Safety Instrumented Systems operated in the Intermediate Demand Mode

S.Eisinger & L.F.Oliveira
DNV GL, Oslo & Rio

K.Tveit & B.Natvig
University of Oslo, Norway

ABSTRACT: When analysing critical systems the demand frequency is crucial. Often the low and the high demand mode are distinguished. In this paper the intermediate demand mode is analysed.
The results from the analyses of the example (two channel) model show that the hazard rate exhibits unexpected behaviour in the intermediate demand region. As far as can be seen from the analysis, the standard Probability of Failure on Demand (PFD) formulas are usable, but they become exceedingly conservative as one moves into the intermediate demand region. On the other hand, usage of the standard formulas for the hazard rate (PFH) (high demand mode) in the intermediate region may lead to non-conservative results. Therefore, whenever a system seems to be operated in this intermediate demand mode, or even only close to it is advisable to perform more accurate analysis compared to standard PFD and PFH formulas. It has been demonstrated that such analysis is readily feasible using modern simulation tools. Operational or maintenance details should be easy to accommodate on top of the issues handled in this article. The knowledge of rare event handling techniques may be necessary. For the operator it is necessary to perform the required tests and documentation after demands in a proper way.

1 INTRODUCTION

For Safety Instrumented Systems, demands on the Safety Function are obviously crucial and may lead to hazards if the Safety System does not react in the specified way. Safety-critical component failures are often not detectable during normal operation. For such systems, if demands happen relatively seldom proof tests may be specified which detect the failures. Obviously proof tests should be performed more frequent than the occurrence of demands. Systems where this is clearly possible are said to be operated in low demand mode. Fire detection represents an example for such a system.

On the other hand systems exist where demands occur relatively frequent and proof tests with an even higher frequency do not make sense. The safety protection must be established in different ways, e.g. through redundancy. Such systems are said to be operated in high demand mode. An example of such a system is given by railway interlocking systems.

Safety Standards like (IEC61508 2010) treat the two demand modes as completely distinguishable with requirements that seem to be separate from each other. Table 1 shows the target failure measures for both low and high demand mode. For low demand mode the average Probability of Failure on Demand (PFD) is used and for high demand mode the average frequency for dangerous failures (PFH). Note that the latter is called Tolerable Hazard Rate (THR) in the railway industry (see (EN50126 1999)). Note also that the PFD cannot directly be used as acceptance criteria - the expected demand rate needs always to be specified. (IEC61508 2010) uses a criterion $\delta < 1y$ (with δ: demand frequency) for the low demand range.

Table 1: Safety Integrity Levels - target failure measures for a safety function (according to IEC61508)

SIL	PFD_{avg} (low demand)	PFH_{avg} (high demand)
4	$\geq 10^{-5} \ldots > 10^{-4}$	$\geq 10^{-9}/h \ldots > 10^{-8}/h$
3	$\geq 10^{-4} \ldots > 10^{-3}$	$\geq 10^{-8}/h \ldots > 10^{-7}/h$
2	$\geq 10^{-3} \ldots > 10^{-2}$	$\geq 10^{-7}/h \ldots > 10^{-6}/h$
1	$\geq 10^{-2} \ldots > 10^{-1}$	$\geq 10^{-6}/h \ldots > 10^{-5}/h$

In reality systems exist, which cannot be clearly placed and might be called intermediate demand mode systems. The present paper discusses this intermediate mode.

The issue of utilising demands as test has not been discussed extensively, but some authors have addressed it with varying focus. In (L.F.Oliveira, R.Youngblood, & P.F.F.Melo 1990) similar systems

as the one discussed here have been analysed. More recently (Y.Liu & M.Rausand 2011) have taken up the issue again using similar systems but focusing on the demand duration. All publications that we are aware of are restricted to the Markov assumption which can be overcome using the analysis techniques discussed here.

2 THE MODELS

The analysis of intermediate demand mode systems is not straight forward due to the fact that there is a combination of periodic tests, repair times and demands. The latter are at least not periodic and are often assumed random, with a constant demand rate δ. In the extreme regions of (very) low demand rate or (very) high demand rate the system reliability can be readily approximated to a good level of accuracy (see Section 2.1). Another complication is given by the component and system level of detail. While failures, repair and proof testing happens on component level, demand and hazards happen on system level. Component level analysis can be performed by (partial) Markov Analysis, but the extension to the system level renders the analysis at least rather complex and limited to the Markov assumptions.

