Using a Fuzzy-Bayesian Approach for Predictive Analysis of Delivery Delay Risk

Ouafae EL Bouhadi[1], Monir Azmani[2], Abdellah Azmani[3], Mouna Atik el ftouh[4]
Intelligent Automation Laboratory, FST of Tangier
Abdelmalek Essaadi University
Tetouan, Morocco

***Abstract*—Although one of the major roles of delivery logistics activities is to ensure a good quality of customer service, certain risks such as damage, delay, return of transported goods occur quite often. This makes risk control and prevention one of the requirements of transport supply chain quality. The article focuses on the analysis of the risk of delay, which is often considered fundamental for the quality of service and as a center of additional costs related to the violation of time windows. Such a risk can harm the image of a supplier, which can even lead to the loss of customers in case of recurrence. The aim of the following case study is the development of a fuzzy-bayesian approach that anticipates, by predictive analysis combining Bayesian networks (BNs) and Fuzzy logic, the possible delays affecting the smooth running of a delivery operation. The results of the implementation of the late delivery risk prediction model are validated by verifying three axioms. In addition, a sensitivity and scenario analysis is performed to validate the model and identify the parameters that have the most adverse impact on the occurrence of such a risk. These results can help carriers/transport providers to minimize potential late deliveries. In addition, the developed model can be used as a basis for different types of predictions in the field of freight transport as well as in other research areas.**

Keywords—Delivery logistics; risk management; predictive analysis; bayesian network; fuzzy logic

I. Introduction

On-time delivery is often considered as a performance indicator in the field of freight transportation [1]. However, any delay in goods delivery to their destination can harm the activities of the actors involved (shipper, carrier, warehouseman, and final customer) and affect their profitability. Therefore, measuring the risk of delay in advance enables reliable decisions to be made when planning deliveries. Consequently, the effect of such a risk can be reduced or even eliminated. This study focuses on the analysis of delivery delay to assist transportation companies in making decisions and optimizing their planning. Despite, the development of a delay risk prediction model is hampered by the unavailability of data. With this in mind, this paper contributes by implementing a fuzzy-bayesian approach to overcome the problem of missing or unavailable data.

The occurrence of delay in freight transport operations is influenced on the one hand by external factors such as road events (accidents, congestion, and weather conditions). On the other hand, by internal factors related to carriers' decisions in terms of resource selection and delivery planning [2]. This paper has studied all of these elements, and it has designed a predictive model integrating the different cause-effect relationships between several factors (internal and external) impacting on-time delivery. For this, it opted for BNs [3] thanks to its advantage of modeling, probabilistic reasoning of uncertain systems [4], and causal analysis.

The article is structured as follows: an overview of the literature related to delivery delay and the application of BNs in the transportation domain is given in Section II. The construction of the fuzzy-Bayesian model and its validation are presented in Section III. Section IV provides a discussion of the results obtained. Finally, the conclusion and further research are covered in Section V.

II. Related Work

In this section, we provide a brief overview of literature regarding our context of study, namely delayed deliveries, and for Bayesian network applications in transportation field.

A. Delayed Deliveries

Delayed deliveries mean goods arriving at their destination out of schedule. According to ATRI (American Transportation Research Institute), in 2016, trucking operations in the U.S. experienced about 1.2 billion hours of delays alone due to traffic congestion. This number of delays generates $74.5 billion in additional operational costs [5]. Because there are other causes of delay besides congestion, these additional costs continue to rise.

According to [6] various factors can disrupt delivery reliability, among them personnel issues, vehicle breakdowns, and poor planning. This means that resource selection by checking vehicle conditions and driver performance in addition to the adequacy of planning have a significant effect on speed of service. Furthermore, other external factors such as accessibility to delivery points, accidents, unforeseen events (public works), and methodological conditions [6]–[8] can cause further delays.

The occurrence of delay largely influences [9]–[12]:

- Additional costs: these costs are typically related to a delay penalty, and vehicle operating costs such as maintenance and vehicle rental in the case of private fleet use. As well as driver-related costs, that include salary and benefits. In addition, delay can lengthen the storage period, resulting in inventory holding costs.

- Warehouse productivity: arriving late at the warehouse may result in overlapping deliveries. This will influence the availability of unloading bays and labor as well as overloading the workforce to wait for late deliveries.
- Customer satisfaction: late deliveries can affect the operations of the consignees. This will have a significant impact on the customer relationship. As a result, customers find themselves in a situation of frustration. In case of recurrence, this situation exposes suppliers to the loss of customers.
- Loss of opportunity to consolidate shipments: the occurrence of delays contributes to the lack of opportunity to consolidate shipments due to the uncertainty of meeting delivery deadlines.
- Carrier profitability: this is generally affected by the additional costs incurred, the damage to the carrier's image due to the loss of customers and the extra time generated which reduces the opportunity to make more deliveries.

The concept of predicting the risk of late delivery is generally studied in the context of improving service quality. In the literature, the prediction of such risk is performed using several techniques from artificial intelligence.

Keung et al. [13] have opted for machine learning methods such as KNN, and ANN to predict shipment delays. The authors [14] used Random Forest and SVM for on-time delivery prediction. Berrones-Sanz [15] also proposed a model for on-time delivery prediction using logistic regression. In order to predict on-time delivery violation, [16] relied on ANN machine learning technique. As for [17], they proposed a neural network-based model to anticipate delivery time.

In addition, BNs are among the most popular prediction methods in various research areas [18]. This paper exploits the potential of BNs in risk prediction to anticipate the occurrence of delay in a delivery operation.

B. *Applications of BNs in the Transportation Field*

BNs are applied in several domains: diagnosis (medical and industrial), risk management, spam detection, fraud detection, data mining, and text mining, etc.[19]. In the transportation domain, BNs are used for different types of prevention. Gregoriades and Mouskos [20] used BNs to quantify the risk of accidents in order to locate black areas. The authors [21] developed a bayesian model for the identification of features affecting the safety of motor carriers. Zhu et al. [22] presented a bayesian approach for contextual (e.g., dynamic traffic information) or non-contextual (e.g., instantaneous driving speed) evaluation of driving behavior. As for [23], they modeled drivers' vehicle use behavior according to time of day. The use of BNs also extends to other transportation axes, namely traffic congestion prediction [24], [25] as well as freight demand prediction [26].

III. Modeling the Risk of Late Delivery using a Fuzzy-Bayesian Approach

A BN is "a graphical probabilistic model consisting of a set of nodes (variables of interest in the domain) and arcs (causal phenomena). In addition to a set of local probability distributions (network parameters)" [3], [27], [28]. A BN allows modeling the effect of a fact or an uncertain event on another via the representation of the causal relationship between those events [29]. In this case, an arc from X to Y can be interpreted as 'X causes Y'. Fig. 1 shows that the knowledge gained about the overlay of fragile products determines the knowledge about the damage of goods.

