Does SIL Live up to Expectations?

Roger Short, HonFIRSE, MIET, CEng, Retired

> *"That, Sir, is the good of counting. It brings everything to a certainty, which before floated in the mind indefinitely"*
>
> *Samuel Johnson*

SUMMARY

The relationship between unsafe failure rate and SIL has always been somewhat equivocal. On one hand the relevant standards insist that SIL is concerned with systematic failures, whose occurrence is inherently unquantifiable, especially where software is concerned. On the other hand the standards contain tables which align tolerable functional failure rates with the corresponding SIL to be attributed with regard to systematic failures. This leads to the tacit expectation that if, for example, a system is developed according to the requirements for SIL4, unsafe functional failures due to design errors in software or hardware will manifest themselves at a rate in the region of 10^{-8} to 10^{-9} per hour.

Empirical confirmation of such performance was long regarded as impracticable as it requires evidence of hundreds of thousands of equipment-years of operation. However, there are now in service in railway signalling and train control applications thousands of individual items of equipment claimed to meet SIL4 requirements This paper sets out to determine whether expectations for SIL have been met by estimating the actual unsafe failure rate, including systematic failures, of the current population of SIL4 units, taking account of all the gaps and uncertainties in the relevant data.

1 INTRODUCTION

In the beginning, when the concept of Safety Integrity Level (SIL) was first applied to railway systems, it was a matter of opinion that systems and software developed according to the requirements relating to SIL4 in the standards for electronic systems [2] and software [3] would achieve a level of safeguard against systematic failure consistent with a tolerable hazard rate of 10^{-8} to 10^{-9} per hour resulting from random failures. It was an expert opinion, held by the groups of experts who had developed the standards, endorsed by the railway industry through the public enquiry phase of the standardisation process, and based on the work of the wider industrial and scientific community who had been developing the principles of system safety.

It was a matter of opinion at that time because the amount of evidence from testing and using such systems was far too small to establish it as a fact, and because there were, and still are, no accepted methods of demonstrating by analysis and calculation that the desired safety performance would be achieved (see for example [6] and [18]). For years it was possible to regard this as an abstruse academic problem, since empirical confirmation of such performance would require evidence of hundreds of thousands of equipment-years of operation. However, there are now in service in railway signalling and train control applications tens of thousands of individual items of equipment claimed to meet SIL4 requirements. They include interlockings, radio block centres, axle counters, ETCS on board units and wayside and on board CBTC subsystems.

This paper attempts to evaluate the extent to which this opinion can now be treated as fact, given two to three decades of experience of a steadily growing population of high integrity programmable electronic systems in use on the railways. It begins with a discussion of the extent to which it is legitimate to expect to find evidence an actual failure rate for allegedly unquantifiable systematic failures and goes on to outline the methodology adopted to cope with the limited data available. The publicly available data is analysed, followed by a discussion of the conclusions which can be drawn.

2 SIL AND FAILURE RATE

2.1 Standards and Expectations

The relationship between unsafe failure rate and SIL has always been somewhat equivocal. On one hand the relevant standards, EN50126 [1] EN50128 [3] and EN50129 [2], insist that SIL is concerned with systematic failures, whose occurrence is treated as being inherently unquantifiable, especially where software is concerned. On the other hand the EN50126 and EN50129 contain tables which align tolerable functional failure rates (TFFR) with the corresponding SIL to be attributed with regard to systematic failures, as in Figure 1 below. This leads to the tacit expectation that if, for example, a system is developed according to the requirements for SIL4, unsafe functional failures due to design errors in software or hardware will manifest themselves at a rate in the region of 10^{-8} to 10^{-9} per hour.

What EN50126 says

TFFR [h^{-1}]	SIL attribution	SIL qualitative measures
$10^{-9} \leq$ TFFR $< 10^{-8}$	4	Defined in sector specific standards
$10^{-8} \leq$ TFFR $< 10^{-7}$	3	
$10^{-7} \leq$ TFFR $< 10^{-6}$	2	
$10^{-6} \leq$ TFFR $< 10^{-5}$	1	

What EN50128 implies

TFFR [h^{-1}]	SIL attribution
$10^{-9} \leq$ TFFR $< 10^{-7}$	3/4
$10^{-7} \leq$ TFFR $< 10^{-5}$	1/2

In EN50128 the qualitative measures for SIL3 and 4 are the same, as are the measures for SIL1 and 2.

