Different Principles Used for Determination of Tolerable Hazard Rates

Dr. Hendrik Schäbe
Institute for Software, Electronics, Railroad Technology
TÜV InterTraffic GmbH
51105 Köln
e-mail: schaebe@iseb.com

1. INTRODUCTION

1.1 The Safety Integrity Level
The concept of safety integrity levels (SIL) has been applied in several standards that are used in different areas of technology. The basic safety standard IEC 61508 [1] defines a SIL for electronic control systems. Also, the SIL is introduced into railway technology by EN 50126 [2], 50128 [3] and 50129 [4]. For defence technology, the British DEF STAN 00 56 [5] defines a SIL.
A SIL is usually associated with a system function or a subsystem and it is used for two purposes:
First, a certain SIL is used to give an interval for a rate of safety critical failures. This characteristic applies to so called „random faults“, i.e. failures that occur in an unpredictable manner. Mostly, these faults are caused and accompanied by intrinsic physical processes such as ageing.
Second, a SIL defines measures to be applied in the design and during the manufacturing process to keep the frequency of occurrence of so called „systematic faults“ small in comparison with random faults. The reason for systematic faults is mainly a design error or a manufacturing process error that causes failures of identical replications of the same type of component or equipment under similar circumstances. These faults might reveal themselves also in the form of common cause failures.
Usually, the higher the SIL, the harder the requirements for the system function. In many cases, SIL4 is the highest SIL, whereas SIL1 is the SIL with the lowest requirements. In addition, there can be system functions that do not even fall into the lowest SIL (SIL1). Sometimes, this is denoted as “SIL0”.
Therefore, it is important to define a SIL for a function or a piece of equipment. This is done in two steps:
• Define a tolerable hazard rate.
• Use a SIL table to identify the SIL from the tolerable hazard rate. The SIL table indicates an interval for tolerable hazard rates for each SIL.

One note needs to be added on the application of a SIL to components, modules, sub-systems and systems. EN 50129 [4] requires to define the SIL for a system function, not for components, modules or other constituents. Therefore, the SIL may be determined only for a system function. For all other units the SIL is the same as for the system function, the unit belongs to. Tolerable rates of dangerous failures can be further apportioned. Consequently, for each piece of equipment, there is a SIL defining the measures to be used to prevent systematic faults and a tolerable rate for random failures.

1.2 Scope and plan of the paper
The safety integrity level is a main characteristic for safety equipment. It describes requirements for the design and manufacturing process. Definition of the SIL starts with the definition of the “tolerable hazard rate”. The railway safety standards EN 50126 and 50129 refer to three quantitative methods. In this paper, we will restrict ourselves to these methods as they are applied in the railway field. Section 2 gives an overview of these principles. In section 3 a comparison of these methods is carried out and experience is presented. In the fourth section, conclusions are drawn and recommendations are given.
2. DIFFERENT METHODS TO DERIVE TOLERABLE HAZARD RATES

The tolerable rate of dangerous failures can be derived using different principles [2].

Globalement Au Moins Aussi Bon (GAMAB),

“All new guided transport systems must offer a level of risk globally at least as good as the one offered by any equivalent existing system.”

As low as reasonably practicable (ALARP),

“Societal risk has to be examined when there is a possibility of a catastrophe involving a large number of casualties.”

Minimum endogenous mortality (MEM),

“Hazard due to a new system of transport would not significantly augment the figure of the minimum endogenous mortality for an individual.”

These principles will be explained later on.

Having obtained the figure for dangerous failures on demand, a Safety Integrity Level (SIL) is defined according to the following table:

Table 1: Definition of SILs [2]

<table>
<thead>
<tr>
<th>SIL</th>
<th>Probability of Failure per demand (Demand Mode)</th>
<th>Rate of dangerous failures per hour (Continuous Mode)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>&lt; 10⁻⁷</td>
<td>&lt; 10⁻¹⁰</td>
</tr>
<tr>
<td>3</td>
<td>≥ 10⁻⁷ to &lt; 10⁻⁶</td>
<td>≥ 10⁻¹⁰ to 0.3 10⁻⁸</td>
</tr>
<tr>
<td>2</td>
<td>≥ 10⁻⁶ to &lt; 10⁻⁵</td>
<td>≥ 0.3 10⁻⁸ to &lt; 10⁻⁷</td>
</tr>
<tr>
<td>1</td>
<td>≥ 10⁻⁵ to &lt; 10⁻⁴</td>
<td>≥ 10⁻⁷ to 0.3 10⁻⁵</td>
</tr>
</tbody>
</table>

