



# Graceful extensibility in asset management: extending the capacity to adapt in managing cyber-physical railway systems

Jan-jaap Moerman<sup>1</sup> · Jan Maarten Schraagen<sup>1,2</sup> · Jan Braaksma<sup>1</sup> · Leo van Dongen<sup>1</sup>

Received: 30 November 2019 / Accepted: 7 January 2021 / Published online: 22 March 2021  
© The Author(s) 2021

## Abstract

Graceful extensibility has been recently introduced and can be defined as the ability of a system to extend its capacity to adapt when surprise events challenge its boundaries. It provides basic rules that govern adaptive systems. Railway transportation systems can be considered cyber-physical systems that comprise interacting digital, analog, physical, and human components engineered for safe and reliable railway transport. This enables autonomous driving, new functionalities to achieve higher capacity, greater safety, and real-time health monitoring. New rolling stock introductions require continuous adaptations to meet the challenges of these complex railway systems as an introduction takes several years to complete and deals with changing stakeholder demands, new technologies, and technical constraints which cannot be fully predicted in advance. To sustain adaptability when introducing new rolling stock, the theory of graceful extensibility might be valuable but needs further empirical testing to be useful in the field. This study contributes by assessing the proto-theorems of graceful extensibility in a case study in the railway industry by means of adopting pattern-matching analysis. The results of this study indicate that the majority of theoretical patterns postulated by the theory are corroborated by the data. Guidelines are proposed for further operationalization of the theory in the field. Furthermore, case results indicate the need to adopt management approaches that accept indeterminism as a complement to the prevailing deterministic perspective, to sustain adaptability and deal effectively with surprise events. As such, this study may serve other critical asset introductions dealing with cyber-physical systems in their push for sustained adaptability.

**Keywords** Graceful extensibility · Sustained adaptability · Asset management · Cyber physical systems

## 1 Introduction

Rolling stock introductions deal with complex railway systems that comprise interacting digital, analog, physical, and human components engineered for safe and reliable railway transport. New rolling stock is characterized by an increasing convergence of information technologies and operational technologies, also referred to as ‘next generation trains’. This enables autonomous driving, new functionalities to achieve higher capacity, greater safety, and real-time health monitoring. The introduction of new rolling stock in an already complex railway system is a big challenge for railway operators as it involves many different business units and organizations

in different stages of the introduction. An introduction takes several years to complete and has to deal with political influences, changing stakeholder demands, new technologies, and technical constraints which cannot be fully predicted in advance. Unfortunately, there is no single solution to overcome these challenges. Therefore, one can expect surprise events and, if not managed properly, fragile railway systems.

The theory of graceful extensibility has recently been introduced as the opposite of brittleness and can be defined as the ability of a system to extend its capacity to adapt when surprise events challenge its boundaries (Woods 2015). It provides a set of basic rules that govern adaptive systems. Its ideas and concepts have been introduced by Woods (2018) as proto-theorems, but, as suggested by Woods (2018), need further empirical testing. This study is a first attempt to assess this new theory and its usefulness in coping with complex cyber-physical systems. Its contribution lies in exploring the explanatory power of the proto-theorems of graceful extensibility in an in-depth historical case study into

✉ Jan-jaap Moerman  
j.moerman@utwente.nl

<sup>1</sup> University of Twente, Enschede, The Netherlands

<sup>2</sup> TNO, The Hague, The Netherlands

a railway rolling stock introduction using pattern-matching analysis (Trochim 1989). Pattern matching analysis involves the specification of a theoretical pattern, the acquisition of an observed pattern, and an attempt to match these two (Trochim 1989). Rolling stock introductions can be considered as the introduction of complex cyber-physical systems, which take, on average, 5 years to complete. By selecting the Fyra V250 case (V250), which has already been subject to several evaluations and reflections [see, for example, Silfhout and Berg (2014)], the authors attempt to identify patterns that may have resulted in failed sustained adaptability, but can provide practical guidance to future rolling stock and other critical asset introductions. The main focus of this study is on human factors and the decision-making processes on different organizational levels within and between organizations involved in the introduction process.

