Developing Resilience Signals for the Dutch Railway System

Willy Siegel[1] and Jan Maarten Schraagen[1,2]

[1]University of Twente/GW, P.O. Box 217, 7500 AE Enschede, The Netherlands
A.W.Siegel@UTwente.nl; J.M.C.Schraagen@UTwente.nl

[2]TNO Behavioral and Societal Sciences, P.O. Box 23, 3769 ZG Soesterberg, The Netherlands
jan_maarten.schraagen@tno.nl

Abstract. A resilience state model for a railway system is proposed consisting of three boundaries putting pressure on the operating state: Safety, Performance (Capacity & Punctuality) and Workload. In order to model the pressure of the boundaries, an additional dimension is added where the slope represents the pressure. By doing so, the model is able to differentiate between internal changes that keep the system in a resilient state or have it move towards brittleness. The resilience state model is also used to develop a quantitative signal model, indicating pressure change of the boundaries. A newly defined resilience signal (RS), a quantitative indication of change in system resilience, can be created with help of the signal model and be used for anticipation during operations. The resulting parametric functions will be evaluated and tuned by empirical testing in further research. Using data from governmental reports on responses to incidents, two empirical cases are worked out using the signal model. The first case shows the correlation between a safety RS and safety risk. The second case analyses a capacity RS and explains the results by the system adaptation process through a multi-layer hierarchy.

1 INTRODUCTION

The Dutch rail is in the process of redesigning its mainly technologically driven system which stems from the previous century. Its focus is on handling disruptions and understand, through the four cornerstones of resilience (Hollnagel, 2009), that its largest gap lies in its ability to anticipate. One aspect of anticipation is to enlarge the operational awareness to the resilience state and stimulate the operators to take a pro-active attitude to explore and anticipate on the unknown and unexpected future. However, a quantitative measure of the resilience state of the whole system is currently lacking in the literature. Woods, Schenk & Allen (2009) compared selected models of system resilience including different concepts to explain resilience phenomena. However, these models do not provide quantitative measures of resilience. As a first step towards resilience quantification, we propose to concentrate on resilience signals (RS), which are measured indications that the resilience of the systems is changing. In a redesign of a system these measurements, and their representations, can be taken into account as well as the tooling

to further analyze and possibly anticipate on the RS. Thus, the aim of this paper is to focus on the resilience of a rail system by developing a model from which quantitative RS can be measured. This basis can be used in further research on the quantification of these signals and this model is also valuable to study, analyze and explain specific cases of the past.

2   THEORETICAL RESILIENCE STATE MODEL FOR A RAILWAY SYSTEM

As a start, Rasmussen's (1997) safe operating envelope was used. In Rasmussen's model, three boundaries – performance, economy and workload – are described to explain the different pressures on the Operating State (OS) which may result in crossing one of the borders or readjust them to create a new steady state. In this framework, the performance boundary is directly linked to the 'safety culture' pressure, the economy boundary to the efficiency pressure and the workload boundary to the 'least effort' pressure. In our adaptation of Rasmussen's model, we have introduced some changes to reflect the nature of a railway system. First, we separated 'performance' and 'safety' to reflect their independent nature, while their mutual influence is made explicit in the new model by "upgrading" safety to a boundary entity, which creates safety pressure. Secondly, we moved the economy boundary backwards creating efficiency pressure on the performance boundary, which creates a performance pressure. This change is justified by the fact that in rail systems economic considerations play a more prominent role in the long run strategy and less in daily decisions. However, the performance pressure, created by capacity growth and punctuality to deliver the planned schedule, plays a major role in daily considerations. The workload boundary stays intact reflecting the human importance within a socio-technical rail system. The result of these changes have been depicted in figure 1 - section I.

