for Real-Time Crash Prediction on Freeways. IEEE Trans. Intell. Transp. Syst. 14, 574–586.
12. Hensher, D. A., Rose, J. M., Rose, J. M., & Greene, W. H. (2005). Applied choice analysis: a primer. Cambridge university press.
13. McFadden, D. 1981. Econometric models of probabilistic choice. In: Manski, D., McFadden (Eds.), A Structural Analysis of Discrete Data with Econometric Applications. The MIT Press, Cambridge, MA.
14. Heck, R. H., & Thomas, S. L. (2020). An introduction to multilevel modeling techniques: MLM and SEM approaches. Routledge.
15. Tipping, M.E., and Bishop, C.M., 1999. Probabilistic principal component analysis. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 61(3), pp.611-622.