One method which overcomes all these difficulties is given by Discrete Event Simulation. It shall be demonstrated that even the Rare Events Problem (see (rareEvents 2015)), which is often a challenge in safety system analysis based on simulation can be solved in a satisfactory way.

As the system to be analysed here clearly involves states, generalised state modelling represents a good choice for model representation both on the component and on the system level. The following generalisations with respect to standard Markov State Models are utilised

- The standard Markov assumption that a state transition is only dependent on the current state is not needed. This means also that the involved statistical distributions do not need to be exponential.

- States can have a structure including sub-state systems as serial or parallel systems. This feature is implemented to counter the general tendency that 'flat' state systems can get rather involved even with a moderate amount of states. For the present purpose components are implemented as parallel sub-state systems. The system level issues are modelled in another parallel sub-state system. In this way the model is kept modular, easy to understand and straightforward to extend to e.g. other system configurations like 2oo4.

- States can have variables related to the whole state system or to sub-systems. This feature turns states into pseudo states in the sense that a state may contain many states as always only the state occupation together with all related variable values fully define the state.

State models thus generalised where proposed by Harel(see (Harel 1987)), which represents also the implementation chosen in this project. The modelling techniques resembles the Petri Net Models (see (Y.Dutuit, F.Innal, A.Rauzy, & J.Signoret 2008)) which have recently been suggested for safety sytem calculations. Both modelling techniques fall into the same class of state-based discrete event simulation, but we believe that the Harel State Charts used here are more intuitive to understand and communicate.

On the component level a rather simple repairable component is modelled. Failures happen with a constant rate λ and are assumed hidden until they are detected by either a demand or a test. Repair takes a time MTTR $= 1/\mu$. In the present article we assume that also this time is exponentially distributed. The simplicity of the model is mostly triggered by the wish to be able to compare our results with previously published results. Most assumptions can be made less stringent and more realistic within the framework of the present analysis.

The Harel State Chart model simulated by ExtendSim for one component is shown in figure 1. The model shows the main states working, undetected failure and repair. The model knows if a failure is detected by demand or proof-test, and is aware of possible demands during the repair time of the system.

Note that this component model has two inputs for the triggers when proof tests are performed or when a demand happens. These are system properties which must therefore come from the super model.

The model of figure 1 runs into a rare-event problem (see (rareEvents 2015)) for high demand rates. This rare event problem is caused by the fact that most demands find the system with all components working and only relatively few demands find one component in the failed state - thereby detecting this failure and initiating repair actions. Even fewer demands cause a system hazard, namely the demands which find both components in a non-working state. Obviously, the system hazard represents the rare event and the many demands which find everything working represent the events which are not really interesting for the analysis, but which use up most of the processing time during a simulation. This problem description contains already the solution to the problem: demands do not really need to be made explicit when not needed - only when at least one component has failed the demand has a function to the system. Moreover, as it is assumed that demands arrive independently from each other, demand generation is not dependent on previous demands and it is thus sufficient to calculate the next demand when a situation arises where this needs to be known, namely when at least one component has failed. This strategy is followed in