This model is useful when reasoning about resilience. For example, Cook & Rasmussen (2005) use different areas in the model to explain the stability of a system: unstable, low-risk stable and high-risk stable. The fact that the boundaries put pressure on the Operating Sate (OS) is added textually with the term 'gradient' and grey areas show the OS jump domain, due to shallow gradients. These gradients are of interest, since they represent the internal pressure on the OS and may indirectly be measured and can help explain the resilience of the system when the OS is located at a specific position. When a gradient is large it represents system resilience against external perturbations, while shallowness represents brittleness. This gradient can be made explicit by adding a depth dimension to Rasmussens model as if it is viewed from above in a landscape of valleys as described by Woods et al. (2009), who related the work of Walker & Holling (2004) to that of Rasmussen (1997). The slope (α) of the valley (see figure 1 section II) describes the internal force gradient or Resilience Engineering as called by Walker, acting on the Operating State (OS) while the vector $\vec{d}$ describes the external perturbations on the OS. $d_P = d \cdot Cos\alpha_P$ represents the pressure of boundary B_P. This third dimension with the valley slope is important to understand the amount of resilience when moving towards one of the boundaries. A small slope is an analog to a small hurdle, representing brittleness, to approach the boundary, while a large slope represents resilience. As an example, figure 1 section III depicts an OS that is moving towards the marginal boundary. There are two options to reflect the change of the internal state. When only the capacity of the system is enlarged and no safety measures are taken, it will result in a brittle state, option a, where the marginal boundary is at stake. However, when measures are taken to enlarge the

safety hurdle as well, as in option b, it may result in a deeper valley maintaining the resilience engineered to cope with a higher capacity.

This theoretical model will be used in the next paragraph to model quantifiable resilience signals (RS) through pressure change of the boundaries.
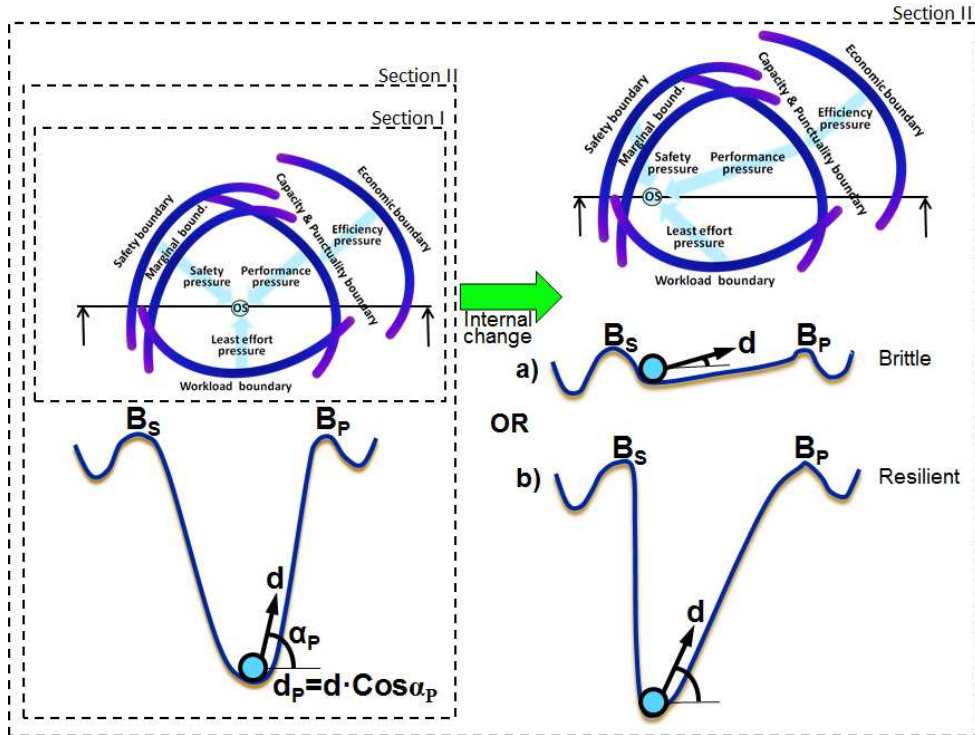

Fig. 1. Resilience state model for a railway system
section I: Rail-sector boundaries putting pressure on the Operating State (OS)
section II: Rail-sector boundaries with resiliency slope α_P, causing pressure d_P
section III: OS move caused by internal change, a or b, influencing system resilience