Figure 1: What the Standards Say

EN50126 says with regard to demonstration of quantified targets "For demonstrating the fulfilment of the quantified targets, only random failures are considered in the evaluation of the TFFR ……. Systematic failures are to be controlled by the adequate process and qualitative requirements stipulated by the SIL qualitative measures and therefore are not contributing to the calculation and quantitative fulfilment of the required safety integrity requirements". But, if in practice systematic failures occur at a rate which significantly exceeds the TFFR then the required safety integrity will not be achieved. Therefore it is not unreasonable of the users of systems developed according to these standards to expect that the achieved rate of unsafe functional failures will match the numerical values stated in the standards.

So far as software is concerned, the relation of expected failure rate to SIL is further complicated by the fact that in EN50128 SIL3 and 4 are allocated the same qualitative measures, as are SIL1 and 2, leading to the implied equivalence shown on the left of Figure 1.

Although the standards as quoted here use the term TFFR, it is clear from the context within which they use this term that they are referring to unsafe failures. In order to minimise any possible confusion, the analysis in the remainder of this paper uses "unsafe functional failure" as a synonym for TFFR.

2.2 Experience not Prediction

This paper does not dissent from the statement quoted above from EN50126 to the effect that SIL measures cannot contribute to the calculation of quantified safety integrity requirements. When the SIL concept was introduced the science of software reliability was not, and still is not, able to predict the probability of errors in software with any degree of accuracy, and the same is true for predicting the probability of design errors leading to systematic failures of hardware. This does not, however, invalidate the concept of determining by means of empirical data the actual failure rate experienced. In this sense it is reasonable to speak in retrospect of the contribution of SIL qualitative measures to the quantitative fulfilment of the required safety integrity requirements.

It is widely recognised, e.g. by Habli [6] and Thomas [18], that the extent to which very low failure rates can be quantitatively demonstrated by means of statistical testing and operational experience is very limited. Shooman [7] has analysed US Federal Aviation Databases and found evidence suggesting the achievement of failure rates

of less than 10^{-7} per hour for software whose failure could contribute to catastrophic failures, but such studies are hard to find.

2.3 Random Failures are Systematic

Although, from the perspective of a design engineer, systematic faults may be non-random in the sense that, *once they are known about*, their effect on the behaviour of the system can be predicted, from the point of view of an operator or maintainer who is not aware of them, their presence is manifested as random failure of the system. The failure of hardware components is subject to the laws of physics: components fail because of combinations of electrical, thermal, mechanical and chemical effects whose results would be entirely predictable if the precise circumstances of the component were known. The apparent randomness of component failures results from lack of knowledge, analogous to the lack of knowledge of design faults before they have been discovered.

"Random" failures are determined by the laws of physics and chemistry, but we don't know when they will happen. "Systematic" failures are determined by design factors, but we don't know when they will happen either (see Figure 2). The difference is in the defences we use.

Thus all failures are systematic from one point of view and random from another point of view. The difference lies in the extent to which the probability of their occurrence can be predicted.