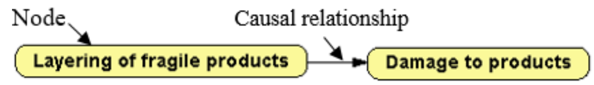

Fig. 1. Example of a Causality Representation.

The study is based on exhaustive bibliographical research, in addition to a survey of experts, in order to identify all the parameters (factors) which can hinder the delivery of a good to its destination in time. As well as the schematization of the dependency relations between these parameters, in order to build the structure of the Bayesian network. The approach followed introduces a methodology for modeling and evaluating the risk of late delivery using BNs by following these steps:

Definition of the BN structure: Determining the relationships between nodes allowed us to design a causal graph based on a three-level architecture.

- The first level represents the input or feeder nodes that have an indirect effect on the risk of delay; they fall into five categories detailed in Table I;
- The second level consists of the intermediate nodes that represent the various intermediate cause factors (direct causes, or cause-effects) that lead to the impact factors encapsulated in the third level (final impacts);
- The third level contains the final impacts that define the factors that directly and negatively influence the arrival of goods at their destination on time.

Generate conditional probabilities of intermediate effects and final impacts: The generation of conditional probability tables will be made using Sugeno's fuzzy inference implementation.

Model validation: This step consists of relying on Bayes theory and posterior probabilities relating to intermediate effects and final impacts, in order to study the effect of a state linked to an input parameter on the envisaged risk. In addition, sensitivity analysis and partial validation are performed to validate the model.

A. *Definition of the BN Structure*

The construction of the Bayesian network architecture can be done in two different ways:

- Objective methods: by using a database to apply the structure's learning methods.

- Subjective methods: by gathering knowledge from experts in the field, through written questionnaires, individual interviews or brainstorming sessions.

The first approach necessitates a large quantity of data. and it may establish dependencies or independencies between some variables that are inconsistent with the experts' opinions [30]. For this reason, researchers prefer the use of the second approach where experts are involved to verify the causal links between the network variables. [31]. In the literature, several researchers have relied on expert knowledge among them [31]–[35].With this in mind, the article used the subjective method.

Based on a survey of experts in the freight transportation field and a literature review of the various factors that drive the occurrence of delay risk, a set of network input parameters are identified and presented in detail in Table I. In addition to the intermediate effects and final impacts presented in Table II. The survey consisted of two questionnaires, the first was conducted to verify and validate the identified variables and the second to establish the causal relationships between the variables.

After identifying the variables (Tables I and II) that will constitute the nodes of the graph and studying the causal relationships between them, the structure of the Bayesian network is developed and illustrated in Fig. 2.

B. *Generation of Conditional Probabilities of Intermediate Effects and Final Impacts*

After building the Bayesian network structure, the next step is to compute the conditional probability tables (CPT) for each variable. These CPTs can be computed based on the knowledge of the experts or using learning algorithms from a database. Although there is no more database in the literature adapted to the parameters identified to build the graph structure, the article is oriented towards the use of subjective methods.

Since the number of conditional probabilities in the developed network is 1032, it is difficult to rely on expert knowledge to evaluate such a large number of probabilities[36]. In fact, in the literature, many researchers have developed models to lower the number of CPTs of a BN. As an example, causal interaction models that have attracted the interest of several researchers such as [37]–[41]. One of the most widely used models is the Noisy-OR model introduced by [39] which allows for the specification of non-deterministic interactions between the parents associated with an effect [42].

In addition, other methods such as fuzzy logic [43] allowing the reduction of numbers of questions asked to experts and the generation of probability tables [2], [44], [45] .With this in mind, the paper relied on this method to first express the experts judgments by fuzzy rules, and then generate the conditional probability tables by a fuzzy inference mechanism. These fuzzy rules are of the type 'If the driver's performance is bad then the occurrence of the accident is high. Here the value of "high" for the accident occurrence is qualitatively represented by a linguistic variable expressed in natural language. For the different rules, the accident occurrence node can be translated into one of the following values: low, medium, high. In addition, for the different nodes of the graph, their influences are revealed by three linguistic values (states) represented in Table III. Since the quantification of these linguistic variables oscillates in interval 0.1, the article opted for the expression of these values in fuzzy form to ascertain the degree of each node's membership in all its fuzzy subsets. As an example, the accident occurrence node is high with 90%, medium with 8% and low with 2%.

The implementation of the approach adopted for generating conditional probability tables is done in three steps:

- Definition of fuzzy variables, their associated linguistic values(variable whose values are qualitative and represent natural language expressions [62] and their membership functions;
- Determination of fuzzy rule bases: a base of "if-then" rules, is used by the "fuzzy inference system" in order to translate the input variables into output [62];
- Development of inference mechanism that forms conclusions based on the fuzzy rules and the input data [63].

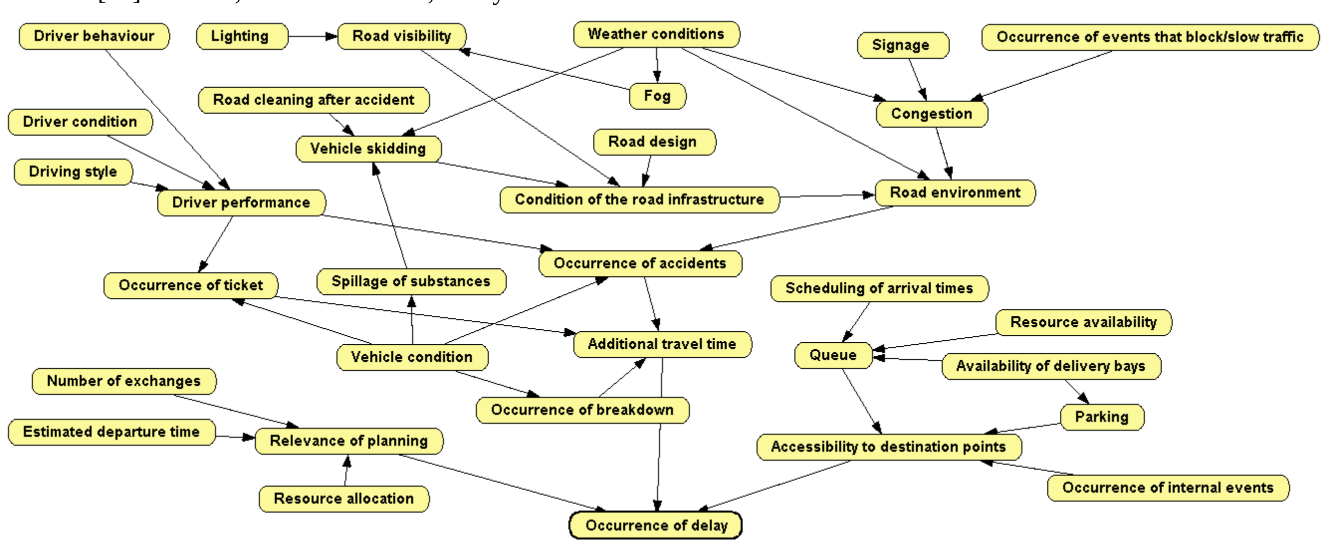

Fig. 2. The Structure of the Bayesian Network Modeling the Risk of Delay of a Delivery.

