INTERNATIONAL
RAILWAY SAFETY COUNCIL
PARIS 2 ▸ 7 OCTOBER 2016

FORECASTING THE DANGER PRECURSORS' EVOLUTION BY DEVELOPING A PREDICTIVE FUNCTION

BORDEAU Vianney

Rail Safety Expert

Department of Safety Control, RATP, Paris, France

Summary

A major component of the risk management policy of RATP (the Paris public transport operator) is to identify, control and monitor the evolution of events that may be signs of severe incidents or accidents in the transport networks.

For nearly ten years, using a statistical analysis model, the studies of the danger precursors of the RATP transport networks contribute to the periodical control and monitoring of events related to the safety. These studies generate recommendations which are systematically analyzed within the transport networks' safety committees and are subject of specific actions (e.g. preventive maintenance, curative maintenance, etc.).

The purpose of these studies is to define a predictive function in order to forecast the danger precursors' evolutions. This function is based on a probabilistic analysis of feedback data that concern safety.

The predictive model gives additional recommendations to the results of the classical model. The applicability of these recommendations is validated by calculating two confidence indexes, one for the prediction and another one for the entire predictive model.

Key words: *Danger precursors / Evolutions' forecasting / Predictive model / Confidence indexes*

1. Background

1.1. RATP Transport networks

The sixteen metro lines, two regional express railway lines, six tramway lines and nearly three hundred and fifty bus lines operated by RATP (Régie Autonomes des Transports Parisiens = Paris Transport Authority) carry several million passengers every day throughout the Paris Region.

In general, transport networks like these lines are said to be "complex systems". This complexity is due to the large amount of equipment used, the many interactions that drive them and the continuously high loads that they carry.

In terms of safety, the characteristics of RATP transport lines introduce constraints that can be conducive to the development of dangerous situations. If these situations are not controlled, they can lead to the occurrence of feared events (operating incidents or serious accidents).

The RATP relies on a set of processes that make up its risk control policy, to control these situations.

1.2. Risk control policy at RATP

Various monitoring, checking and preventive or corrective maintenance procedures have been set up to control risks related to RATP's public transport mission.

Among these procedures, trend charts that are directly concerned by our work are important tools in our risk control policy. Their purpose is to evaluate the risk level of systems that are sensitive for safety.

These trend charts are produced based on a strong concept: "*Pay attention to all precursor signs to reduce their frequency of occurrence, which is the only method of reducing the probability of occurrence of feared events*".

RATP adopts this principle and makes use of trend charts for specific monitoring of these different transport networks with a global and transverse overview of safety. This approach is focussed on particular events that have already been identified as precursor signs of dangerous situations.

The method of analysing these events is the main activity impacted by the work that we will be presenting. This method is based on a statistical analysis developed specifically for RATP's activities. This method is referred to as the "classical analysis method" throughout the remainder of this publication.

2. Purpose of the approach

The purpose of the work carried out is to improve the classical analysis method by building an anticipation function into it. In practice, this function must supply information about the future trend of changes in danger precursor events.

The benefit derived from the addition of such a function lies in the increased sensitivity of classical analysis tools. In the final analysis, the new function must make it possible to anticipate degradation "not yet" observed so that appropriate preventive or remedial actions can be anticipated to control risks.

This publication presents the principles of the predictive analysis method that we have created for construction of the new anticipation function.

3. Presentation of the classical analysis model

3.1. Definition of danger precursor events

A danger precursor event is defined as being an event forming part of an incident chain and that could lead to feared events (operating incidents or serious accidents).

In general, danger precursor events reveal the degree of safety of installations to which they belong. When their occurrence frequency increases, the degree of safety of the system becomes degraded and the probability of occurrence of a feared event increases.

In practice, they are identified within work groups composed of members of the different expertise clusters of RATP's operations, engineering and maintenance departments. Since this identification can depend on technical, environmental and organisational changes, the list of monitored precursor events can change over time. The principles of this identification have already been described in a publication for Lambda-Mu 18 [3].

Almost two hundred danger precursor events have now been identified and are monitored on all RATP's transport networks. Figure 1 presents a simplified example illustrating the concept of danger precursor events in the railway industry.