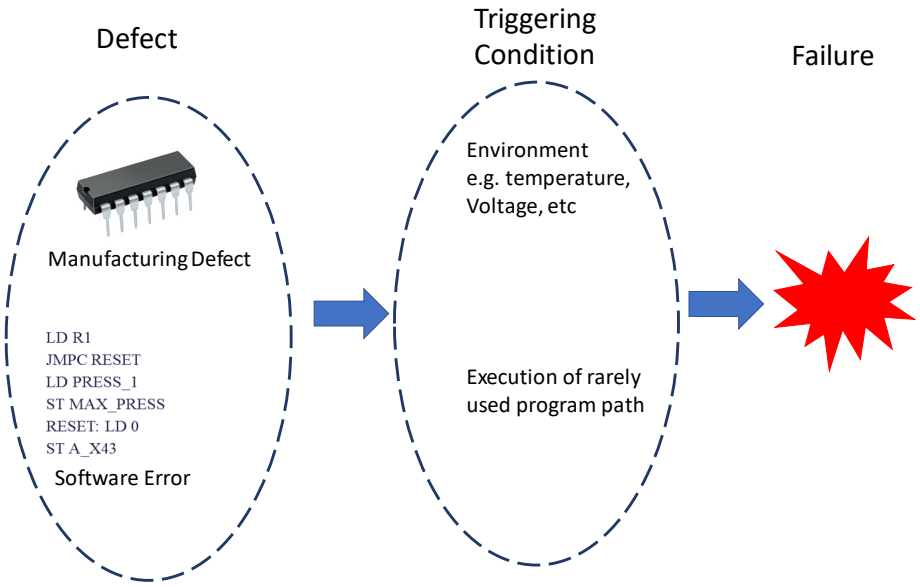

Figure 2: Systematic Errors Cause Random Failures

The failure rates calculated in this paper are intended to encompass both random and systematic failures.

3 METHODOLOGY

3.1 Calculating Failure Rate

The method used here to calculate the rate of unsafe functional failures of signalling systems is simple in concept. If a total of N installations has suffered F unsafe failures in a time period T, then the unsafe functional failure rate λ_{uff} can be calculated using the following expression which is based on reference [4]

$$\lambda_{uff} = F / (NT) \quad (1)$$

This apparently simple calculation is rendered difficult by a number of factors:

- The almost complete absence of data relating to unsafe failures (F). See 3.1 below.
- The limited data concerning the number of installed systems (N). See 4 below.

- The fact that the number of installations (N) will have grown significantly over the long period (T) needed to provide data which is significant in relation to failure rates in the region of 10^{-8} to 10^{-9} per hour. See 4 below.

The scope of this paper is limited to SIL4 systems because they are critical to safety, and because they are more widespread than systems of lower SIL and are easier to identify and enumerate.

3.2 Accidents or Failures?

Railway accidents are rare events which inevitably, and rightly, attract a great deal of public attention. Consequently, if there are no reports of accidents caused by unsafe failure of a SIL4 signalling system then it is possible to be confident that no such accidents have occurred. The same cannot be said for unsafe failures. Unsafe failures which did not result in accidents are not necessarily publicised, so that absence of evidence of unsafe failures cannot be treated as evidence of the absence of unsafe failures.

The approach adopted in this paper is to estimate the accident rate, given zero accidents, as set out in section 3.2 below, and then to determine the unsafe failure rate that this accident rate implies. The implied unsafe failure rate is obtained by analogy with the ratio of other types of failures to railway accidents which can be calculated from publicly available data as explained in section 5.2 below.

3.3 Estimation from Zero-Accident Data

Using equation (1) to estimate the accident rate by setting the value of F to zero given that no accidents have occurred leads to a conclusion that the accident rate is zero and that there will never be any accidents in future. This would clearly be unreasonably optimistic. An approach often adopted in quantitative risk assessment for estimating the probability of an event when no such events have been observed is to proceed as if one event has occurred. The accident rate λ_{acc} would then be given by

$$\lambda_{acc} = 1 / (NT) \qquad (2)$$

This may seem a rather arbitrary approach, but Bailey [5] demonstrates that if the events (accidents or failures) concerned follow the familiar binomial distribution then 1/NT gives a reasonable upper bound on the accident or failure rate. Given how sparse is the data on which the values of N and T are based, a conservative upper bound supporting a conclusion that the actual accident or failure rate is not greater than this would seem to be the most appropriate value to use in support of the argument in this paper.

4 DEPLOYMENT STATISTICS

4.1 Data Sources

The principal source statistics of numbers of units installed has been manufacturers' data sheets, and only a few of these include such information. It has been assumed that for each product considered the number in service has grown linearly at a constant rate from the earliest date mentioned in the source from the publication date of the source material. It has also been assumed that all units in service at the time of publication of the source material have remained in service to the present day, but the growth rate has not been extrapolated and it is assumed that no further units have been deployed. This latter assumption is almost certainly pessimistic.