TABLE I. DESCRIPTION OF THE INPUT PARAMETERS OF THE CAUSAL GRAPH

Variable class	**Variable name**	**Description**
Parameters related to the road	Road design	It refers to the geometric nature of the road, such as the number of lanes and direction of traffic. [46].
	Lighting	Lighting conditions (artificial light, daylight, darkness, twilight) have a considerable effect on the occurrence of road accidents, due to the adaptation of speed with visibility conditions [47].
	Road cleaning after accident	Improper cleaning of accident sites can result in the vehicle skidding due to the presence of body fluids, fuel and other debris.
Traffic parameters	Signage	Traffic signal control considers the control of traffic signals and the existence of stop sign and yield sign [48] .
	Weather conditions	Weather conditions can disrupt driver capabilities, pavement friction, road infrastructure conditions, and vehicle stability and maneuverability, due to reduced visibility, extreme temperatures, precipitation, high winds, and lightning [49]. As a result, they affect traffic demand (carriers postpone or cancel planned delivery operations), traffic safety (accident rates), and traffic flow relationships (changes in the fundamental traffic flow variables, volume, speed, and density influence the capacity of a road system) [50].
	Occurrence of events that block/slow traffic	Traffic can be disrupted by the occurrence of a variety of events, including: Unexpected events such as road accidents, vehicle breakdown in the middle of the road, land subsidence and public works [51]; Irregular social events such as political demonstrations, diplomatic visits [51], sporting events; Regular events such as festive events (religious, national or international holidays) and vacation departure; In addition, the delivery period also influences traffic. Delivery during peak hours can slow down traffic.
Parameters related to personnel (drivers, planners, ...) or planning systems	Driver behavior	Driver behavior is considered one of the major sources of traffic accidents [52]. These behaviors include failure to observe speed limits, safety distance, and poor vehicle handling that manifests itself in hard braking, and hard acceleration, especially when visibility is reduced, and when climbing hills [53].
	Driving style	A driver's habitual driving style that includes calmness behind the wheel, level of attention contribute to traffic accidents [54], [55].
	Driver condition	The poor condition of drivers is often due to fatigue, alcohol consumption, negative emotions (worries and fear of arriving late), drowsiness, headaches, respiratory illnesses, and fever [56].
	Number of exchanges	The increase in the number of exchanges resulting from indirect deliveries leads to more loading/unloading operations [57], and therefore an increase in the risk of delay due to uncontrolled/estimated time at the exchange points.
	Departure time	A travel time estimate is said to be accurate if it could help improve the quality of service by delivering the goods on time [58]. For this, the departure time must be sufficient to cover the planned route within the required time window.
	Resource allocation	Personnel, vehicle condition (breakdown), and delivery points are among the factors that cause delay to occur [6]. Therefore, optimal allocation of resources (vehicles, warehouses, personnel, etc.) to routes helps to reduce the risk of delay.
Vehicle-related parameters	Vehicle condition	Vehicle condition is one of the causes of accidents [59]. Poor physical condition of the vehicle can cause vehicle breakdown and hence the occurrence of delays. In addition, negligence in checking the administrative condition of vehicle can be the cause of vehicle ticketing.
Parameters related to delivery areas (warehouses)	Availability of delivery bays	The lack of adequate delivery bays for different vehicle sizes leads to long queue [60]. In this case, drivers try to park their vehicles far away or illegally near the delivery site. This leads to long queue.
	Resource availability	The lack of human resources in the delivery areas is one of the factors that lead to long queue[60]. Similarly, the lack of material resources (handling equipment, forklifts, etc.) can aggravate this problem.
	Internal events	Internal events such as strikes over working conditions can slow down or block access to delivery areas.
	Scheduling of arrival times	Synchronizing inbound and outbound delivery operations within a warehouse requires reliable arrival times [61]. Lack of scheduling of arrival times (delivery deadlines) with carriers can cause overlapping operations of multiple deliveries, subsequently resulting in long queues.

TABLE II. PRESENTATION OF INTERMEDIATE EFFECTS AND IMPACTS ON DELAY

Intermediate effects		**Final Impacts**
• Road visibility • Vehicle skidding • Condition of the road infrastructure • Spillage of substances • Congestion • Occurrence of ticket	• Road environment condition • Driver performance • Queue • Parking • Occurrence of accidents • Occurrence of breakdown	• Additional travel time • Accessibility to destination points • Relevance of planning • Occurrence of delay

TABLE III. STATES OF THE BAYESIAN NETWORK NODES

Nodes	Linguistic values
Road design	Good, medium, bad
Lighting	Good, medium, bad
Fog	Low, medium, heavy
Road visibility	Good, medium, bad
Condition of the road infrastructure	Good, medium, bad
Road cleaning after accident	Appropriate, medium, inappropriate
Spillage of substances	Low, medium, strong
Vehicle skidding	Low, medium, significant
Weather conditions	Normal, medium, extreme
Occurrence of events that block/slow traffic	Low, medium, high
Signage	Good, medium, bad
Congestion	Low, medium, high
Road environment	Good, medium, bad
Vehicle condition	Good, medium, bad
Driver performance	Good, medium, bad
Driver behavior	Good, medium, bad
Driving style	Good, medium, bad
Driver condition	Good, medium, bad
Occurrence of accidents	Low, medium, high
Occurrence of breakdown	Low, medium, high
Occurrence of ticket	Low, medium, high
Additional travel time	Low, medium, high
Number of exchanges	Low, medium, high
Resource allocation	Good, medium, bad
Relevance of planning	Good, medium, bad
Estimated departure time	Good, medium, bad
Scheduling of arrival times	Good, medium, bad
Occurrence of internal events	Low, medium, high
Resource availability	High, medium, low
Availability of delivery bays	High, medium, low
Queue	Fast, medium, slow
Parking	Very close, close, far
Accessibility to destination points	High, medium, low
Occurrence of delay	Low, medium, high

To better assimilate our proposed approach, we explain CPTs generation for 'Congestion' node. In this case, the inference mechanism aims at determining the probabilities of the occurrence of congestion with respect to the states of the nodes: Signage, weather conditions, occurrence of events blocking/disrupting the traffic.

The membership function used for the different nodes of the graph is Gaussian, as it provides less error compared to other triangular and trapezoidal functions [64]. An example of the description of the membership function for the weather variable is provided in Fig. 3.

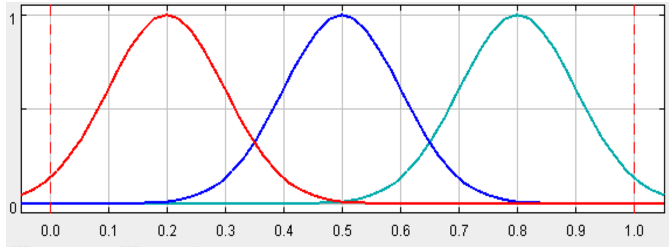

Fig. 3. Membership Functions for the Weather Variable.