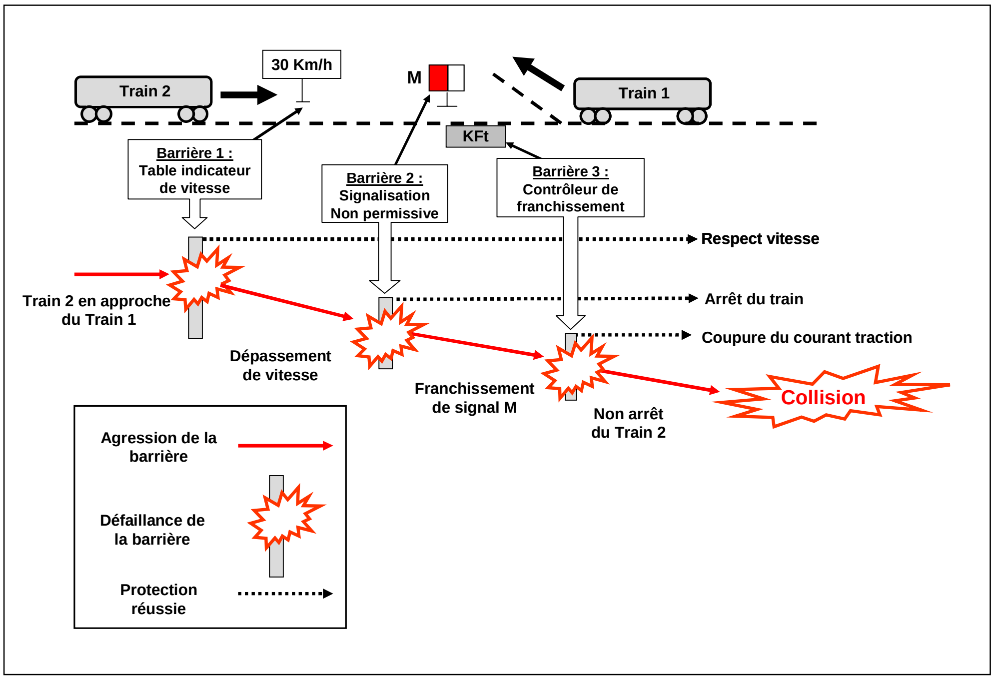

Figure 1: Example of an incident chain that could lead to a collision between two trains

In this example, the studied situation is a Train 1 that is in the process of shunting, while a Train 2 is approaching. Train 1 cannot be shunted in complete safety unless Train 2 has stopped before passing signal M.

The feared event is a collision between two trains by one train catching up with the other or a sideways collision.

- This risk is prevented by the following systems:
  - o the Speed Indicator Table (TIV) showing a speed limit of 30 km/h so that the train can stop before it reaches signal M;
  - o the non-permissive manoeuvre signal M, red light on, informing the driver that he must stop before the signal;
  - o the crossing controller KFt that automatically triggers emergency braking of the train if the non-permissive signal is not respected.
- The precursor events for the feared event in this example are:
  - o speed limit exceeded;
  - o crossings of non-permissive signals;
  - o unavailable states of crossing controllers.

3.2. Principles of "Danger alert trend charts"

The RATP produces trend charts called "Danger alert trend charts" for monitoring of danger precursor events. The purpose of these tables is to evaluate the safety state of installations at the end of every month depending on the analysis of all listed precursor events.

Danger alert trend charts are generated based on the analysis of several indicators chosen and designed to be representative of changes to safety aspects of installations.

In practice, if one or more indicators are degraded, the risks associated with the operation of the installations increase and their safety levels are degraded. Alerts are issued automatically when this type of situation is identified.

Danger alert trend charts contribute to setting up actions to maintain (or restore) the safety of installations to an acceptable level, through this warning process.

The purpose of this new function that we are presenting now is to anticipate reporting the issued alerts so that treatment plans necessary and appropriate to deal with the announced degradations can be implemented early.

3.3. Mechanism for formulating "classical" alerts

In practice, alerts announcing a degradation in the safety level of installations are issued in the form of recommendations for action. There are two types of these recommendations:

- "Apply increased vigilance", when drifts start to appear;
- "React", when the safety level is becoming unacceptable.

They are systematically forwarded to the decision making body (i.e.general management and managers of the departments concerned) and a specific review is carried out for all recommendations during safety committee meetings of the different transport networks.

Recommendations are formulated as a function of statistical analyses made on data containing counts of occurrences of danger precursor events. The currently applied statistical model is derived from work done by RATP in partnership with LAMSADE (Systems Analysis and Modelling Laboratory for Decision Making Assistance at Paris Dauphine University [1] [2].