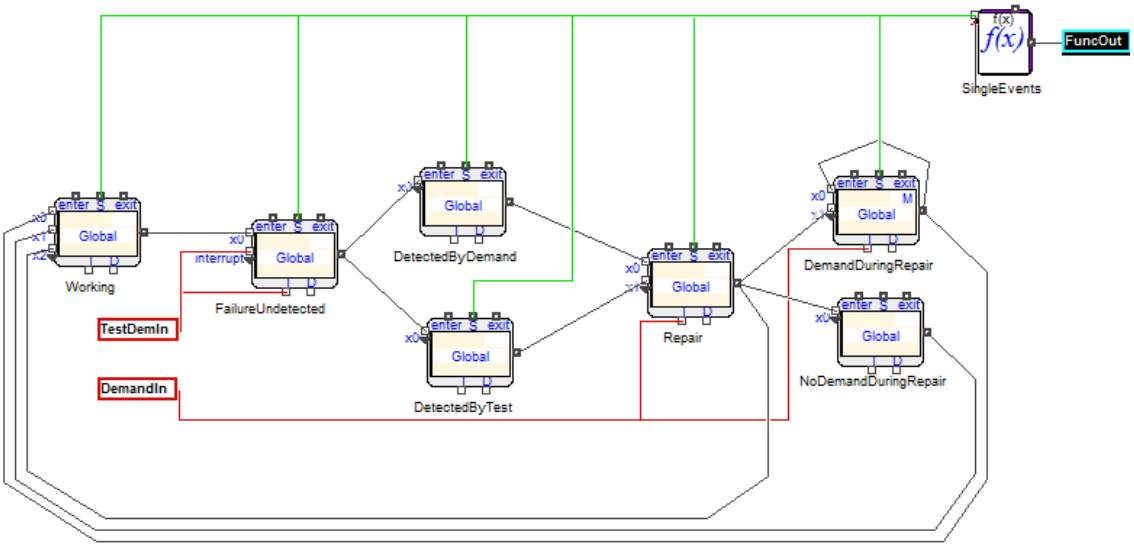

Figure 1: State model of component sub-system

a variant model to figure 1, where the demand input is omitted and the time for the next demand is kept as a system variable. The time for the next demand is calculated by any component which fails and is available for all components in the system.

A similar rare event problem exists in the low demand mode region. When the demand frequency gets low the hazard frequency gets likewise low, but system proof tests are still performed using valuable processing time. Similar to the discussion above, observing that most tests are not actually important for the analysis (namely the tests when all components are working), tests can simply be generated when needed, i.e. when at least one component has failed. In the case of the proof test there is complete dependency between tests, such that it is again possible to calculate the next proof test time at any time of the simulation using the formula 1.

$$t_{\text{nextTest}} = t + \tau - (t \mod \tau) \tag{1}$$

Also in this case the test input is omitted and the time for the next test is kept as a system variable which is only updated 'just in time', when at least one component fails.

In high demand mode a single component system does not really make sense in critical applications: either the failure mode in question can be excluded as incredible or redundancy is needed as it is impossible to detect failures and bring the system into a safe state if there is only one component and a high demand frequency. This article is restricted to two component systems as the simplest extension to a single system. The two component system is shown in figure 2 and follows the same rules as given in (L.F.Oliveira, R.Youngblood, & P.F.F.Melo 1990). "C1" and "C2" represents the single channel system illustrated in figure 2.

We distinguish between two models:

online model During repair the system is fully in use. This includes also the possibility that demands are received during repair, even if both components are not working.

offline model The system is still in use if one component has failed. If both components have failed and the failures are detected, the system is taken offline for repair.

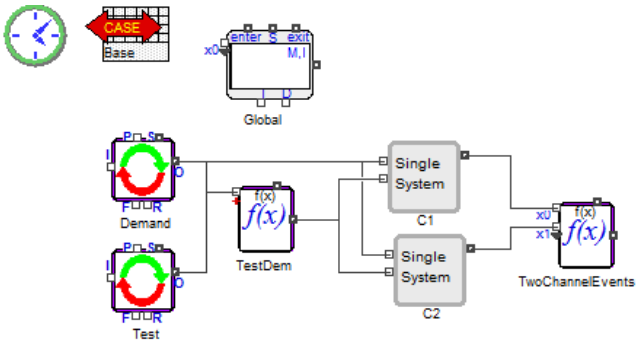

Figure 2: Direct model - reliability model of two component system

The related state diagram is shown in figure 3, implementing the states

State 1 both channels are up

State 2 one channel is up, and the other is down, but failure is undetected

State 3 both channels are down, but failures are undetected

State 4 one channel is up, and the other is under repair (its failure has been detected due to demand)

State 5 one channel is down, but undetected, and the other is under repair

State 6 both channels are down, and their failures have been detected due to demand. Note that the transitions from state 6 are somewhat different from (L.F.Oliveira, R.Youngblood, & P.F.F.Melo 1990) due to the fact we assume that both repairs can be done simultaneously.

Figure 2 represents the two component system model while figure 3 represents the component sub-model, which resides in the blocks “C1” and “C2” of the system model. The two figures 2 and 3 illustrate very well the different approach in Harel State Charts modelling compared to traditional state charts. In many ways the system model of figure 2 resembles a Reliability Block Diagram (ref. (A.Høyland & M.Rausand 1994)), but it is in fact more than that because the ”TwoChannelEvents” block keeps track of which state each of the components are in at all times. In that way this block contains the relevant states that are illustrated in the state diagram. The model in figure 2 is very well modularised and can be extended to more components in a straightforward way. The model of figure 3 does not offer that. Moreover, Markov modelling is also limited when it comes to the choice of distributions, maintenance details and system safety strategy. This simplified model has mainly been chosen for comparison with previous work.