After defining the membership functions for the congestion nodes and its antecedents, subsequently, a fuzzy rule base is determined to evaluate the variation of the congestion node with respect to the states of its parent nodes or its causes. This fuzzy rule base is detailed in Table IV.

TABLE IV. FUZZY RULES OF THE 'CONGESTION' NODE WITH ITS PARENT NODES

Rule number	Signage	Weather conditions	Occurrence of events	Congestion
1	Bad	extremes	High	high
2	Bad	extremes	Medium	high
3	Bad	extremes	Low	medium
4	Bad	medium	High	high
5	Bad	medium	Medium	medium
6	Bad	medium	Low	medium
7	Bad	normal	High	medium
8	Bad	normal	Medium	medium
9	Bad	normal	Low	low
10	Medium	extremes	High	high
11	Medium	extremes	Medium	medium
12	Medium	extremes	Low	medium
13	Medium	medium	High	medium
14	Medium	medium	Medium	medium
15	Medium	medium	Low	medium
16	Medium	normal	High	medium
17	Medium	normal	Medium	medium
18	Medium	normal	Low	low
19	Good	extremes	High	high
20	Good	extremes	Medium	medium
21	Good	extremes	Low	medium
22	Good	medium	High	medium
23	Good	medium	Medium	medium
24	Good	medium	Low	low
25	Good	normal	Low	medium
26	Good	normal	Medium	low
27	Good	normal	Low	low

The inference mechanism is triggered by initializing the input variables with precise values representing the peak of the Gaussian distribution and activating a set of fuzzy rules. This mechanism uses the Sugeno inference system. The aim is to identify degrees of membership to each fuzzy subset. The calibration of these degrees provides the CPTs of the BN.

Consider the example of inferring the congestion node knowing that the signage is average, the methodological conditions are normal, and the occurrence of traffic blocking/interfering events is low. The initial values of the input variables represented in rule 18 of Table VI will feed the fuzzy inference system. The latter is implemented with the Fispro software.

The results of the fuzzy inference are illustrated in Fig. 4, thereafter; the different results are aggregated, in order to combine them into a single value for each state. This value is the result of the union of the different conclusions of the rules activated with the max method.

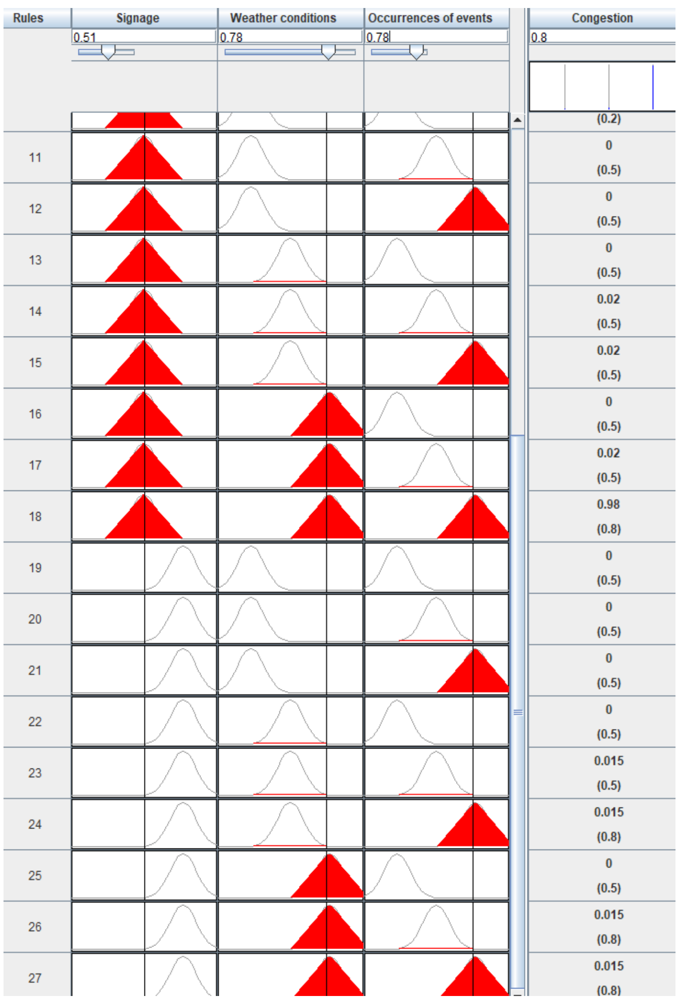

Fig. 4. Fuzzy Inference of the Variable 'Congestion.

The conclusions of the activated rules of rule 18 are summarized in Table V. Thus, the low value of congestion is 0.98, and the average value is 0.020. Now, since each possibility of a fuzzy subset must be greater than zero, a value of 0.001 is given to the null probability, in this case the high value is 0.001.

Congestion (low) = max (0.98, 0.015, 0.008) = 0.98.

Congestion (medium) = max (0.015, 0.008, 0.020) = 0.020.

Congestion (high) = 0.001.

The conditional probabilities for the different states of the variable 'congestion' of the rule 18 is calculated in the following way:

P (Congestion = low | Occurrence events = low, Weather = normal and Signage = medium) = 0.98/(0.98+0.020+0.001)=0.979.

P (Congestion = medium | Occurrence of events = low, Weather = normal and Signage = medium) = 0.020/(0.98+0.020+0.001)=0.019.

P (Congestion = high | Occurrence of events = low, Weather = normal and Signage = medium) = 0.001/(0.98+0.020+0.001)=0.001.

TABLE V. DEGREE OF MEMBERSHIP FOR EACH FUZZY SUBSET OF THE VARIABLE 'CONGESTION

Rule activated	**Language value of the output variable**	**Degree of membership**
R27	Low	0.015
R26	Low	0.015
R24	Low	0.015
R23	Medium	0.015
R18	Low	0.98
R17	Medium	0.020
R15	Medium	0.020
R14	medium	0.020
R9	Low	0.008
R8	medium	0.008
R6	medium	0.008

C. Anticipation of Scenarios and Interpretation of Results

After constructing the Bayesian network using Open Markov tool, it is used to deduce the probabilities of certain events by setting evidences (states) for certain nodes and study their effects through the propagation of their probabilities on the child nodes. At this level, the input parameters' impact on the occurrence of delay is studied through four scenarios listed as follows:

- Scenario 1: favorable road environment and delivery areas input parameters, unfavorable transport company input parameters.
- Scenario 2: unfavorable road environment and delivery areas input parameters, favorable transport company input parameters.

- Scenario 3: favorable road environment and delivery areas input parameters, favorable transport company input parameters.
- Scenario 4: unfavorable road environment and delivery areas input parameters, unfavorable transport company input parameters.

The input parameters related to the transport company correspond to those related to the vehicles, the staff (drivers, planners ...) and the planning systems.

The input parameters related to the road environment and delivery areas correspond to those related to the road, traffic and warehouses.