The remainder of this paper is structured as follows. Section 2 briefly introduces the theory of graceful extensibility and connects it to current Industry 4.0 challenges. Section 3 explains the research approach and summarizes the pattern-matching analysis technique. After the case introduction in Sect. 4, case results are presented in Sect. 5. This section concludes by stating the (mis)matches of the patterns of graceful extensibility and provides possible explanations, supported by relevant literature. As this study is part of a research project aimed at increasing reliability in rolling stock introductions, Sect. 6 includes the main conclusions and possible future research directions.

### 1.1 Dealing with surprise events in critical asset introductions by means of graceful extensibility

Industry 4.0 is currently a much-discussed topic that has the potential to affect entire industries by transforming the way goods are designed, manufactured, delivered, and paid for. The rapid adoption and application of pervasive digital technologies in several industries not only radically changes products and services, but also fundamentally reshapes organizations (Yoo et al. 2012). Hermann et al. (2016) identified four industry 4.0 components based on their review of academic and business publications: the concepts of cyber-physical systems (Akanmu and Anumba 2015), the Internet of things (Porter and Heppelmann 2014, p. 4), and the Internet of services (Andersson and Mattsson 2015) are closely linked. These concepts enable the so-called ‘smart factory’, which is based on the idea of a decentralized production system, in which “human beings, machines and resources communicate with each other as naturally as in a social network” (Kagermann et al. 2013, p. 19). However, as more heterogeneous modules, originally produced by diverse actors, are combined to create innovations, organizations increasingly run the risk of complex systemic failure or other forms of unintended consequences (Perrow 1984). This is

also reflected in the observation made by Baheti and Gill (2011) who state that the diversity of models and formalisms in the development of cyber-physical systems at the component level poses a serious problem for verifying the overall correctness and safety of designs at the system level. Therefore, organizations should look for ways to deal with an increase in surprise events.

Projects dealing with complex systems, such as the introduction of new rolling stock, have certain characteristics that require consideration to be managed successfully. Understanding and dealing with surprise events and the unknown are a major challenge in project management. For example, Ramasesh and Browning (2014) present a conceptual framework for dealing with unknowns in project management. These unknowns could be foreseen but for various reasons (e.g., barriers to cognition) are not. Furthermore, in managing unforeseen events, Saunders et al. (2016) observed high reliability practices in their study [see, for example, Weick et al. (2008)] into safety-critical projects. However, these practices are often fragile in nature and dependent on key individuals. The concept of system resilience is another approach in dealing with complex systems. Four lines of inquiry were identified to capture different senses of resilience and reducing risks of sudden failures in complex systems (Woods 2015): rebound, robustness, graceful extensibility, and architectures for sustained adaptability. Previous research has shown that the effort invested to improve fitness, leads to systems that are robust to stressors they were designed to handle, yet fragile to unexpected events and design errors (Carlson and Doyle 2000, p. 2529). While improving the system regarding certain criteria, the same improvements produce severe brittleness when surprise events occur. Brittleness is defined as the rapidity of a system’s performance decline when it nears or reaches one or more boundary. Brittle systems experience rapid performance collapses, or failures, when events challenge boundaries (Woods 2015). The opposite of brittleness is Graceful Extensibility (GE), or how to extend adaptive capacity in the face of surprise events (Woods and Branlat 2011). In accordance with Woods (2018, p. 6), a surprise is here defined as: “Given bounds on adaptive capacity, there are events which will occur that fall near and outside the boundaries; thus, surprise is model surprise where base adaptive capacity represents a partial model of fitness”.

The theory of GE explains the contrast between successful and unsuccessful cases of sustained adaptability. Sustained adaptability refers to the ability to continue to adapt to changing environments, stakeholders, demands, contexts, and constraints (Woods 2018). The theory of GE is strongly linked to concepts in control systems. Control systems are in many ways a simple form of adaptive system, and theory specifies how to ensure stability (adequate adaptive performance) given well-defined targets and well-modeled

disturbances. Graceful extensibility also is a play on the concept of software extensibility from software engineering. Software engineering emphasizes the need to design, in advance, properties that support the ability to extend capabilities later, without requiring major revisions to the basic architecture, as conditions, contexts, uses, risks, goals, and relationships change (Woods 2018).