It shall be noted, that the “demand mode” is abandoned in EN 50129 [4]. Moreover, the figures for the continuous mode have been related to the figures from IEC 61508 [1] by a factor of 0.1 in EN 50129 [4].

The table has to be used in the following way. For a rate of dangerous failures, the coinciding class, i.e. the SIL, is searched up in the table, i.e. the table is only used from the right to the left. Then, design measures have to be applied during the design process. The design measures to be applied are also given in the standard. In many cases, these design measures are similar to those given by IEC 61508 [1].

A very sensitive task is the definition of the tolerable rate of dangerous failures.

2.1 Basic understanding

In the sequel, two main terms are explained.

The **tolerable hazard rate** is a rate of dangerous events that can be caused by a piece of equipment. This rate is a target value that has to be met by the equipment. It is derived from a certain principle and is chosen such that the risk arising from the equipment is small.

The **rate of dangerous failures** is the rate of occurrence of failures of a piece of equipment that is achieved by a particular design. So, if the rate of dangerous failures is smaller than the tolerable hazard rate, the piece of equipment can be considered as safe.

2.2 The ALARP principle

The ALARP principle is based on frequency classes and severity classes.

Severity classes can be defined as described in table 2.

The frequency classes are usually defined in steps delimited by a factor of 10. An example is given in table 3. Then, three regions are defined for combinations of severities and frequencies:

I: Intolerable risk, either severity or frequency must be reduced.
T: Tolerable risk, should be reduced. However, risk reduction might be stopped when the costs are too high.
N: Negligible, no action is necessary.

Table 2: Severity classes (example)

<table>
<thead>
<tr>
<th>Safety Class</th>
<th>Failure Consequence</th>
<th>Severity Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insignificant</td>
<td>Minor injuries</td>
<td>IV</td>
</tr>
<tr>
<td>Marginal</td>
<td>Major injuries</td>
<td>III</td>
</tr>
<tr>
<td>Critical</td>
<td>1 fatality</td>
<td>II</td>
</tr>
<tr>
<td>Catastrophic</td>
<td>10 fatalities</td>
<td>I</td>
</tr>
<tr>
<td>Desastrous</td>
<td>100 fatalities and</td>
<td>0</td>
</tr>
<tr>
<td>more</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Frequency Categories (example)

<table>
<thead>
<tr>
<th>Description</th>
<th>Frequency Range (in events per year)</th>
<th>Category Designation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequent</td>
<td>10^-1</td>
<td>A</td>
</tr>
<tr>
<td>Probable</td>
<td>10^-2</td>
<td>B</td>
</tr>
<tr>
<td>Occasional</td>
<td>10^-3</td>
<td>C</td>
</tr>
<tr>
<td>Remote</td>
<td>10^-4</td>
<td>D</td>
</tr>
<tr>
<td>Improbable</td>
<td>10^-5</td>
<td>E</td>
</tr>
<tr>
<td>Incredible</td>
<td>10^-6</td>
<td>F</td>
</tr>
</tbody>
</table>

Table 4: ALARP region (example).

<table>
<thead>
<tr>
<th>Frequency</th>
<th>A</th>
<th>T</th>
<th>I</th>
<th>I</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td></td>
<td>T</td>
<td>T</td>
<td>I</td>
<td>I</td>
</tr>
<tr>
<td>C</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td></td>
<td>I</td>
</tr>
<tr>
<td>D</td>
<td>N</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>IV</td>
<td>III</td>
<td>II</td>
<td>I</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Within the ALARP method, collective risks are considered. Starting from the ALARP region, for each technical subsystem requirements for tolerable hazard rates in the different severity classes are derived. It must be shown that the tolerable hazard rates of all components and sub-systems of the overall system meet the ALARP requirement.