The four scenarios corresponding to different configurations of the input node states are detailed in Table VII.

After feeding the BN with the states of each scenario, the inference mechanism provides probability propagation over intermediate effects and final impacts to quantify delay occurrence. Table VIII, shows the probability distribution for some nodes in the network.

TABLE VI. CONDITIONAL PROBABILITY TABLE OF THE 'CONGESTION' NODE

					Conditional probability		
Rule number	**Signage**	**Weather conditions**	**Occurrences of events**	**Congestion**	**Low**	**medium**	**high**
1	bad 0.22	extremes 0.18	high 0.21	high	0.001	0.015	0.984
2	bad 0.22	extremes 0.18	medium 0.52	high	0.001	0.020	0.979
3	bad 0.22	extremes 0.18	low 0.78	medium	0.001	0.979	0.020
4	bad 0.22	medium 0.5	high 0.21	high	0.001	0.020	0.979
5	bad 0.22	medium 0.5	medium 0.52	medium	0.001	0.988	0.011
…	…	…	…	…	…	…	…
18	medium 0.51	normal 0.78	low 0.78	low	0.979	0.020	0.001
19	good 0.81	extremes 0.18	high 0.21	high	0.001	0.015	0.984
20	good 0.81	extremes 0.18	medium 0.52	medium	0.006	0.988	0.006
…	…	…	…	…	…	…	…
25	good 0.81	normal 0.78	high 0.21	medium	0.015	0.984	0.001
26	good 0.81	normal 0.78	medium 0.52	low	0.979	0.020	0.001
27	good 0.81	normal 0.78	low 0.78	low	0.979	0.020	0.001

TABLE VII. DESCRIPTION OF THE SCENARIOS ACCORDING TO THE VALUES OF THE INPUT PARAMETERS

Input parameter	**Value**			
	Scenario 1	***Scenario 2***	***Scenario 3***	***Scenario 4***
Road environment and delivery areas parameters				
Road design	good	bad	good	bad
Lighting	good	bad	good	bad
Post-accident road cleaning	appropriate	inappropriate	appropriate	inappropriate
Signage	good	bad	good	bad
Weather conditions	normal	extremes	normal	extreme
Occurrence of events that block/impede traffic	low	high	low	high
Availability of delivery bays	high	low	high	low
Resources Availability	high	low	high	low
Occurrence of internal events	low	high	low	high
Scheduling of arrival times	good	bad	good	bad
Transport company parameters				
Driving style	bad	good	good	bad
Driver condition	bad	good	good	bad
Driver behavior	bad	good	good	bad
Vehicle condition	bad	good	good	bad
Resource allocation	bad	good	good	bad
Estimated departure time	bad	good	good	bad
Number of exchanges	high	Low	low	high

TABLE VIII. PROBABILITY DISTRIBUTION FOR BAYESIAN NETWORK NODES

Node	State	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Congestion	Low	0.9790	0.0010	0.9790	0.0010
	Medium	0.0020	0.0150	0.0020	0.0150
	High	0.0010	0.9840	0.0010	0.9840
Occurrence of a breakdown	high	0.9790	0.0010	0.0010	0.9790
	Medium	0.0200	0.0020	0.0020	0.0174
	Low	0.0027	0.9790	0.9790	0.0010
Occurrence of an accident	Low	0.0022	0.0012	0.9574	0.0010
	Medium	0.0418	0.9574	0.0416	0.0173
	high	0.9559	0.0214	0.0010	0.9817
	Low	0.0209	0.9166	0.9166	0.0052
Occurrence of ticket	Medium	0.0052	0.0592	0,0592	0.0209
	high	0.9738	0.0242	0.0242	0.9738
Accessibility to destination points	High	0.9635	0.0010	0.9635	0.0010
	Medium	0.0355	0.0155	0.0355	0.0155
	Low	0.0010	0.9835	0.0010	0.9835
Delay	Low	0.0017	0.0050	0.9718	0.0010
	Medium	0.0861	0.9624	0.0272	0.0160
	High	0.9121	0.0327	0.0010	0.9830

In the case of the first scenario, the occurrence of congestion is low with a probability of 97.9%. Concerning the occurrence of a breakdown and ticket, and accident are important with respectively 97.9% and 97.3% and 95.5%. As for the accessibility to the destination, point tends to be high with 96.3%. Therefore, the probability of the occurrence of delay is important with a value of 91.1%.

For the second scenario, the probability of congestion occurrence is high with a probability of 98.4%. Regarding the occurrence of accident is more likely to be 'medium' with a probability of 95.7%, as for the occurrence of breakdown and ticket are low with respectively 97.9% and 91.6%, In addition, the accessibility to the destination point tends to be low with 98.3% which leads to an average risk of occurrence of delay of 96.2%

For the third scenario, the inference model predicts a low probability of congestion occurrence with a rate of 97.9%. Also, the risks related to the occurrence of an accident, breakdown and a ticket are low with respectively 95.7%, 97.9% and 91.6%. In addition, the accessibility to the destination point tends to be high with 96.3%. Therefore, the probability of delay occurrence is very low with 97.1%.

For the fourth scenario, congestion tends to be high with a rate of 98.4%. Also, the risks of the occurrence of an accident, a breakdown or a ticket are high with respectively 98.1%, 97.9% and 97.3%. As for the accessibility to the destination, point is low with a percentage of 98.3%. Hence the highest probability of the occurrence of delay with 98.3%.

Based on the results of the inference of the first two scenarios, we can conclude that the parameters related to the transport company have more impact on the occurrence of delivery delay than those related to road and delivery environment.

D. Sensitivity Analysis

The sensitivity analysis identifies the factors ranked according to those that have more impact on the probability of a node [65]. The Fig. 5 shows the tornado diagram [66] of the "delay occurrence" sensitivity analysis. According to the figure, variables such as vehicle condition, resources availability, estimated departure time, number of exchanges, and resource allocation contribute more to the occurrence of late deliveries. Poor vehicle condition is the main parameter influencing the risk of delay. This can be justified by its resulting events such as the occurrence of breakdowns and accidents, as these have a significant impact on the occurrence of such risk.

E. Partial Validation

In order to validate the proposed model, the paper uses partial validation allowing the verification of three axioms presented by [67] and opted by various researchers such as [68], [69]: (1) The occurrence of change (increase/decrease) in the probability of the parent node thus changes the probability of the child node; (2) The values of the child nodes must be consistently affected by the changes made to the probability distributions relative to the parent node; In the case of a node with more than one parent (e.g., x and y), the overall effect of x and y must be greater than the individual effects of parent x or parent y.

For example, when the probability of "road design = good" increases from 70% to 75%, the probability of "road infrastructure condition = good" and "road environment = good" increases from 77% to 80% and 84% to 86% respectively. In light of this variation, when the probability of "Occurrence of events that block/slow traffic=low" increases from 74% to 78%, the probability of "congestion=low" and "road environment=good" increases from 78% to 82% and 87% to 90% respectively, which is consistent with the axiom. Similarly for the other nodes of the graph.