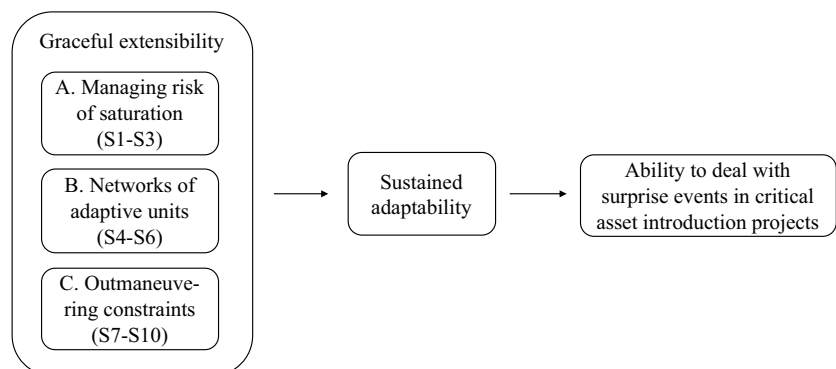
Graceful extensibility is defined as the ability of a system to extend its capacity to adapt when surprise events challenge its boundaries (Woods 2018). See, for example, Wears et al. (2008) for how medical emergency rooms adapt to changing, and high, patient loads. At the heart of the theory of GE lies the fundamental concept of managing risk of saturation via regulating the Capacity for Maneuver (CfM), both at the level of an adaptive unit and at the level of a network where neighboring adaptive units interact as risk of saturation increases (Woods 2018). In prior research, Woods and Branlat (2011) identified three basic patterns of maladaptation. The three basic patterns are decompensation (lack of capacity to adapt when disturbances cascade), working at cross-purposes [local versus global (mal) adaptive behavior] and getting stuck in outdated behaviors when relying on past successes. The theory of GE is presented as 10 proto-theorem (S1–S10) divided into three subsets (Fig. 1) that express the fundamentals that govern adaptive systems. Proto-theorems in subset A (S1–S3) capture how the CfM is regulated to manage and reduce the risk of saturation. Subset B (S4–S6) addresses what is required in order for a layered network to sustain adaptability. It captures several basic processes which influence how adaptive units will act when a neighbor is at risk of saturation and whether units will act in ways that extend or constrict the CfM of the unit at risk. Subset C (S7–S10) captures how constraints, such as perspective bounds and mis-calibration of adaptive capacity, can be addressed. Expanding on the work on GE done by Woods (Woods 2018, p. 23), this study proposes to explore the explanatory potential applying pattern-matching analysis to the 10 statements of GE.

## 1.2 Research approach

This study explores the explanatory potential of graceful extensibility using pattern-matching analysis. A systematic research design was adhered to in a single confirmatory descriptive case study (Yin 2003). Pattern matching analysis (Trochim 1989) at the very least involves the specification of a theoretical pattern, the acquisition of an observed pattern, and an attempt to match these two. What matters are the patterns of the outcomes, not the outcomes themselves. Trochim (1989, p. 357) describes pattern-matching techniques as distinct from the traditional hypothesis testing in that “pattern matching encourages the use of more complex or detailed hypotheses and treat(s) observations from a multivariate rather than a univariate perspective”. In case-study research, pattern-matching techniques are designed to enhance the rigor of the study; if the empirically found patterns match the predicted ones, the findings can contribute to, and strengthen the internal validity of the study, and can result in the confirmation of the propositions (Yin 2003). Furthermore, Yin (2003) emphasizes that, irrespective of design, data analysis using pattern matching is entirely appropriate for all case-study designs if its use is consistent with the purpose of the study and the research questions to be answered. Since qualitative research often lacks precision, an important suggestion is to avoid postulating very subtle patterns, so that pattern-matching results deal with gross matches or mismatches whose interpretation is less likely to be challenged (Yin 2003). Several factors need to be considered in the research design when using pattern-matching analysis as described by Trochim (1989, p. 357). These factors are: conceptualization of the theory, the level of generalization, the value of reanalysing historical data, treating relevant data as a whole rather than a collection of individual outcomes, and the procedures required to provide evidence for a match.