This model is based on the analysis of trend indicators generated specifically for monitoring danger precursor events. Each of these trend indicators is a mathematical algorithm representative of a characteristic of the change in precursors.

In the classical model, three indicators have been chosen and defined so as to represent:

- A view of the monthly calendar change in which the same month is studied over several years (5 years). This makes it possible to take seasonal effects into consideration, since the solicitation on the system can vary significantly depending on the period in the year (eg lower passenger flow in August, bad weather conditions in winter, etc.);
- A short term view of changes over a 6-month period;
- A long term view of changes over a 5-year period; This view can show up changes for which the variation is not perceptible with a short term view.

4. Predictive model principle

4.1. Origin of input data

Operation of a predictive model consists of observing the state of the system being studied at the present time (time "t"), and then comparing this state with its historic variations (times "t-1", "t-2", "t-3", …, "t-n") to be able to deduce the state that it can be expected to reach in the near future (at time "t+1", or even times "t+2", "t+3", etc.), taking account of mechanisms of change in the system.

One of the prerequisites necessary for the creation of a predictive model of a given system is to construct a database for learning and understanding mechanisms by which this system changes.

RATP has a large database built up from Operating Experience covering several decades on these different transport networks. Safety-related elements of this operating experience (for example equipment malfunctions, serious incidents, accidents, etc.), are the starting point for the predictive model that we created.

Danger precursor events have been archived in a special manner in a dedicated database for more than 10 years, within the framework of a classical analysis model. Our predictive model makes use of this data source in priority.

4.2. General principle of the selected predictive model

Our predictive model is based on the construction of a simple mathematical model that uses a number of indicators representative of changes in studied systems. These indicators that are intended to be predictive and are called "predictive indicators" are calculated and compared with threshold values so as to define the state of the studied system at a given moment.

The criteria defining the state of the system are compiled and then used in a probabilistic model that forms the learning function of our model.

A deliberate decision was made that the method should not require any understanding of the physical or organisational methods involved in studied systems, so that change mechanisms not found by conventional expertise centres can be highlighted. It is based solely on the analysis of the probabilistic model derived from operating experience.

This method is inspired from analysis principles used by the *Société de Calcul Mathématique* (Mathematical Calculation Society). These principles have already been applied in different activity sectors (eg forecasting of variations in the price of real estate, electricity or raw materials, forecasting of world automobile sales, etc.) [3].

4.3. Definition of predictive indicators

As for the classical analysis method, input data to our predictive model are processed using several indicators that highlight particular characteristics of a change in the number of occurrences of danger precursor events.

We selected three indicators for our predictive model that our similar to indicators used in the classical model:

- Monthly predictive indicator;

- Short term predictive indicator;
- Long term predictive indicator

The common principle adopted for these three indicators is based on the fact that as the indicator considered increasingly characterises a degrading state of the system, the number of occurrences of danger precursor events in this system will increase, with a corresponding increase in the associated risk(s).

4.3.1. Monthly predictive indicator;

The monthly predictive indicator is based on a comparison of data for the current month with two threshold values. These values are determined for a given month based on data for the same month during the previous 5 years. The calculation of threshold values is an average calculation weighted by the age and the value of each data.

Thus for the low threshold that is the "optimistic" average, the weight of the data decreases as the data becomes older, and the weight of the data increases as its value reduces.

For the high threshold that is a "pessimistic" average, the weight of the data decreases as the data becomes older, and the weight of the data increases as its value increases.

The value of the monthly predictive indicator is compared with these two thresholds and interpreted in three ways according to the following three criteria:

- For values less than the low threshold, the safety state of the installations is not significantly degraded
  ➔ For the monthly calendar change, the system is in the **GREEN**.
- For values between the low threshold and the high threshold, the safety state of the installations is starting to degrade
  ➔ For the monthly calendar change, the system is in the **ORANGE**.
- For values higher than the high threshold, the safety state of the installations is degrading quickly or is degraded
  ➔ For the monthly calendar change, the system is in the **RED**.

For the predictive model, these three states (**GREEN** $_{\text{monthly}}$ / **Orange** $_{\text{monthly}}$ / **RED** $_{\text{monthly}}$) are recognised through the monthly calendar change of danger precursor events.