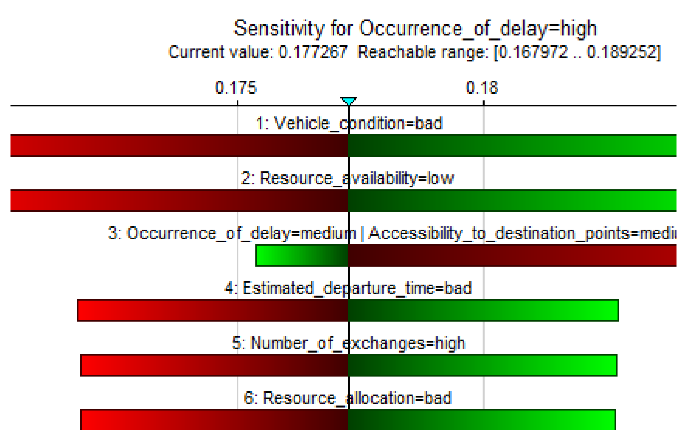

Fig. 5. Sensitivity Analysis for Occurnce of Delay.

IV. Discussion

In this paper, an analysis of potential delivery delays is developed using a Bayesian fuzzy model. BNs have been used in various existing studies to analyze the risk of delay in different areas such as maritime transport[45], [70], rail transport [71] and air transport [72]. Compared to the existing works, the developed model focuses on the occurrence of delay in the road freight transport domain. This model consists of many internal and external factors that cause delay to occur. Sensitivity analysis and interpretation of the results of four scenarios show that the internal factors related to the transport companies have a stronger effect on the occurrence of late deliveries than the external factors. This means that the transport company decisions in terms of resource selection, in addition to planning relevance, have a considerable effect on delivery reliability than road and delivery events such as congestion, weather conditions and availability of delivery bays. Therefore, optimized resource (physical and material) allocation to the right routes and a smart routing planning design plays an important role in ensuring on-time deliveries. In order to validate the results of this model, an example test of three axioms is performed.

V. Conclusion

The respect of delivery deadlines is crucial to ensure the quality of logistics services. Unfortunately, it often happens that this deadline is not respected, leading to a series of more or less unfortunate consequences. This article focuses on predicting the risks of delays in deliveries with the aim of anticipating them in order to either avoid them or be better prepared to deal with them and reduce the impact of poor quality of service. To do this, the article proposes a fuzzy Bayesian model combining the Bayesian approach and fuzzy logic in order to monitor the occurrence of delivery delays. This model is based on a set of factors causing the delay of deliveries and their causal relationships represented by a causal graph. Fuzzy logic intervenes by the generation of fuzzy rules based on conditional probability tables. Such a model is particularly effective, especially for the type of problem dealt with through this article. It is positioned as an excellent alternative to deep learning for making predictions in the absence of massive data.

One of the limitations of the proposed model is the identification of all the factors that cause delivery delays, the lack of integration of all these factors can affect the effectiveness of this model in correctly predicting possible delays.

The generalization of the model to all risks that can degrade the quality of delivery services such as damage, theft and loss, is a potential perspective to the work presented here. Similarly, a comparative study between a Bayesian-fuzzy model and a model derived from deep-learning around the prediction of delivery risks and the management of the crises that follow are very promising avenues of research.

Acknowledgment

This research is supported by the Ministry of Higher Education, Scientific Research and Innovation, the Digital Development Agency (DDA) and the National Center for Scientific and Technical Research (CNRST) of Morocco (Smart DLSP Project - AL KHAWARIZMI IA-PROGRAM).

References

[1] R. D. Alcoba et K. W. Ohlund, « Predicting On-time Delivery in the Trucking Industry », p. 4, 2017.

[2] M. ATIK EL FTOUH, « Modélisation d'un smart écosystème numérique pour une gestion optimisée et collaborative du transport urbain des marchandises, agrégeant pour l'aide à la décision plusieurs méthodes et techniques de l'intelligence artificielle », 2020.

[3] J. Pearl, Bayesian networks: A model of self-activated memory for evidential reasoning. UCLA, Computer Science Dept, 1985.

[4] J. Chen et al., « shinyBN: an online application for interactive Bayesian network inference and visualization », BMC Bioinformatics, vol. 20, no 1, p. 711, déc. 2019, doi: 10.1186/s12859-019-3309-0.

[5] A. Hooper, « Cost of Congestion to the Trucking Industry: 2018 Update », p. 41, 2018.

[6] A. Mckinnon, J. Edwards, M. Piecyk, et A. Palmer, « Traffic congestion, reliability and logistical performance: A multi-sectoral assessment », International Journal of Logistics-research and Applications - INT J LOGIST-RES APPL, vol. 12, oct. 2009, doi: 10.1080/1367556090 3181519.

[7] O. Korbaa, A. Boufaied, et C. Lajimi, « Assessing and modelling transport delays risk in supply chains », IJAOM, vol. 9, no 4, p. 225, 2017, doi: 10.1504/IJAOM.2017.10010983.

[8] T. Bektas, « Freight Transport and Distribution », p. 286, 2017.

[9] H. Chevroton, Y. Kergosien, et J.-C. Billaut, « Problème de tournée avec pénalités de départ et de retard », p. 9, 2019.

[10] A. Fowkes, P. E. Firmin, G. Tweddle, et T. Whiteing, « How highly does the freight transport industry value journey time reliability - and for what reasons? », International Journal of Logistics, vol. 7, mars 2004, doi: 10.1080/13675560310001619259.

[11] Q. Gong, Q. Miao, B. X. Wang, et T. M. Adams, « Assessing Public Benefits and Costs of Freight Transportation Projects: Measuring Shippers' Value of Delay on the Freight System », Art. no CFIRE 04-14, juill. 2012.

[12] V. Lukinskiy, N. Pletneva, V. Gorshkov, et P. Druzhinin, « Application of the Logistics "Just in Time" Concept to Improve the Road Safety », Transportation Research Procedia, vol. 20, p. 418-424, 2017, doi: 10.1016/j.trpro.2017.01.068.

[13] K. L. Keung, C. Lee, et Y. Yiu, A Machine Learning Predictive Model for Shipment Delay and Demand Forecasting for Warehouses and Sales Data. 2021, p. 1014. doi: 10.1109/IEEM50564.2021.9672946.

[14] A. Durjoy, A. Bristy, et M. Hasan, E-commerce shipping: Prediction of on time delivery of products using Data Mining. 2021.

[15] L. Berrones-Sanz, « PREDICTING ON-TIME DELIVERIES IN TRUCKING: A MODEL BASED ON THE WORKING CONDITIONS OF DRIVERS », AT, vol. 7, no 2, p. 47-53, juin 2021, doi: 10.22306/atec.v7i2.105.

[16] R. Shabbar et A. Kasasbeh, Proactive Event Management using ANN with PSO Prediction in Transport Processes. 2017.