Some of the interlocking systems considered below pre-date the family of European Standards, [1],[2] and [3], against which safety related electronic systems and software are currently judged. The inclusion of these systems as sources of evidence of the effectiveness of the standards may seem questionable. However, the introduction of these standards did not bring about a radical change in approaches to design, engineering or safety assurance. Far from ushering in a revolution in safety, they merely codified the good practice which had been developed over a period of years. The people who designed the earlier generations of electronic interlockings followed and enhanced this good practice, and some of them were instrumental in the production of the actual European Standards. From this perspective there is justification for considering these systems as having been developed in conformity with these standards *avant la lettre*.

4.2 ERTMS Onboard Units

The most precise and comprehensive publicly available data is provided by the overview of ERTMS deployment statistics on the UNIFE website [11]. This includes a bar chart showing the number of vehicles equipped with ERTMS for each two-year period from 2010 to 2018 inclusive, reaching a total of 12,950 vehicles in 2018. Summing the number of vehicles in each period multiplied by the number of hours in two years gives a gross figure of 900 million equipment hours. However, as railway vehicles have a duty cycle which is significantly less than 24 hours per day, and as the heading of the UNIFE bar chart is "Number of Vehicles Equipped with ERTMS *Contracted*" (Author's italics), one third of this figure has been assumed to be a reasonably conservative value for the contribution of ERTMS onboard units to the global performance of SIL4 systems, i.e. they have contributed $3x10^8$ equipment hours of accident-free operation.

4.3 Signal Interlockings

There are in Europe perhaps a dozen or more types of signal interlocking variously described as electronic, solid state or computer-based but all with a good claim to compliance with the requirements of the relevant European Standards for SIL4. Lists of countries in which these systems have been installed can be readily found on the internet, but in only a few cases has it been possible to find the number of interlockings installed and the period for which they have been in service. Consequently, data for only four interlocking systems is included in this analysis. The author apologises to those manufacturers whose products are not mentioned, and stresses that this is not a reflection on the qualities of their systems but merely a result of the author's failure to find the relevant data.

4.3.1 SSI

Since its first installation in 1985 SSI has become the mainstay of Britain's national network, and it has also been widely installed in other countries. Information obtained from the two suppliers of SSI by the author in 2007 showed that at that time approximately 750 SSI systems were in service. Assuming that production and installation had continued at a uniform rate from 1985, then the number of interlocking-hours which would have accumulated by mid-2007 is given by the formula:

½ x (no. interlockings) x (no. hours in 22 years)

Neglecting additional installations since 2007, for which there is no data, then the total to date is:

{½ x (no. interlockings) x (no. hours in 22 years)} + {(no. interlockings) x (no. hours in 12 years)} (3)

$= 1.5x10^8$ equipment hours

4.3.2 Elektra

According to reference [12] in 2016 there were 310 examples of the Elektra interlocking system in service. The first Elektra system was put into operation in 1989 [13]. Substituting the appropriate values into (3) above gives a total of just over 50 million interlocking hours, so that the contribution of Elektra to the overall estimate is:

$0.5x10^8$ equipment hours

4.3.3 LockTrac L90

According to [14], in 2014 LockTrac 6111 ESTW L90 had been successfully operated for more than 25 years and "systems currently in service switch over 20,000 signals and more than 8,000 points". Assuming an average of 50 signals per interlocking suggests that the population of L90 interlockings grew to 400 in the 25 years leading up to 2014. Substituting the appropriate into expression (3) above makes the contribution of LockTrac L90 to the overall estimate:

$0.6x10^8$ equipment hours

4.3.4 Westrace

Reference [15] states that by 2014 Westrace had already been proven in over 1,200 applications worldwide, while [16] indicates that Westrace dates from the 1990s. For the purposes of this paper it is assumed that the applications to which [15] refers were made over a period of 20 years. Using these values in expression (3) above results in the following contribution of Westrace to the overall estimate:

1.5×10^8 equipment hours

4.4 Systems and Equipment not Included

It has been possible to obtain deployment statistics for only part of the high-integrity signalling systems currently in use. The following brief review indicates the scale of the uncounted systems.

The four interlocking systems for which deployment data was found are not likely to have a greater share of the market than those for which data was not found, considering that the latter include some of the most long-established products.

The Radio Block Centres which for part of the ERTMS system can also be assumed to now make a significant contribution to the accumulated equipment hours of high integrity systems, although probably a smaller contribution to date than that of interlockings.