Now, hazard reduction has to take place as long as the system falls into the “T” (tolerable) region or the “I” (intolerable) region. The process may be stopped in the “T” region if the effort of further hazard reduction is too high. Threshold values for sub-systems and functions can be computed partitioning the ALARP region for the functions or sub-systems.

2.3 Minimum Endogenous Mortality

The minimum endogenous mortality is based on an individual risk [6]. Consideration starts at the point of the lowest rate of mortality for human individuals. The rate is minimal for a 15 year old individual and reads 2 \cdot 10^{-4} per year. From the requirement that a technical system shall not contribute more than 5%, it can be derived that a technical system shall not lead to a fatality of a single person at risk with a rate larger than 10^{-5} per year. This figure can then be apportioned further to sub-systems.
2.4 GAMAB
The GAMAB principle can be explained in a very simple way. If $\lambda_{\text{dangerous,old}}$ and $\lambda_{\text{dangerous,new}}$ are the rates of occurrence of dangerous events for the old and new systems, respectively, then the GAMAB principle requires

$$\lambda_{\text{dangerous,new}} \leq \lambda_{\text{dangerous,old}}$$  \hspace{1cm} (1)

Although this inequality is very simple, one aspect shall be noted. Assume a system, as e.g. a train protection system is operated on demand. Then, the rate of dangerous failures is defined as a product of the rate of demands (number of trains over time, $\lambda_{\text{demand}}$) and the probability of failure on demand ($p_{\text{fail}}$), i.e.

$$\lambda_{\text{dangerous}} = \lambda_{\text{demand}} \times p_{\text{fail}}.$$  \hspace{1cm} (2)

3. COMPARISON BETWEEN METHODS TO OBTAIN A TOLERABLE HAZARD RATE
In this section, different methods will be compared that are used to obtain the tolerable hazard rate in railroad technology [2, 4]. It shall be noted that the examples presented below are simplified considerations inspired by real technical projects. Simplification has been used for two reasons:

- prevent disclosure of proprietary information,
- ease the presentation of the material.

3.1 Example for Comparison between ALARP and MEM
This comparison is interesting since the ALARP principle is based on the conception of collective risk, whereas the MEM principle is based on the conception of individual risk.

Assume a hypothetical automatic people mover system controlled by a computer based automatic train protection system (ATP) consisting of 20 modules (track circuits or combinations of them), where each part is controlled by one computer sub-system. For sake of simplicity, failures not arising from the ATP are assumed to lead to 90% of fatalities, so that 10% of all fatalities are caused by the ATP.

3.1.1 ALARP
For a hypothetical automatic people mover system, an ALARP region is assumed as in table 5.

Let us now focus on the train protection system (ATP). For the ATP we take 10% of the figures given in the table above, to leave 90% for other failures or hazards that might cause fatalities. However, failures of the ATP will mostly lead to accidents with about 1..10 fatalities. Accidents with a larger number of fatalities arise from events such as fire. Therefore, only the boundaries of the first two classes with 1..10 fatalities will be taken into account for the ATP. Assume, an ATP failure leads to fatalities with probability of 0.05. Then the limits for the ATP are given in the following table.

<table>
<thead>
<tr>
<th>No. of fatalities</th>
<th>ALARP lower bound in events per year</th>
<th>ALARP upper bound in events per year</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$10^{-4}$</td>
<td>$10^2$</td>
</tr>
<tr>
<td>10</td>
<td>$10^{-6}$</td>
<td>$10^3$</td>
</tr>
<tr>
<td>100</td>
<td>$10^{-7}$</td>
<td>$10^4$</td>
</tr>
<tr>
<td>1000</td>
<td>$10^{-9}$</td>
<td>$10^6$</td>
</tr>
</tbody>
</table>
Table 6: Limits for the ATP

<table>
<thead>
<tr>
<th>No. of fatalities</th>
<th>from ALARP lower bound in events per year</th>
<th>from ALARP upper bound in events per year</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$2 \times 10^{-4}$</td>
<td>$2 \times 10^{-2}$</td>
</tr>
<tr>
<td>10</td>
<td>$2 \times 10^{-6}$</td>
<td>$2 \times 10^{-3}$</td>
</tr>
</tbody>
</table>