3   QUANTIFIABLE RAILWAY RESILIENCE SIGNAL (RS) MODEL

The challenge is to translate the above theoretical resilience state model to concrete measurements. Measuring the resilience boundaries with the relative position of the Operating State is a difficult task. The boundaries of a socio-technical system have a subjective character and are based on acceptance and behavior of the community. Cook & Rasmussen (2005) give an example of the marginal creep determined by socio-technical processes and still "only" describe the phenomena, while quantification is not mentioned. As a first step, we propose to measure the pressure of these boundaries and their change over time, which may be considered as a signal triggering further analysis of the situation and possible anticipation, one of the four cornerstones of resilience mentioned by Hollnagel (2009). The focus is on a rail system which has a specific nature and architecture to be used for measuring pressure change due to 1) safety, 2) capacity and 3) workload.

3.1 Safety

Safety plays a role in many aspects in which a safety pressure could be measured. To quantify the concept, we have chosen to measure the safety pressure in the nominal safety sequence of a rail system. A train collision can either occur on the same track or on the crossing of two tracks. We will analyze the safety process in those situations, where the signaling system plays a role, to understand the sequence from which quantitative measurements can be taken to identify a growing pressure. These two nominal situations are depicted in figure 2. In general, rail systems have a technically separated logistic system, using the infrastructure, and a safety system, guarding it independently. According to a logistic plan and the position of trains, the logistic system requests rail path allocations from the safety system. The independent conventional safety system is based on sections that can be occupied by one train at a time and will allocate one rail path to a switch. As depicted in fig 2a, train A occupies section 1, the signal before the section is red, signaling an approaching train to stop. The signal before section 2 is yellow, signaling an approaching train to reduce its velocity and only the signal before section 3 may be green when a rail-path has been allocated for the following train B. The cross-track situation is depicted in figure 2b, where the path for train A, combined through sections 3b, 2b & 1b, allocates switch S1, connecting sections 1b & 1a. Train B is kept on a distance through red and yellow signals before sections 1a & 2a respectively.
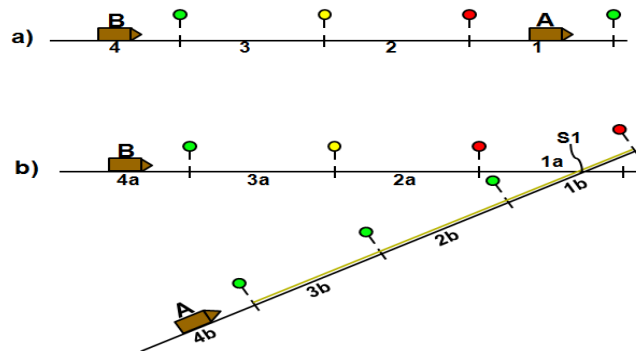

Fig. 2. Safety signalling – one train per rail section and one rail track per switch

When passing a red signal, commonly known as a Signal Passed at Danger or SPAD (Hollywell, 2005), safety is at stake and the more frequently this occurs, the higher the probability of an accident. The number of SPADs may be used to express the amount of safety pressure. However, this can be extended by measuring the number of yellow signal passages and even by the number of red and yellow signal approaches. The latter is justified by following the safety sequence, and the deviation from the optimal safe "green wave". In the "green wave" the train has only green signals, until it gets to its station-stop, and does not need to decelerate until then. With the "green wave" as reference, the pressure of the safety boundary can be expressed as a function of red and yellow signal approaches and passages. These variables can be extended by the number of inhibited rail-paths, due to occupation of switches by another rail path. Thus:

$\alpha_S = f_S$(no. of: SPADs, yellow signal passages, yellow/red signal approaches, switch inhibited rail-paths)

A safety RS can be defined when the slope decrease is larger than a predefined threshold, Threshold-RS_S, indicating that the internal system is becoming less resilient as seen in

fig.1-IIIa. When assuming a monotonic function, the change of the slope $\Delta\alpha_S$ can be estimated by the cumulative weighted changes of the function variables:

$\Delta\alpha_S = K_{1S}(\Delta$ SPADs$) + K_{2S}(\Delta$ yellow-passages$) + K_{3S}(\Delta$ red-approaches$) + K_{4S}(\Delta$ yellow-approaches$) + K_{5S}(\Delta$ switch-inhibits$)$; Safety RS: $\Delta\alpha_S <$ Threshold-RS$_S < 0$

where the weights K_{1S}, K_{2S}, K_{3S}, K_{4S}, K_{5S} and Threshold-RS$_S$ need further empirical investigation.