[17] A. M. Ghaithan, I. Alarfaj, A. Mohammed, et O. Qasim, « A neural network-based model for estimating the delivery time of oxygen gas cylinders during COVID-19 pandemic », Neural Comput & Applic, vol. 34, no 13, p. 11213-11231, juill. 2022.

[18] S. Borujeni, N. Nguyen, S. Nannapaneni, E. Behrman, et J. Steck, Experimental evaluation of quantum Bayesian networks on IBM QX hardware. 2020.

[19] O. PARENT et J. EUSTACHE, « Les Réseaux Bayésiens A la recherche de la vérité ». Université Claude Bernard Lyon, 2007.

[20] A. Gregoriades et K. C. Mouskos, « Black spots identification through a Bayesian Networks quantification of accident risk index », Transportation Research Part C: Emerging Technologies, vol. 28, p. 28-43, mars 2013, doi: 10.1016/j.trc.2012.12.008.

[21] S. Hwang, L. N. Boyle, et A. G. Banerjee, « Identifying characteristics that impact motor carrier safety using Bayesian networks », Accident Analysis & Prevention, vol. 128, p. 40-45, juill. 2019, doi: 10.1016/j.aap.2019.03.004.

[22] X. Zhu, Y. Yuan, X. Hu, Y.-C. Chiu, et Y.-L. Ma, « A Bayesian Network model for contextual versus non-contextual driving behavior assessment », Transportation Research Part C Emerging Technologies, vol. 81, p. 172-187, août 2017, doi: 10.1016/j.trc.2017.05.015.

[23] D. Li, T. Miwa, et T. Morikawa, « Modeling time-of-day car use behavior: A Bayesian network approach », Transportation Research Part D: Transport and Environment, vol. 47, p. 54-66, août 2016, doi: 10.1016/j.trd.2016.04.011.

[24] T. Afrin et N. Yodo, « A probabilistic estimation of traffic congestion using Bayesian network », Measurement, vol. 174, p. 109051, avr. 2021, doi: 10.1016/j.measurement.2021.109051.

[25] Y. CAO, C. WANG, Y. YANG, J. XU, et Y. GAO, « Multicause Automatic Real-time Identification of Urban Road Traffic Congestion Based on Bayesian Network », 30 décembre 2020.

[26] M. Petri, G. Fusco, et A. Pratelli, « A New Data-Driven Approach to Forecast Freight Transport Demand », juin 2014, vol. 8582, p. 401-416. doi: 10.1007/978-3-319-09147-1_29.

[27] C. J. Butz, S. Hua, J. Chen, et H. Yao, « A simple graphical approach for understanding probabilistic inference in Bayesian networks », Information Sciences, vol. 179, no 6, p. 699-716, mars 2009, doi: 10.1016/j.ins.2008.10.036.

[28] S. Verron, « Diagnostic et surveillance des processus complexes par réseaux bayésiens », 2007.

[29] T. D. Nielsen et F. V. JENSEN, Bayesian Networks and Decision Graphs. Springer Science & Business Media, 2009.

[30] F. Ghasemi, O. Kalatpour, A. Moghimbeigi, et I. Mohammadfam, « Selecting Strategies to Reduce High-Risk Unsafe Work Behaviors Using the Safety Behavior Sampling Technique and Bayesian Network Analysis », J Res Health Sci, vol. 17, no 1, p. 372, mars 2017.

[31] I. Mohammadfam, F. Ghasemi, O. Kalatpour, et A. Moghimbeigi, « Constructing a Bayesian network model for improving safety behavior of employees at workplaces », Applied Ergonomics, vol. 58, p. 35-47, janv. 2017, doi: 10.1016/j.apergo.2016.05.006.

[32] S. A. Abdulkareem, Y. T. Mustafa, E.-W. Augustijn, et T. Filatova, « Bayesian networks for spatial learning: a workflow on using limited survey data for intelligent learning in spatial agent-based models », Geoinformatica, vol. 23, no 2, p. 243-268, avr. 2019, doi: 10.1007/s10707-019-00347-0.

[33] R. Kaya et B. Yet, « Building Bayesian networks based on DEMATEL for multiple criteria decision problems: A supplier selection case study », Expert Systems with Applications, vol. 134, p. 234-248, nov. 2019, doi: 10.1016/j.eswa.2019.05.053.

[34] S. Zhu, X. Cai, J. Lu, et Y. Peng, « Analysis of factors affecting serious multi-fatality crashes in China based on Bayesian network structure », Advances in Mechanical Engineering, vol. 9, p. 168781401770414, juin 2017, doi: 10.1177/1687814017704145.

[35] P. Huang et al., « A Bayesian network model to predict the effects of interruptions on train operations », Transportation Research Part C: Emerging Technologies, vol. 114, p. 338-358, mai 2020, doi: 10.1016/j.trc.2020.02.021.

[36] S. Renooij, « Probability elicitation for belief networks: issues to consider », The Knowledge Engineering Review, 2001, doi: 10.1017/S0269888901000145.

[37] D. Heckerman et J. Breese, « Causal independence for probability assessment and inference using Bayesian networks », Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, vol. 26, p. 826-831, déc. 1996, doi: 10.1109/3468.541341.

[38] P. J. Lucas, « Bayesian network modelling through qualitative patterns », Artificial Intelligence, vol. 163, p. 233-263, déc. 2003, doi: 10.1016/j.artint.2004.10.011.

[39] J. Pearl, Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann, 1988.

[40] S. P. D. Woudenberg, L. C. van der Gaag, et C. M. A. Rademaker, « An intercausal cancellation model for Bayesian-network engineering », International Journal of Approximate Reasoning, vol. 63, p. 32-47, août 2015, doi: 10.1016/j.ijar.2015.05.011.

[41] Y. Xiang et N. Jia, « Modeling Causal Reinforcement and Undermining for Efficient CPT Elicitation », IEEE Transactions on Knowledge and Data Engineering, vol. 19, no 12, p. 1708-1718, déc. 2007, doi: 10.1109/TKDE.2007.190659.

[42] S. Srinivas, « A Generalization of the Noisy-Or Model », in Uncertainty in Artificial Intelligence, Elsevier, 1993, p. 208-215. doi: 10.1016/B978-1-4832-1451-1.50030-5.

[43] L. A. Zadeh, « Fuzzy sets », Information and Control, vol. 8, no 3, p. 338-353, juin 1965, doi: 10.1016/S0019-9958(65)90241-X.

[44] C. LAMAAKCHAOUI, « Modélisation d'un Ecosystème Numérique Favorisant la Transformation Digitale et l'Aide à la Décision Marketing et agrégeant plusieurs Concepts et Méthodes de l'Intelligence Artificielle, un Système Multi-Agent et son Ontologie », ABDELMALEK ESSAADI, 2018.

[45] N. H. M. Salleh, R. Riahi, Z. Yang, et J. Wang, « Predicting a Containership's Arrival Punctuality in Liner Operations by Using a Fuzzy Rule-Based Bayesian Network (FRBBN) », The Asian Journal of Shipping and Logistics, vol. 33, p. 95-104, juill. 2017.