Axle counters use programmable electronic technology and, although one axle counter evaluator may provide train detection for a number of track sections, where axle counters are used there will be multiple axle counter evaluators for each interlocking. It would be reasonable to believe that the number of equipment hours accumulated by axle counters might be at least an order of magnitude greater than the combined total for interlockings.

Some interlocking systems include remote terminal units or object controllers distributed around the controlled area to provide a vital interface between the central interlocking unit and lineside equipment such as signals, point machines and track circuits. These are themselves high integrity programmable electronic subsystems. The number installed in any particular interlocking area depends on the type of interlocking system and the details of the specific application. For an SSI system these units are known as Trackside Functional Modules, and there can be up to 63 of them for each central interlocking unit. If the equipment hours collectively accumulated by such units were to be included, they might increase the total for interlockings by an order of magnitude.

Statistics for metros would also have made a significant contribution if they had been available. A simple thought experiment suffices to indicate the possible scale: 50 metros with automatic train control provided by CBTC or a similar system, each having 50 trains would require 5,000 on board units, assuming two cabs fitted per train. The possible significance of the contribution from metros can be seen by comparing this with the figures for ERTMS in 4.2 above.

5 ACCIDENT AND FAILURE RATES

5.1 Accident Rate

The sum total of equipment hours obtained in 4.2 and 4.3 above amounts to 7.1×10^8 equipment hours. Using this for the value of NT in equation (2) above gives an accident rate of 1.4×10^{-9} per equipment hour. At first sight this immediately reassuring: the accident rate has been found to be close to the magical figure of 10^{-9} per equipment hour. Given the pessimistic assumption that one accident has occurred when the actual number to date is zero, and given the large number of systems not included due to lack of data, there are strong grounds for confidence that the real accident rate is substantially lower than this. The standards, however, associate SIL with unsafe failures rather than accidents, and it is now necessary to consider what unsafe functional failure rate is implied by this calculated failure.

5.2 Unsafe Functional Failure Rate

5.2.1 Absence of published data

Unsafe functional failures have certainly happened. The author is aware from first-hand accounts of two or three such failures. The expression “two or three” should be taken literally here: there were three incidents, but one occurred during pre-commissioning testing and so from some viewpoints might not be considered a failure. Publicly available data relating to unsafe functional failures of electronic and software based systems is sparse almost to the point of being non-existent. The safety overview report published by the European Union Agency for Railways [9] does include data for signalling wrong side failures in the category “Precursors to Accidents”, showing an average of just under 500 failures per year for the years 2011 to 2015. This presumably covers all types of signalling equipment, and since no details are provided of the types of system/equipment or the causes of the failures it not possible to draw any conclusions about the unsafe failure rates of electronic systems or equipment.

There is no doubt that within the railway industry suppliers and infrastructure managers collect and analyse records of failures in FRACAS or similar systems, but for valid reasons they are not likely to be published. It is also quite possible that such records are to a significant extent incomplete with regard to the types of failure considered here. Anyone who has been involved in the maintenance of electronic equipment will be aware of the problem of intermittent failures which manifest themselves and then disappear without trace until their next appearance. There is no reason to suppose that unsafe functional failures could not exhibit similar behaviour. If a signal were to display a proceed aspect when it was not safe to proceed but there was nobody in a position to see it, then it is not likely that the failure would be detected.

5.2.2 Inferring failure rate from accident rate

Following the occurrence of an unsafe functional failure there are usually a number of factors which might prevent a resulting accident, so that the accident rate is lower than the unsafe functional failure rate. One class of failure which often does not result in an accident, and for which there is ample published data is the signal passed at danger (SPAD).

During the 1980s and 1990s there was much concern in Great Britain over the continuing occurrence of SPADs. This eventually resulted in the decision to install a system of automatic train protection on the whole of the national railway network. The studies which led to this decision included a large amount of statistical analysis which showed, among other things, that over the years 1992 to 2001 there were 4 accidents caused by SPADs (see Table 1 in [10]), while throughout this period SPADs classed as having the potential to cause an accident were occurring at a rate of just over 200 per year (see Figure 2 in [17]). Accidents due to SPADs were thus occurring at a rate of one per 500 SPADs. If the same ratio applies to our one assumed accident, then this implies that there might have been 500 unsafe functional failures over the same period.