3.2 Capacity

The change in transport capacity can be directly measured by the actual rail-track usage. In rail reports, for example the report of the Dutch Ministry of Infrastructure and Environment - Inspection of Environment and Transport (2011), the number of kilometers driven on the amount of available track is used as a measurement of capacity utilization. This measurement is relevant but a result of planning activities before operation. During operation, the capacity pressure is largely influenced by delays and infrastructure withdrawal due to malfunctioning and maintenance. The capacity is thus a function of all these parameters:

$\alpha_P = f_P($Driven km's/km track, delays, km infrastructure withdrawal$)$

Similar to the safety RS, a capacity RS can be defined when the slope decrease is larger than a predefined threshold, Threshold-RS$_P$. The change of the slope $\Delta\alpha_P$ can be estimated by the cumulative weighted changes of the function variables:

$\Delta\alpha_P = K_{1P}(\Delta$ driven km's/km track$) + K_{2P}(\Delta$ km infrastructure withdrawal$) + K_{3P}(\Delta$ delays$)$

Performance RS: $\Delta\alpha_P <$ Threshold-RS$_P < 0$

The delay measurement is a complex one still to be tackled in the appropriate context. The weights K_{1P}, K_{2P}, K_{3P} and Threshold-RS$_P$ need further empirical investigation as well.

3.3 Workload

The main driver for managing the rail operation is the prepared logistic plan. When no deviation occurs from this plan, the system can practically be run automatically and causes work under-load. When deviations occur, the plan is updated or direct commands to the infrastructure are executed. The workload can be estimated through these measurable system activities. Neerincx (2003) proposed a model to measure the workload during operations, as a function of three variables: task switches, task time duration and task complexity. Each of the above system activities could be seen as part of a task sequence providing the possibility to count the task switches and task duration. Complexity in this model is defined by Rasmussens SRK-levels (Rasmussen, 1986). Each change of the plan needs reasoning and is at least rule-based but mostly knowledge based. This modeling makes it possible to estimate the workload with indirect measurements and can be expressed as follows:

$\alpha_W = f_W($plan adaptations, direct actions on the infra$)$;

$\Delta\alpha_W = K_{1W}(\Delta$ plan adaptations$) + K_{2W}(\Delta$ direct actions on the infra$)$;

Workload RS: $\Delta\alpha_W <$ Threshold-RS$_W < 0$

The workload change can be estimated by the change of plan adaptations and the number of actions on the infrastructure. The workload RS can be defined when the slope decrease is larger than a predefined threshold, Threshold-RS$_W$. The weighted relations K_{1W}, K_{2W} and Threshold-RS$_W$ need to be worked out empirically.

4   TWO EMPIRICAL CASES USING THE RESILIENCE SIGNAL (RS) MODEL

The RS model described above needs to be verified and tuned according to empirical testing within the railway *operations* itself, which will provide the needed detailed information. However, in the public domain information may be gleaned from reports describing situations on a national level and on a yearly basis. The report from the Dutch Ministry of Infrastructure and Environment - Inspection of Environment and Transport (2011) on red signal passage, chosen for its relevance to the safety boundary, and the report from the Dutch Competition Authority (2010) on the rail capacity, chosen for its relevance to the performance boundary, were analyzed and provided two cases with respect to the RS model.

4.1   Correlation between a safety RS and safety risk

One of the assumptions of the safety signal model, worked out in the previous paragraph, is that a growing number of SPADs indicates a growing safety pressure. This assumption, among others, needs to be proven empirically, since a SPAD does not always create a high risk situation. For example, if no train is in the block behind a red signal the probability of a collision is very low. The aspect of differentiating between a SPAD with a high or low risk has been worked out by the Dutch Ministry of Infrastructure and Environment (2011) that tracks and reports yearly on the status of SPADs. It uses for each SPAD the SPAD Risk ranking methodology, a standard of the RSSB (Rail safety and standards board), which takes into account among others the relative positions, velocities, infra setting, etc. to calculate the risk of a serious accident. These figures, the number of yearly SPADs with a high risk, have been extracted from the report as well as the total number of yearly SPADs, and plotted against each other. The result for the period 2007-2011 shows that the two variables are highly correlated (Pearson r = 0,98), justifying the intuitive safety RS model assumption that a growing number of SPADs indeed signals a growing pressure on the safety boundary. The other assumptions that red signal stopping, Yellow signal passage and crossing rail-paths are safety signals as well, still need to be tested empirically.