4.3.2. Short term predictive indicator

The long term predictive indicator is used to evaluate the change in the number of danger precursor events over a short period of 6 months. The value of the short term predictive indicator is determined from the calculation of the slope of the curve of the variation of the danger precursor event considered over 6 months.

The value of the short term predictive indicator is interpreted in two ways using two criteria as follows:

- If the indicator is negative (the slope over the last six months is negative) then the safety state of the installations is not significantly degraded
  ➔ For the short term change, the system is in the **GREEN**.
- If the indicator is zero or positive (the slope over the last six months is positive or zero) then the safety state of the installations is beginning to degrade, is degrading quickly or is already degraded
  ➔ For the short term change, the system is in the **RED**.

For the predictive model, these two states (**GREEN** $_{\text{short term}}$ / **RED** $_{\text{short term}}$) are recognised through the short term change of danger precursor events.

4.3.3. Long term predictive indicator

The long term predictive indicator is used to evaluate the change in the number of danger precursor events over a period of 5 years continuously The value of the long term predictive indicator is determined in two calculation steps.

Firstly, five annual averages of the number of events are calculated over periods of 12 consecutive months in the last five years considered.

Secondly, the slope of the curve showing the variation of these five annual averages is calculated.

The calculated slope that forms the value of the long term predictive indicator is interpreted in two ways using two criteria as follows:

- If it is negative, the safety state of the installations is not significantly degraded
  ➔ For the long term change, the system is in the **GREEN**.
- If it is positive or zero, then the safety state of the installations is beginning to degrade, is degrading quickly or is already degraded
  ➔ For the long term change, the system is in the **RED**.

For the predictive model, these two states (**GREEN** $_{\text{long term}}$ / **RED** $_{\text{long term}}$) are recognised through the long term change in danger precursor events.

4.4. Combination of predictive indicators

Globally, the analysis of three predictive indicators recognises:

- Three state criteria for the monthly calendar change: **GREEN monthly** / **Orange monthly** / **RED monthly**
- Two state criteria for the short term change: **GREEN short term** / **RED short term**
- Two state criteria for the long term change: **GREEN long term** / **RED long term**

The combination of these criteria generates a model with twelve triplets, that for our predictive model constitutes all possible states of change of danger precursor events.

Table 1 below shows the twelve triplets in our model.

	Monthly predictive indicator;	Short term predictive indicator	Long term predictive indicator
Triplet 1	GREEN monthly	GREEN short term	GREEN long term
Triplet 2	ORANGE monthly	GREEN short term	GREEN long term
Triplet 3	RED monthly	GREEN short term	GREEN long term
Triplet 4	GREEN monthly	RED short term	GREEN long term
Triplet 5	ORANGE monthly	RED short term	GREEN long term
Triplet 6	RED monthly	RED short term	GREEN long term
Triplet 7	GREEN monthly	GREEN short term	RED long term
Triplet 8	ORANGE monthly	GREEN short term	RED long term
Triplet 9	RED monthly	GREEN short term	RED long term
Triplet 10	GREEN monthly	RED short term	RED long term
Triplet 11	ORANGE monthly	RED short term	RED long term
Triplet 12	RED monthly	RED short term	RED long term

Table 1: Combination of triplets representing the different states of change of danger precursor events

4.5. Quantification of probabilities of an increase in danger precursor events

The combination of twelve triplets represents the state of a danger precursor event at a given time in the present or the past (at times "t", "t-1", "t-2", "t-3", …, "t-n"). The prediction must evaluate if the occurrence of this event will increase or reduce in the near future (at time "t+1", or even at times "t+2", "t+3", etc.).

Two states of change of danger precursor events are selected for the prediction:

- The state in which the occurrence of precursors in "increasing" ➔ Sign of a degradation in the safety level,
- The state in which the occurrence of precursors in "reducing" ➔ Sign of an improvement in the safety level,

The probability of degradation is quantified by saving the data pair formed firstly by the triplets observed for each month M of the history, and secondly by the state of change of the precursor for months "M+1".

Quantification is then made by counting the number of times in which the occurrence of precursor events for a given triplet in month M was "up" in month "M+1".

This approach applied to each of the twelve defined triplets can be used to produce a statistical prediction table that takes account of all quantifications made based on available operating experience.

For each prediction, when the proportion of "increased" precursor occurrences expressed as a percentage is more than 50%, this indicates that the studied precursor is likely to be degraded during the following month. The probability of the prediction increases as this value increases. This value expressed as a percentage is used as a prediction confidence index to evaluate the reliability of the prediction.