In terms of their impact on the safe operation of the railway SPADs and unsafe signalling failures have comparable effects. In the case of a SPAD a train makes a movement without authority while in the case of an unsafe signalling failure a train is given an authority for an unsafe movement. It can thus be argued by analogy that the factors which prevent most SPADs from resulting in an accident, such as there not actually being another train in the conflict zone at the time of the incident, or a train driver observing a conflict and reacting in time to avert an accident, might apply to a similar extent in the case of a signalling unsafe failure.

Figure 3 below uses the “Swiss Cheese” model of accident causation introduced by James Reason [19] to illustrate the relation between failures and accidents. Assuming that the layers of defence represented by the slices are equally effective for SPADs and signalling wrong side failures, the Swiss cheese can be reverse engineered to deduce the unsafe functional failure rate (λ_{uff}) implied by the accident rate (λ_{acc}).

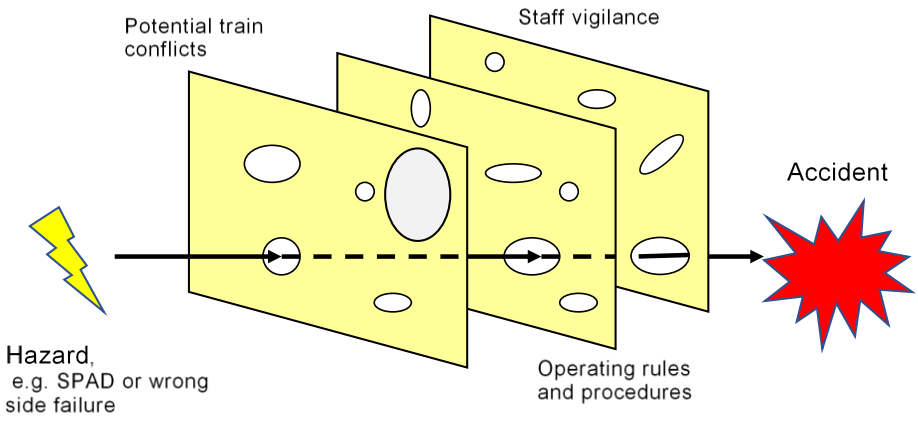

Figure 3: The Swiss Cheese of Signalling Failure

Of course, the SPAD/accident ratio quoted above is based on British signalling and operating practice and might have different values for other railways, but it does seem reasonable to infer that the unsafe failure/accident ratio would be in the region of 100-1000 to 1, i.e.

$$10^3 \geq \lambda_{uff} / \lambda_{acc} \geq 10^2$$

Rounding the calculated accident rate to 1×10^{-9} and applying this ratio gives the value which can be claimed on the basis of the evidence quoted in this paper for the implied unsafe functional failure rate as:

$$10^{-6} \geq \lambda_{uff} \geq 10^{-7} \text{ unsafe functional failures per equipment hour}$$

This is a higher value for the unsafe functional failure rate of SIL4 systems than indicated by the standards, but this is to a significant extent the result of incomplete data on the size of the population of SIL4 systems in service. Section 4.4 above has summarised the types of system which have not been included due to lack of population data. The next section indulges in some speculation about the conclusion which might be reached if that data were available.

6 DISCUSSION AND SPECULATION

The reader with enough patience to have reached this point will by now be experiencing a sense of anti-climax: all these numbers, all this argument, only to conclude that SIL4 has not yet surpassed the intended safety performance of SIL2. It is now time to indulge in a little speculation by cautiously estimating what the result of the calculation might have been if data for some of the systems and equipment described in 4.4 above had been available.

Let us suppose that:

- By including data for all types of interlocking in service, the number of equipment hours due to interlockings is doubled.
- There have been 25 times as many remote terminal units as interlockings in operation over the same lifespan as the interlockings.
- There have been 25 times as many axle counter evaluators as interlockings in operation over the same lifespan as the interlockings.

These seem to be conservative assumptions, but that is only an opinion!