The V250 case, which will be further introduced in Sect. 4, was selected as an example of failed sustained adaptability. Train services were canceled after 2 months in operation, after the introduction had already been delayed

**Fig. 1** Connecting graceful extensibility (Woods 2018) to sustained adaptability in in critical asset introductions dealing with surprise events



for several years. The case includes many data sources since a parliamentary inquiry was also part of the evaluation of the V250 (Parliamentary Inquiry Committee Fyra 2015). Consistent with the approach developed by Yin (2003), each data source was initially collected and analyzed resulting in 430 coded items (refer to Fig. 2). Sources (videos, published and unpublished reports, internal memos) were coded using the qualitative research software of Atlas.ti. The first step in pattern matching is developing a proposition prior to undertaking the study (Trochim 1989). A theoretical pattern is a hypothesis about what is expected in the data. The observed pattern consists of the data that is used to examine the theoretical model. To the extent that patterns match, one can conclude that the theory predicts the observed pattern and receives support. Based on the proposition which will be introduced in Sect. 5, focused coding resulted in 176 items. Data were analyzed on an organizational level (level of generalization) and theoretical patterns were selected in advance. The conceptualisation of the theoretical patterns was proposed by Woods (2018) and adopted in this study in the pattern-matching process.

Following Fig. 2, based on the items arrived at by means of focused coding (176 items), empirical patterns were constructed from the case findings to compare it to the theoretical patterns as defined by the theory on GE. The sub-statements within each of the 10 proto-theorems were categorized as sub-patterns and compared to the empirical patterns identified in the case. In instances where patterns did not match, alternative explanations were explored and discussed with former key players in the V250 introduction. Part of the pattern-matching analysis (central in Fig. 2) has been included in the Appendices, so that readers have the opportunity to compare results for themselves. Furthermore, Sect. 5.2 includes a detailed example of the matching process in the case as presented in Fig. 2.

### 1.3 Case introduction

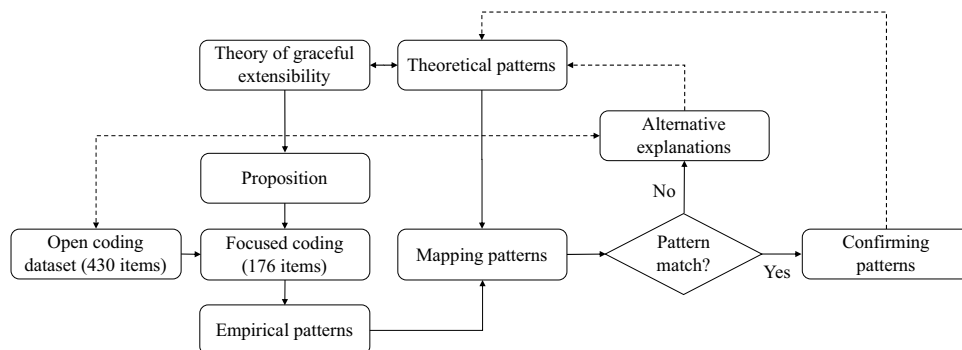
To gain an understanding of the V250 introduction and its context, this section summarizes the timeline of the introduction and highlights the context in which surprise events

occurred. As summarized by Johns (2006), an understanding of a context contributes to an understanding of the entities embedded within that context. It affects the cognition, affect, and behavior of individuals embedded within it. Context influences processes and interrelationships between constructs, as well as the meaning that people ascribe to events or themselves.