The blocks in addition to “C1” and “C2” in the system model of figure 2 have the following purpose

Global Global variable settings which are available for all sub-state models. In our case these are λ, δ, μ, τ, $t_{\text{nextDemand}}$ and t_{nextTest}. Note that the first three of these could be component level variables (and be chosen different from eachother). Here they are only added for convenience, since they are chosen equal for all components.

Demand The demand generator. This block triggers demands and communicates them to the components.

Test The test generator. This block triggers proof tests and communicates them to the components.

TestDem Since the component blocks need only to know the demands and the combined demands and tests, “TestDem” generates the combined signal from these triggers.

The model of figure 2 represents the model without treatment of rare events. For optimised treatment of rare events the model must be modified into the model shown in figure 4.

Clearly the explicit generation of demands and tests is not present any more in figure 4. In the case of online repair another rare event problem is revealed, namely the demands during repair of both components, which become many events in the case of high demand frequency. Instead of explicitly generating these demands, only the state ‘DemandDuringRep’ is modelled. When this state finishes a representative number of additional demands is sampled through a Poisson distribution according to the demand rate and the time interval. This issue represents a solution to a system level rare event problem.

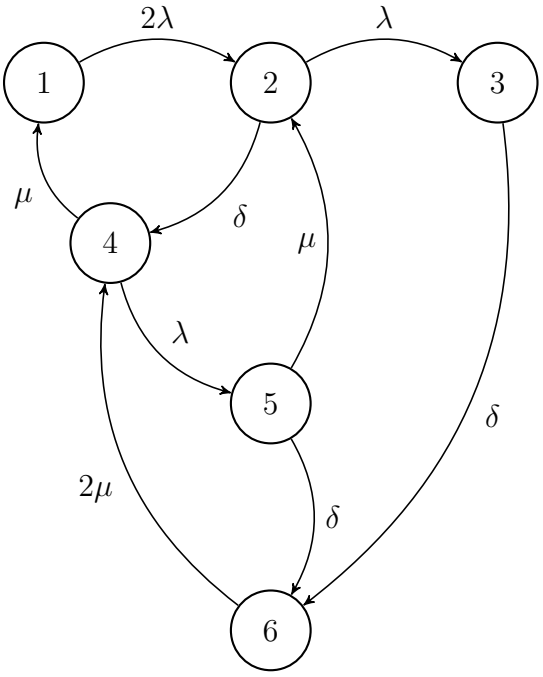

Figure 3: State diagram of the two-channel model where each transfer of state is exponentially distributed with the corresponding parameter.

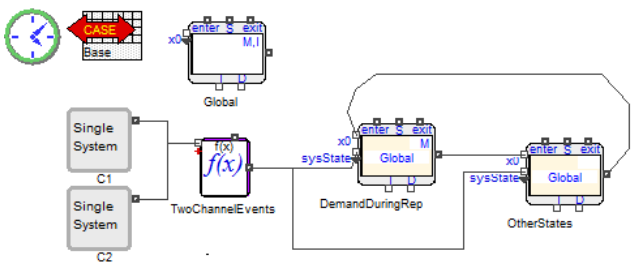

Figure 4: Reliability model of two component system optimised for dealing with the rare event problem

With respect to figure 3 the hazard rate for an offline case is found by:

$$\hat{\eta}_{\text{offline}} = \frac{\text{\# events in states 3 and 5}}{\text{simulation time}} \tag{2}$$

because these two states represent hazardous events when a demand occurs. For the online case the hazard rate is:

$$\hat{\eta}_{\text{online}} = \frac{\text{\# events in states 3, 5 and 6}}{\text{simulation time}} \tag{3}$$

where the additional state 6 represents the additional demands during repair discussed above.

2.1 *Asymptotes*

The asymptotes for the hazard rate for small and high demand rates can be calculated analytically.

In the low demand range the demands de-couple from the failures such that the traditional PFD can be calculated for a two channel system. The hazard rate becomes (see (IEC61508 2010), part 6, B.3.2.2)

$$\eta_{\text{low demand}} \simeq \delta \cdot 2\lambda^2 \left(\frac{\tau}{2} + \frac{1}{\mu}\right) \cdot \left(\frac{\tau}{3} + \frac{1}{\mu}\right) \tag{4}$$

This formula can be derived through Markov analysis or through reasoning about failure rates and equivalend down times

In the high demand range the repair time dominates the hazards. In the case of offline repair the state 5 of figure 3 dominates. I.e. one channel is under repair and the other fails and the failure is detected by the demand. This leads to the formula

$$\eta_{\text{high demand offline}} \simeq \frac{2\lambda^2\mu}{\lambda^2 + 2\lambda\mu + \mu^2} \tag{5}$$

In the case of online repair, the additional failures during the time when both components are repaired come in addition and are dominant for very high demand rates. The respective formula becomes

$$\eta_{\text{high demand online}} \simeq \frac{\delta\lambda^2}{\lambda^2 + 2\lambda\mu + \mu^2} \tag{6}$$

The last two equations are either obtained through calculating the equilibrium Markov solutions or through approximations with respect to repairable systems (see also (L.F.Oliveira, R.Youngblood, & P.F.F.Melo 1990).