Taking the lower bound computed for the ALARP lower bound for 10 fatalities gives a value of $2 \times 10^{-6}$/year. Therefore, the tolerable hazard rate will be about $2 \times 10^{-10}$/hour coinciding with SIL3. The ATP consists of 20 identical modules (e.g. computer sub-systems) controlling 20 parts of the people mover system, such that failure of each of these can lead to an accident. Then the tolerable rate of dangerous failures for a single module (computer sub-system) would be $10^{-11}$/h. In addition to this rate of dangerous (random) failures, measures according to SIL3 would have to be applied for the computer sub-systems. Repeating the computation for the upper bound we arrive at a value of $10^{-8}$/h for a computer sub-system.

### 3.1.2 MEM

Now, the MEM principle will be applied. The natural death rate is $2 \times 10^{-4}$/year (endogenous mortality rate). Allowing a technical system to increase this rate by 5% gives a tolerable rate of $10^{-5}$/year for the people mover. From this rate, 10% will be assigned to ATP failures, leaving 90% for other failures and hazards that might lead to fatalities. This gives a rate of $10^{-6}$/year. Taking into account that ATP failures lead to fatality with probability of 0.05, the tolerable rate for the ATP becomes higher since not every dangerous ATP failure leads to a fatality. Thus, a tolerable rate of $2 \times 10^{-5}$/year for dangerous failures is derived for those parts of the ATP system the individual is exposed to. This is approximately $2 \times 10^{-9}$/h. From this rate SIL3 is obtained. Now, according to the MEM principle, the given individual is not exposed to the entire system at the same time. It is exposed to possibly 2 computer sub-systems of the ATP at the same time (when changing from one area to another) but never to the entire system. Hence, the admissible rate for a computer sub-system is $10^{-9}$/h.

With the two methods we have obtained the results in table 7.

The following observations can be made:
- The rate derived by MEM falls into the interval derived from ALARP.
- The SILs derived are equal for both methods.
- For the ALARP method there is some freedom in choosing the ALARP region. As a result, also a bound can be derived for the equipment. The number of identical computer sub-systems can influence the tolerable rate for one single computer sub-system, but not the SIL.
- For the MEM method, additional factors must be taken as pre-supposition, e.g. the endogenous mortality rate.

Note: For real systems, probabilities as e.g. the probability that ATP failure leads to fatalities might differ.

Table 7: Results for MEM.

<table>
<thead>
<tr>
<th>Method</th>
<th>SIL</th>
<th>Tolerable hazard rate for 1 computer sub-system of ATP</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALARP</td>
<td>SIL3</td>
<td>$10^{-11}...10^{-8}$</td>
</tr>
<tr>
<td>MEM</td>
<td>SIL3</td>
<td>$10^{-9}$</td>
</tr>
</tbody>
</table>

### 3.2 Example for Comparison between GAMAB and MEM

In this example we will study a driver assistance system which gives information on maximal acceptable velocities to the driver. This system is used as a supplementary system to the usual
signal system and does not replace it. Hence, it has no safety functions but assists the driver. Note that, the example resembles the methodology but not the real details for the information system.

### 3.2.1 Use of GAMAB

The use of GAMAB has to compare the old technical system with the new one. We compare the following two situations:

- The driver has to obtain the velocity using two different schedules which are supplied to him on paper.
- The driver reads the data from the display.

We will not assume that the data are wrong in either case. In the first case, wrong data can be obtained only by human error. Since the driver has to carry out a complex task, we assume an error probability of 0.01, see [7]. In the second situation, either the driver reads a wrong value or the electronic system displays a wrong value. The action of the driver is a simple action since only a display has to be red. In this case, an error probability of 0.001 is applicable. The failure probability of the electronic system is denoted by \( p \) (probability of failure on demand). Hence, the overall probability of obtaining a wrong value is 0.001+\( p \).