[46] P. Álvarez, M. A. Fernández, A. Gordaliza, A. Mansilla, et A. Molinero, « Geometric road design factors affecting the risk of urban run-off crashes. A case-control study », PLOS ONE, vol. 15, no 6, p. e0234564, juin 2020, doi: 10.1371/journal.pone.0234564.

[47] A. K. Jägerbrand et J. Sjöbergh, « Effects of weather conditions, light conditions, and road lighting on vehicle speed », SpringerPlus, vol. 5, no 1, p. 505, avr. 2016, doi: 10.1186/s40064-016-2124-6.

[48] X. Zou et W. Yue, « A Bayesian Network Approach to Causation Analysis of Road Accidents Using Netica », Journal of Advanced Transportation, vol. 2017, déc. 2017, doi: 10.1155/2017/2525481.

[49] P. Pisano et L. C. Goodwin, « Surface Transportation Weather Applications », p. 11, 2002.

[50] T. H. Maze, M. Agarwal, et G. Burchett, « Whether Weather Matters to Traffic Demand, Traffic Safety, and Traffic Operations and Flow », Transportation Research Record, vol. 1948, no 1, p. 170-176, janv. 2006, doi: 10.1177/0361198106194800119.

[51] M. Rao et K. R. Rao, « Measuring Urban Traffic Congestion – A Review », International Journal for Traffic and Transport Engineering, vol. 2, p. 286-305, déc. 2012, doi: 10.7708/ijtte.2012.2(4).01.

[52] E. Taniguchi, R. G. Thompson, et T. Yamada, « Concepts and Visions for Urban Transport and Logistics Relating to Human Security », in Urban Transportation and Logistics, CRC Press, 2014.

[53] T. F. Fwa, « Transport and Logistics in Asian Cities », in Urban Transportation and Logistics, CRC Press, 2014.

[54] K. Bucsuházy, E. Matuchová, R. Zůvala, P. Moravcová, M. Kostíková, et R. Mikulec, « Human factors contributing to the road traffic accident occurrence », Transportation Research Procedia, vol. 45, p. 555-561, janv. 2020, doi: 10.1016/j.trpro.2020.03.057.

[55] L. Eboli, G. Mazzulla, et G. Pungillo, « How to define the accident risk level of car drivers by combining objective and subjective measures of driving style », Transportation Research Part F: Traffic Psychology and Behaviour, vol. 49, p. 29-38, août 2017, doi: 10.1016/j.trf.2017.06.004.

[56] F. Alonso, C. Esteban, J. Sanmartín, et S. A. Useche, « Reported prevalence of health conditions that affect drivers », Cogent Medicine, vol. 4, no 1, p. 1303920, janv. 2017, doi: 10.1080/2331205X.2017 .1303920.

[57] E. Taniguchi et R. G. Thompson, Éd., City Logistics: Mapping The Future. Boca Raton: CRC Press, 2014. doi: 10.1201/b17715.

[58] H.-E. LIN et R. Zito, « A review of travel-time prediction in transport and logistics », Proceedings of the Eastern Asia Society for Transportation Studies, vol. 5, janv. 2005.

[59] V. Ratanavaraha et S. Suangka, « Impacts of accident severity factors and loss values of crashes on expressways in Thailand », IATSS Research, vol. 37, no 2, p. 130-136, mars 2014, doi: 10.1016/j.iatssr.2013.07.001.

[60] J. G. V. Vieira, J. C. Fransoo, et C. D. Carvalho, « Freight distribution in megacities: Perspectives of shippers, logistics service providers and carriers », Journal of Transport Geography, vol. 46, p. 46-54, juin 2015, doi: 10.1016/j.jtrangeo.2015.05.007.

[61] S. Van der Spoel, C. Amrit, et J. Hillegersberg, « Predictive Analytics for Truck Arrival Time Estimation: A Field Study at a European

Distribution Center », International Journal of Production Research, vol. In Press, p. 1-21, janv. 2016, doi: 10.1080/00207543.2015.1064183.

[62] H. Sattar et al., « Smart Wound Hydration Monitoring Using Biosensors and Fuzzy Inference System », Wireless Communications and Mobile Computing, vol. 2019, p. e8059629, déc. 2019, doi: 10.1155/2019/8059629.

[63] M.-D. Pop, O. Proștean, T.-M. David, et G. Proștean, « Hybrid Solution Combining Kalman Filtering with Takagi–Sugeno Fuzzy Inference System for Online Car-Following Model Calibration », Sensors, vol. 20, no 19, Art. no 19, janv. 2020, doi: 10.3390/s20195539.

[64] S. Mandal, J. Choudhury, et S. Bhadra Chaudhuri, « In Search of Suitable Fuzzy Membership Function in Prediction of Time Series Data », International Journal of Computer Science Issues, vol. 9, mai 2012.

[65] M. Aydin, E. Akyuz, O. Turan, et O. Arslan, « Validation of risk analysis for ship collision in narrow waters by using fuzzy Bayesian networks approach », Ocean Engineering, vol. 231, p. 108973, juill. 2021, doi: 10.1016/j.oceaneng.2021.108973.

[66] D. S. N. A. Shamsuddin et al., « Dynamic Hazard Identification on SOFC system using Bayesian Network », International Journal of Integrated Engineering, vol. 14, no 2, Art. no 2, juin 2022.

[67] B. Jones, I. Jenkinson, Z. Yang, et J. Wang, « The use of Bayesian network modelling for maintenance planning in a manufacturing industry », Reliability Engineering & System Safety, vol. 95, no 3, p. 267-277, mars 2010, doi: 10.1016/j.ress.2009.10.007.

[68] B. Kamal et M. Aydın, « Application of fuzzy Bayesian approach on bankruptcy causes for container liner industry », Research in Transportation Business & Management, vol. 43, p. 100769, juin 2022.

[69] Z. Liu, Q. Ma, B. Cai, X. Shi, C. Zheng, et Y. Liu, « Risk coupling analysis of subsea blowout accidents based on dynamic Bayesian network and NK model », Reliability Engineering & System Safety, vol. 218, p. 108160, févr. 2022, doi: 10.1016/j.ress.2021.108160.

[70] F. Goerlandt et S. Islam, « A Bayesian Network risk model for estimating coastal maritime transportation delays following an earthquake in British Columbia », Reliability Engineering & System Safety, vol. 214, p. 107708, oct. 2021, doi: 10.1016/j.ress.2021.107708.

[71] F. Corman et P. Kecman, « Stochastic prediction of train delays in real-time using Bayesian networks », Transportation Research Part C: Emerging Technologies, vol. 95, p. 599-615, oct. 2018, doi: 10.1016/j.trc.2018.08.003.

[72] C.-L. Wu et K. Law, « Modelling the delay propagation effects of multiple resource connections in an airline network using a Bayesian network model », Transportation Research Part E: Logistics and Transportation Review, vol. 122, p. 62-77, févr. 2019.