4.2   Analysis of a capacity RS and a safety RS

Does a resilience signal (RS) always imply that the resilience of the system is degraded? This is in theory not the case, since a signal implies a growing pressure on the boundaries but the total impact on the whole system still needs to be analyzed. As an empirical example, we have taken the capacity RS reported by the Dutch Competition Authority (Nederlandse Mededingingsauthoriteit - NMa) (2010), stating a yearly capacity grow of the Dutch rail infrastructure utilization, in the period between 2005 and 2009. This situation could be described by the resilience state model in figure 1-III, where the operating state is moving towards the safety boundary, due to a capacity growing pressure. This may lead towards a brittle situation when also a growing pressure on the safety boundary is seen, as in option a, or it may lead to sustain the resilience of the system, as in option b, where appropriate internal measurements are taken. To draw a conclusion, additional data have been used from Dutch Ministry of Infrastructure and Environment report (2011) where SPADs in the period between 2007 and 2011 have been recorded. The two sets of data, in the overlapping period between 2007 and 2009, have been plotted against each other. The result shows clearly that while the capacity utilization is increasing, the SPADs are decreasing, meaning that the internal system has organized itself in a resilient manner as

shown in option 3b. This surprising result invites further analysis, which can be extracted from the Infrastructure and Environment report (2011) as well.

Already in the 1990s, the system logged a growing number of SPADs. This triggered the Railned company to write a report on the status of the rail system (Gotz, 2002). This report was presented to the parliament (Peijs, 2004). It was only after the collision in Amsterdam on May 21 2004 that different rail groups were triggered to set up a parliamentary steering committee on the subject of SPADs. This committee defined the following targets:

50% reduction of the 2003 SPADs to be achieved in 2009

75% reduction of 2003 SPAD risks to be achieved in 2009

These targets were adopted by the Minister and presented to parliament. Accordingly, the steering committee set up a program to achieve the targets divided into 4 parts (Rail branch steering committee SPADs, 2009 - the year of the Barendrecht accident on September 24):

1) A program for Train-drivers; 2) Automatic system for influencing trains (ATB) revised version; 3) Emplacement analysis; 4) Setup regulations

The above case shows the adaptive capability of the socio-technical railway system over a period of many years and accidents. This process can be described by a multi-level hierarchy (Rasmussen, 1997) depicted in figure 3.
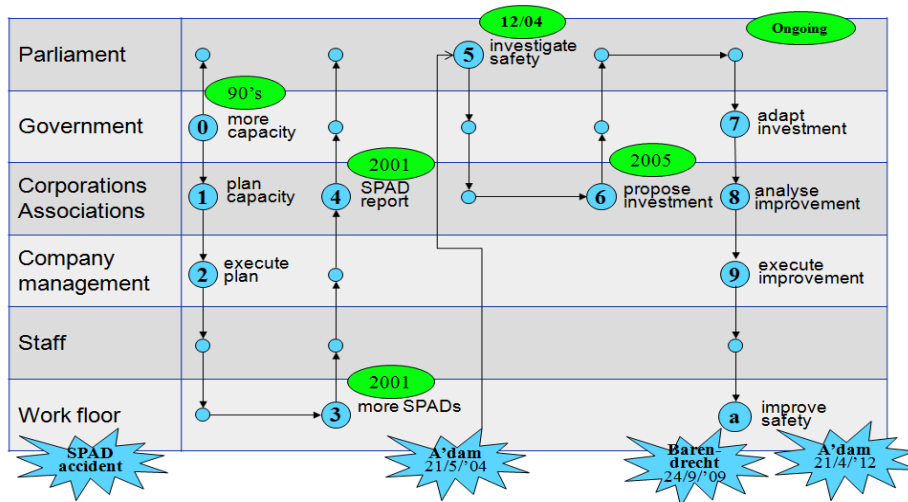

Fig. 3. Multi-level hierarchy explaining the adaptation process

Although the main trigger to come into action were the accidents, the initial trigger was the growing number of SPADs - a safety RS.