Applied to the complete history, the model quantifies the probabilities associated with the change of all triplets, by counting the number of times in which the precursor was "up" and the number of times that it was "down", during a reference period. This provides a means of generating a statistical prediction table for each precursor.

Table 2 present an example prediction table obtained with one of the danger precursor events that we have analysed.

	Évolution statistique au mois « M+1 »			
	Nombre de Hausses constatées	Nombre de Baisses constatées	Probabilité d'évolution	
			Hausse	Baisse
Triplet 1	0	0	-	-
Triplet 2	0	0	-	-
Triplet 3	0	0	-	-
Triplet 4	14	25	35,9%	64,1%
Triplet 5	30	20	60,0%	40,0%
Triplet 6	23	4	85,2%	14,8%
Triplet 7	0	0	-	-
Triplet 8	0	0	-	-
Triplet 9	0	0	-	-
Triplet 10	4	6	40,0%	60,0%
Triplet 11	10	4	71,4%	28,6%
Triplet 12	3	0	100,0%	0,0%

Triplets qui ne se sont jamais présentés

Triplet suivi à 85,2% d'une Augmentation et à 14,8% d'une Baisse

Triplet toujours suivi d'une Hausse

Table 2: Example statistical prediction table

Three types of triplets are observed based on this example:

- Triplets that have never occurred during the life of the installation considered. These can be identified in the prediction table by the number of increases and the number of reductions both being zero during month "M+1".
- Triplets for which the history indicates that only increases (or only reductions) will be possible in month "M+1". These are identified with a probability of an increase (or reduction) equal to 100%.
- Triplets for which the history indicates that increases as well as reductions in the number of occurrences of the precursor event considered will be possible during month "M+1". These triplets have a non-zero probability of an increase or a reduction.

4.6. Principles of the anticipation function

Predictions formulated by our model are made using a set of operations that make up the new anticipation function. This function is based on the use of statistical prediction tables that are defined for each danger precursor event and are permanently updated with operations and maintenance data.

The anticipation function performs the following operations:

1) The three predictive indicators of the precursor event studied for the current month "M" are calculated.

   For instance:

   - Monthly predictive indicator ➔ **RED $_{\text{monthly}}$**;
   - Short term predictive indicator ➔ **RED $_{\text{short term}}$**;
   - Long term predictive indicator ➔ **RED $_{\text{long term}}$**;

2) The indicator triplet for the precursor event studied for month "M" is defined.

   For example, **Triplet 6** for the **{RED $_{\text{monthly}}$ / RED $_{\text{short term}}$ /GREEN $_{\text{long term}}$}** predictive indicators combination<

3) A predicted change for the next month "M+1" is given based on probabilities of change defined in the prediction table.

   For example, according to the prediction table in Table 2, the probability that the number of occurrences of the danger precursor event for Triplet 6 will increase during the next month will be **85.2%**. The probability of it reducing will be 14.8%.

4) The prediction formulated in the classical analysis model is included, taking account of the safety level of the installation that is given by the classical model.

   For example, when an increase in the number of occurrences of the danger precursor event is predicted:

- If it induces the beginning of a drift in the safety level, then an early vigilance alert is given one month in advance;
- If it induces a significant degradation in the safety level, then an alert requiring a reaction is given. This can be materialised in different ways depending on the configuration (eg increased monitoring of the equipment concerned, preventive maintenance operations, etc.).

4.7. Evaluation of the degree of confidence given to the model

The degree of confidence of the model consisting of the predictions of each statistical table is evaluated by checking these predictions for all data in the history.

In practice, each state of the precursor event (triplet in month "M") and its change during the next month (up or down in month "M+1") is compared with the predictions in the statistical prediction table of the precursor. The model is more reliable if the predictions are more accurate.

As a first approach, for this study phase it is considered that a model is reliable for a given precursor when at least 2/3 of the history of the precursor is verified by the model.

The proportion of correct predictions (as a ratio of 10) is used as a prediction confidence index, by which the reliability of the model can be evaluated.

The uncertainty associated with this type of prediction is quantified and forms an applicability indicator of the method. There can be situations in which it is impossible to make any prediction. In this case, all the model does is to apply the existing methodology. There is no question about the validity of the principles currently used to produce trend charts. In other words, if it is impossible to predict a trend, the calculations and the resulting analysis are exactly the same as in the classical analysis model.