3 RESULTS

The problem at hand and the models introduced in section 2 contain the following parameters

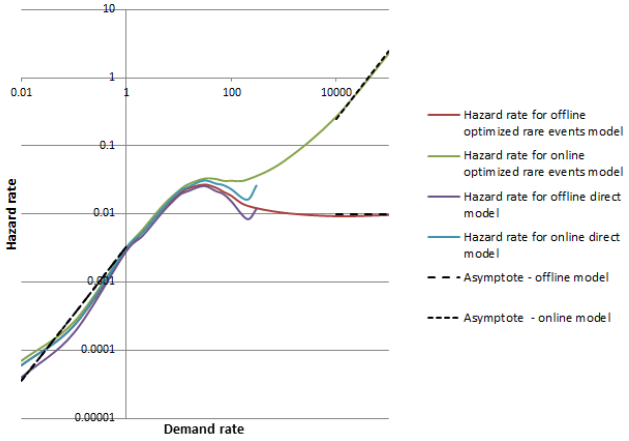

Figure 5: Offline repair results for the direct model and the rare events optimised model for $\tau = 0.1$ over a wide demand range

Failure rate λ The rate at which the components of the system fail. It is assumed that λ is constant and that the failure rates of all components of the system are equal.

Demand rate δ The rate of demands on the safety system. This is a system parameter.

Proof test interval τ The interval for proof tests of the system components. It is assumed that proof tests are performed periodically and that all components are tested at the same time.

Repair rate μ The repair rate $\mu = 1/\text{MTTR}$ for a component after a failure is detected. Within generalised state modelling it is not necessary to assume a constant repair rate. In any case, when a failure mode is known it is often more realistic to assume a constant repair time. Still, in this paper a constant rate is assumed for easy comparison with previous work.

Without loss of generality $\lambda = 1$ is set throughout this paper, i.e. the time unit is set equal to the mean time between failures of a single component. As repair rate $\mu = 200$ is used as a 'typical' repair rate.

Results for $\tau = 0.1$ are shown in figure 5. It is clearly seen that the direct model is limited in the demand range at least in the high demand mode area both for online and offline repair. For demand rates above about 100λ the simulation times for the direct model become too long to be practicably feasible. In the area where both models can produce results, the results coincide well within statistical accuracy. The rare events problem in the low demand range does not become visible for demand rates down to 0.01λ.

With the choice of time scale as λ and $\mu = 200$ as typical values, this leaves two parameters to be varied in a suitable range and the resulting system hazard rate. The results for offline repair are shown in figure 6. Similar results for online repair are shown in figure 7. As the frequency of proof-tests decreases, the intermediate mode has a greater effect on the system. There is a larger deviation from the simulated results

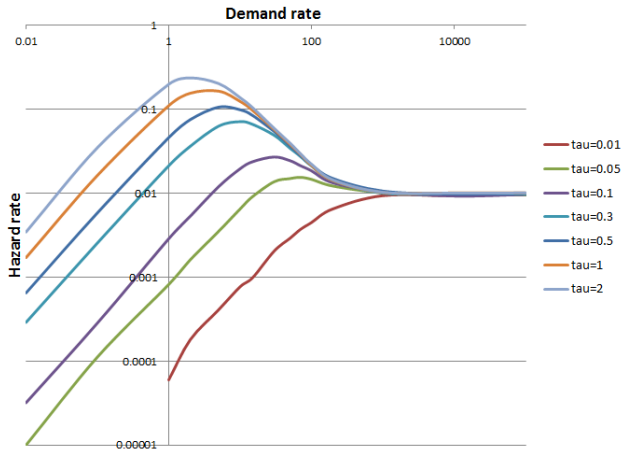

Figure 6: Offline repair results using the rare events optimised model for various τ over a wide demand range. $\mu = 200$