5   CONCLUSIONS AND NEXT RESEARCH STEPS

Quantification of resilience system-state attributes is important to enhance operational awareness and stimulate rail traffic operators to take an attitude to explore and anticipate the unknown and unexpected future. We have chosen to explore resilience signaling rather than the resilience boundaries themselves since these boundaries are not tangible

because, on the one hand, resilience is about regions beyond the standard behavior of the system and, on the other hand, the boundaries are uncertain (Cook & Rasmussen, 2005) and constantly moving, due to the socio-technical nature of the system. Signals are by their nature not solid but give a clue on possible events that may occur.

In this paper, we have developed resilience signals (RS) with a focus on the Dutch railway system. Most probably these results are applicable to other railway systems and to other semi-governmental transport systems. These expectations can only be verified after the modeling has been adjusted according to further empirical investigation within the Dutch rail environment, where more operational parameters will be imposed maturing the results for usage in real-time operations.

The empirical example in this article has a limited significance, due to data of a short three year period with coarse granularity in yearly and national units, but is a good case showing that resilience signals (RS) are by nature not strong signals and need further investigation, to draw correct conclusions. This is exactly the role of the system operators at the anticipation cornerstone.

REFERENCES

Cook, R., & Rasmussen, J. (2005). "Going solid": a model of system dynamics and consequences for patient safety. *Quality & safety in health care*, *14*(2), 130–4. doi:10.1136/qshc.2003.009530

Dutch Competition Authority (Nederlandse Mededingingsauthoriteit - NMa). (2010). *Vijfde Spoormonitor NMa*. Retrieved from
 http://www.nma.nl/images/Vijfde_Spoormonitor_NMa_origineel22-157761.pdf

Dutch Ministry of Infrastructure and Environment - Inspection of Environment and Transport. (2011). *STS-passages 2011 analyse 2007-2011*. doi:ILT-2012/14980

Gotz, W. W. J. (2002). *Trends in veiligheid van het spoorwegsysteem in nederland* (pp. 1–38). doi:RnV/02/M10.008.076

Hollnagel, E. (2009). The four cornerstones of resilience engineering. In C. P. Nemeth, E. Hollnagel, & S. Dekker (Eds.), *Resilience Engineering Perspectives. Vol.2, Preparation and restoration* (pp. 117–134). Ashgate studies in resilience engineering.

Hollywell, P. (2005). ORAM: a structured method for integrating human factors into SPAD risk assessment. In J. Wilson, B. Norris, T. Clarke, & A. Millas (Eds.), *Rail human factors* (pp. 189–200). Ashgate.

Neerincx, M. A. (2003). Cognitive task load analysis: allocating tasks and designing support. In E. Hollnagel (Ed.), *Handbook of cognitive task design* (pp. 283–305). Mahwah, NJ: Lawrence Erlbaum Associates.

Peijs, K. M. H. (2004). *Veiligheid en het railvervoer* (pp. 1–5). doi:Report code: 29893, ISSN 0921-7371

Rasmussen, J. (1986). *Information Processing and Human-Machine Interaction. An Approach to Cognitive Engineering*. North-Holland. Retrieved citeulike-article-id:794973

Rasmussen, J. (1997). Risk management in a dynamic society: a modelling problem. *Safety Science*, *27*(2/3), 183–213.

Walker, B., & Holling, C. (2004). Resilience, adaptability and transformability in social-ecological systems. *Ecology and Society*, *9*(2). Retrieved from http://profesores.usfq.edu.ec/fdelgado/Ecologia Humana/articulosdigitales/Walker.pdf

Woods, D., Schenk, J., & Allen, T. (2009). An initial comparison of selected models of system resilience. *Resilience engineering perspectives* (pp. 73–94).