5. Conclusion

The work done by RATP and presented in this publication has been used to define a reference method for a predictive model applicable for monitoring danger precursor events. Figure 2 summarises the principles of this model.

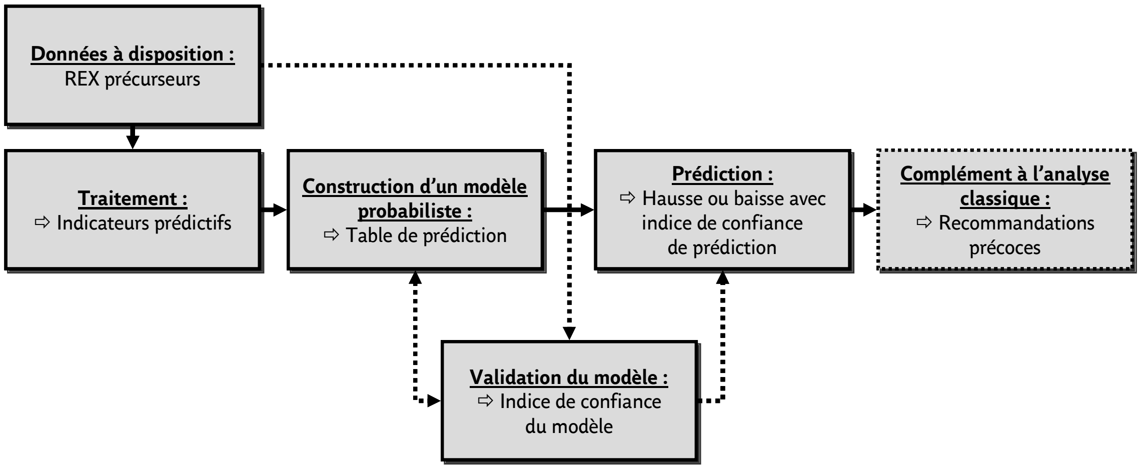

Figure 2: Principle of the selected predictive model

The reference method for the developed anticipation function is based on the analysis of predictive indicators generated so that they can be included in a probabilistic model that does not require an understanding of any phenomenon explaining the mechanism for change in the studied systems.

The uncertainty associated with this type of prediction is quantified and forms an applicability indicator of the method. In this respect, it should be noted that situations can occur in which it is impossible to make any prediction.

The predictive model was applied on almost half of all danger precursor events. The results of this test demonstrate that the predictive model, as it is defined for this first study phase, functions with satisfactory confidence indexes for more than a third of the events that were monitored.

These results have to be considering remembering that it is a probabilistic model for which the predictions are not systematically true. Consequently, the predictive results are always accompanied by prediction confidence indexes and a model confidence index used to give an opinion about the reliability of the complementary recommendations given.

Not that for another third of the monitored events, there is not enough history for our predictive model that is based on operating experience to provide a prediction with a satisfactory confidence index.

In terms of prospects, the work presented is still in progress and there are many approaches for improvement of the predictive model. Approaches considered at the present time are related to the definition of predictive indicators based on new mathematical models with a simpler prediction table than the current model composed of twelve triplets.

For this work, like the predictive model that we have presented, any improvement will be based on a validation process based on a check of the model made on the existing and future history of the precursors.

In the long term, if the applicability parameters of the model are good (high confidence indexes), it will be possible to envisage integrating the predictive model into the process for monitoring danger precursor events.

6. References

[1] T.-J. TANZI, B. ROY, M. FLAGES, D. VONCKEN, Indicateurs de dangerosité appliqués au transport collectifs (Dangerosity indicators applied to public transport), Publication Lambda-Mu 12 (Montpellier), p. 703 to 708, 2000

[2] D. VONCKEN, B. ROY, Indicateurs d'alerte des dangers (Rapport final) (Danger alert indicators (Final report)), RATP Document, 2004

[3] Société de Calcul Mathématique (SA), Construction d'indicateurs prospectifs par des méthodes probabilistes robustes (The use of robust probabilistic methods for the construction of prospective indicators), 2003 http://scmsa.eu/fiches/SCM_Indicateurs_prospectifs.pdf

[4] F. PARRENNES, A. COINTET, Maîtrise des risques d'un système de transport - Identification des précurseurs de dangers (Risk control in a transport system - Identification of danger precursors), Publication Lambda-Mu 18 (Tours), p. 454 to 461, 2012