Now it is required that the new technical solution shall be at least as good as the old one, i.e. the failure probability shall be smaller:

\[
0.001+p < 0.01.
\]

Resolving for \( p \) gives

\[
p < 0.099.
\]

Using the table for definition of the SIL, we obtain a SIL0 which coincides with the result obtained by MEM.

In this case the following observations can be made:

- The computation based on the GAMAB principle is much shorter than that based on the MEM principle.
- The GAMAB principle requires to analyse the existing (old) system.
- The GAMAB method can give good and quick results if the new and the old system are both easy to analyse.

### 3.2.2 MEM

Figure 1 shows the chains of events that can lead to derailment or collision.

From the event tree in figure 1, the rate of dangerous events can be computed as follows:

\[
0.01 \times \lambda \times 1 \times 0.01 \times 0.001 \times 0.01 \times 0.05 + 0.2 \times (0.1 \times \lambda \times 0.05 + 0.9 \times \lambda \times 0.05 \times 0.2) \times 0.0000001 \times (0.01 + 0.001) \times 0.05 = 5.0156 \times 10^{-11} \times \lambda.
\]

Here, \( \lambda \) denotes the rate that the information system gives a wrong value for the permitted velocity. In order to arrive at a requirement for the information system, a target value has to be set up for a dangerous event arising from a failure of the information system.

Starting from the MEM principle, a tolerable hazard rate of \( 10^{-5} \) /year which is nearly \( 10^{-9} \) /hour is derived for a technical system. We assign a fraction of 10% from this value to the signalling system, which is a sub-system of the overall rail system. Hence, we obtain for the signalling system and supporting systems \( 10^{-10} \) /hour. Taking into account that the information system supports the signalling system, we assign a fraction of 1% from the value for the signalling system to the information system. We arrive at a threshold value of \( 10^{-12} \) /hour which shall not be exceeded by dangerous events caused by the information system. Finally, we arrive at the requirement

\[
5.0156 \times 10^{-11} \times \lambda < 10^{-12} /\text{hour}.
\]

Solving for \( \lambda \), we get

\[
\lambda < 0.02 /\text{hour}.
\]

Consequently, the information system has to be assigned SIL0, i.e. weaker requirements than for SIL1 can be applied.
Figure 1: Event Tree for Collision

1

Person inside train 0.05

Collision with obstacle 0.01

Deraillage leading to death 0.001

Dangerously high speed

Failure of ATP 0.0000001

Driver does not realize failure (although known track)

Next signal shows "Stop" 0.001

Velocity is substantially high 0.01

Driver does not realize that value is wrong 1

Information system gives wrong values for permitted velocity

Information system gives wrong values for permitted velocity

Reduction of permitted velocity

Driver does not realize signals 0.05

Reduction of permitted velocity by more than 25 times 0.2

Reduction in normal speed 0.1
3.3 GAMAB and ALARP
From a comparison of ALARP and GAMAB one can see that
- GAMAB is not directly related either to collective or individual risk
- GAMAB might highly depend on rate of dangerous failures of the current system
- in many cases, GAMAB can easily be applied.

4. Conclusions
Three principles for obtaining tolerable hazard rates and, thus Safety Integrity Levels, have been studied. Simple examples are given, where the principles have been compared. Under reasonable assumptions, all three principles give tolerable hazard rates that have the same order of magnitude and hence, yield the same safety integrity level (SIL).
The MEM principle is based on the concept of individual risk, whereas ALARP is based on collective risks. GAMAB refers to neither of those concepts explicitly.
The MEM principle allows to compute directly the tolerable hazard rate for a function or equipment, however, needs assumptions on apportionment of risks, represented by several factors that are used in the computation.
The ALARP concept works with an upper and a lower bound. It is more straightforward in obtaining safety requirements (tolerable rate). However, it needs additional considerations, as long as the rate of dangerous failures of a system function still falls inside the ALARP region, i.e. the rate of dangerous failures is above the lower bound.
GAMAB is easy to apply, since it compares two systems and might compare only small parts. However, the old system needs to be analysed and the results can depend on the rate of dangerous failures of the old system.
Properly applied, all principles can yield comparable results. Due to a certain impreciseness of all three methods, in several cases it might be worthwhile to apply several of them and to compare the results.

References