Substituting these values into (2) and including the ERTMS data from 4.2, while using the mid-range value for the failure/accident ratio:

$$\lambda_{uff} = 500 / [2(NT)_{int} + 2(NT)_{int}(25 + 25) + (NT)_{ERTMS}]$$

$$= 1x10^{-8} \text{ unsafe functional failures per equipment hour (rounded to one significant figure)}$$

Which would bring the unsafe functional failure rate achieved to date just to the boundary of the SIL4 band. This is a conjecture which needs to be confirmed by more data.

7 CONCLUSION

The use of programmable electronic systems in railway signalling applications has been growing steadily for more than 30 years, and it seems likely that enough operational experience could have been accumulated by now to demonstrate that the unsafe functional failure rates specified in the standards is being achieved with regard to both random and systematic failures. The available published data was not sufficient to be able to come to a definitive conclusion in this study, but the results hold out the promise that if more data on populations of equipment could be obtained it would be possible to convert expert opinion on the effectiveness of SIL into empirical fact.

8 REFERENCES

1. EN50126: Railway Applications - The Specification and Demonstration of Reliability, Availability, Maintainability and Safety (RAMS).
2. EN50129: Railway applications - Communication, signalling and processing systems - Safety related electronic systems for signalling.
3. EN 50128: Railway applications - Communications, signalling and processing systems - Software for railway control and protection systems.

Note: No dates have been included for references [1] to [3] because each of these standards have been revised within the timespan covered by this paper. In the opinion of the author these revisions do not significantly affect the arguments used in the paper.

4. P. Jalote, B. Murphy, M. Garzia, B. Errez, Measuring Reliability of Software Products, Microsoft Corporation, ISSREE, 2004. https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/tr-2004-144.pdf
5. Bailey, R. Estimation from zero-failure data. Risk Analysis, 1997; 17(3):375–380
6. Habli I, Hawkins R, and Kelly T. Software safety: relating software assurance and software integrity, *Int. J. Critical Computer-Based Systems*, Vol. 1, No. 4, pages 361-383, 2010
7. Shooman, M. 'Avionics software problem occurrence rates', Proceedings of the 7th International Symposium on Software Reliability Engineering, White Plains, NY, pp.53–64, 1996
8. Quigley, J. and Revie, M. Estimating the Probability of Rare Events: Addressing Zero Failure Data, Risk Analysis, Vol. 31, No. 7, 2011.
9. Railway Safety in the European Union – Safety Overview 2017, ISBN 978-92-9205-383-3
10. Evans, A. Fatal Train Accidents on Britain's Main Line Railways: End of 2016 Analysis, Imperial College London, May 2017, https://www.imperial.ac.uk/people/s.conner/document/2987/FTAB2016/?FTAB2016.pdf
11. The European Rail Traffic Management System, UNIFE, http://www.ertms.net/?page_id=58
12. Locktrac 6131 Elektra Electronic Interlocking System, Thales, 2016, https://www.thalesgroup.com/sites/default/files/database/d7/asset/document/locktrac_6131_elektra_en_1.pdf

13. Erb, A. Safety Measures of the Electronic Interlocking System "Elektra", Safecomp '89, Vienna 1989

14. Electronic Interlocking System LockTrac 6111 ESTW L90, Thales 2014, https://www.thalesgroup.com/sites/default/files/database/d7/asset/document/05_p184753_thales_leaflet_locktrac6111_en.pdf

15. Trackguard Westrace Mk II, Siemens AG 2014, https://www.mobility.siemens.com/mobility/global/SiteCollectionDocuments/en/rail-solutions/rail-automation/electronic-interlockings/trackguard-westrace-mk2-en.pdf

16. http://www.railway-strategies.com/2007/11/05/westinghouse/

17. Speirs, F. and Johnson, C. Signals Passed at Danger: A Case Study in the Application of Visualisation Techniques, http://www.dcs.gla.ac.uk/~johnson/papers/Speirs/spad1.pdf

18. Thomas, M. Engineering Judgement, 9th Australian Workshop on Safety Related Programmable Systems, Brisbane 2004.

19. Reason, J. "Beyond the Organisational Accident: the need for Error Wisdom on the Front Line", http://bmjjournals.com

20.