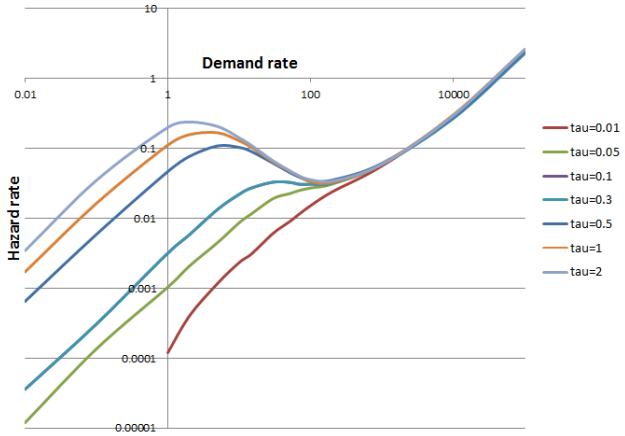

Figure 7: Online repair results using the rare events optimised model for various τ over a wide demand range. $\mu = 200$

and the asymptotic formulas normally used for PFD and PFH calculation.

The asymptotes as discussed in section 2.1 are confirmed well in all plots, as illustrated for one case ($\tau = 0.1$) in figure 5. When the demand rate increases the hazard rate for the offline model approaches towards the hazard rate given by the assymptotic equation 5. For the online model, equation 6 shows that the hazard rate increases with the demand rate.

The plots exhibit an unexpected pair of extreme points which are most marked for large proof test times. The top point is due to the fact that demands become effective as tests when the demand rate increases. In this way failures of single components are detected earlier, reducing the chance for double failures and hazards. On the other hand there is the repair time which contradicts this effect since failures and demands during repair can increase the hazard rate. The asymptotic formula which explains the low demand region does not take these effects into account. The effect diminishes when the proof test interval is reduced and seems to vanish altogether for very small proof test intervals. There is a similar dependency on μ which is not elaborated here. Together these results confirm the above explanation of the pair of extreme points.

4 DISCUSSION AND CONCLUSIONS

The results from section 3 show clearly that the hazard rate exhibits unexpected behaviour in the intermediate demand region. As far as can be seen from the analysis, the standard PFD formulas are usable, but they become exceedingly conservative as one moves into the intermediate demand region. According to (IEC61508 2010) the PFH formula should be used for $\delta > 1y$. In this case the asymptotic hazard rate renders non-conservative results in the intermediate region . Therefore, whenever a system seems to be operated in this demand mode, or even only close to it, it is advisable to perform more accurate analysis compared to standard PFD and PFH formulas. It has been demonstrated that such analysis is readily feasible using modern simulation tools. Operational or maintenance details should be easy to accommodate on top of the issues handled in this article. The knowledge of rare event handling techniques may be necessary.

As the usage of demands as effective tests is crucial for gaining the advantage of an improved hazard rate it is important that

- demands are properly recorded in relevant systems
- the necessary tests on the components are performed and the results recorded, such that the demand can actually be used as an effective test

The analysis performed here can be extended towards a number of additional points in order to better understand the details. Without claiming completeness the following issues would be interesting

- systematic analysis of the dependencies on the repair rate μ
- more realistic distributions (e.g. constant repair time)
- other system architectures (e.g. more general koon architectures including also common cause failures)
- other possible maintenance strategies (e.g. take the system offline when only one working component is left)

REFERENCES

A.Høyland & M.Rausand (1994). *System reliability theory; models, statistical methods and applications* (2 ed.). Wiley.

EN50126 (1999). *Railway applications - The specification and demonstration of Reliability, Availability, Maintainability and Safety (RAMS)* (1 ed.). CENELEC.

Harel, D. (1987). Statecharts: A visual formalism for complex systems. *Science of Computer Programming 8*, 231–274.

IEC61508 (2010). *Functional safety of electrical/electronic/programmable electronic safety-related systems* (2.0 ed.). iec.ch: IEC.

L.F.Oliveira, R.Youngblood, & P.F.F.Melo (1990). Hazard rate of a plant equipped with a two-channel protective system subject to a high demand rate. *Reliability Engineering and System Safety 28*, 35–58.

rareEvents (2015). Rare event sampling. *Wikipedia. en.wikipedia.org*.

Y.Dutuit, F.Innal, A.Rauzy, & J.Signoret (2008). Probabilistic assessments in relationship with safety integrity levels by using fault trees. *Reliability Engineering and System Safety 93*, 1867–1876.

Y.Liu & M.Rausand (2011). Reliability assessment of safety instrumented systems subject to different demand modes. *Journal of Loss Prevention in the Process Industries 24*, 49–56.