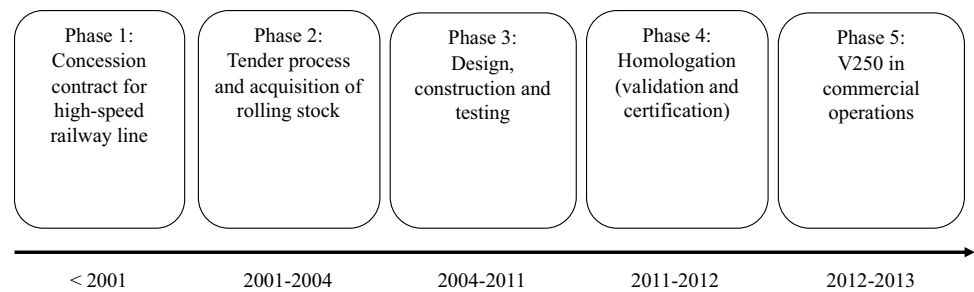
The main objective of the V250 introduction was introducing a high-speed train on the high-speed railway networks HSL Zuid (Netherlands) and Line 4 (Belgium) to connect to existing high-speed railway networks in Europe. The introduction of the V250 was characterized by the introduction of new digital train systems (e.g., the European Railway Traffic Management System, ERTMS), which needed to be integrated with other mechanical systems. This has a great impact on the behavior of the joint human–machine system (Hollnagel and Cacciabue 1999) in meeting the demands from the environment and maintain control. The main stakeholders in the introduction consisted of private and public entities which included the railway operators, the suppliers, the governments of Belgium and the Netherlands, the authorizing bodies (supervisors), the Designated Body, the Notified Body, the infrastructure managers, and the maintenance supplier. This is not a complete list, but serves as an indication of the large number of stakeholders and their interest in the V250 introduction. Figure 3 indicates the timeline of the introduction, with a lead time of over 11 years.

- Phase 1: Concession contract for high-speed railway line. In 1996, in the context of the liberalization of the European Railway industry, the Dutch government ‘privatized’ the main railway operator in the Netherlands (Nederlandse Spoorwegen, NS), but remained its sole shareholder. Furthermore, train and track systems were legally separated. This emergent market orientation has led, among other things, to a strongly legalistic approach to the construction of the HSL Zuid and the acquisition of the V250 trains. Furthermore, requirements imposed by the government to implement a new (unproved) European safety system (ERTMS) in a cross-border high-

**Fig. 2** Data analysis and pattern-matching process V250 case-study design based on Trochim (1989)



**Fig. 3** Timeline v250 introduction (Parliamentary Inquiry Committee Fyra 2015)



speed infrastructure network (HSL Zuid and Line 4), using innovative high-speed trains, increased complexity.

- Phase 2: Tender process and acquisition of rolling stock. The imposed requirements for rolling stock limited the scope for maneuver in the highly regulated tendering process. Due to the small number of trains and the high development costs per train, only one candidate contractor remained. The Dutch and Belgian railway operators signed a turnkey contract with this supplier. A turnkey contract is one under which the contractor is responsible for both the design and construction of rolling stock, ready for commercial use at the agreed price and by a fixed date. The main reason for a turnkey approach in the purchasing agreement was to outsource risks as both railway operators had little experience in designing and constructing high-speed rolling stock. However, this restricted the opportunity to monitor (and influence) the design, construction, and testing processes, which prevented an early anticipation of issues regarding maintaining and operating the V250 trains.
- Phase 3: Design, construction, and testing. An overly optimistic estimate of the delivery times beforehand resulted in unrealistic planning in all phases. Detailed timetables of deliverables by the contractor were lacking, and as a result, assumptions were made. Eventually, this resulted in a delay of five years in the delivery of products and services by the contractor (Parliamentary Inquiry Committee Fyra 2015, p. 5). Furthermore, testing was delayed due to a lack of clear (testing) requirements for ERTMS and a late delivery of the infrastructure of HSL Zuid. As the ERTMS system was in its early development phase, updates were prescribed each time. This led to a great deal of uncertainty and delay in solving technical problems to establish a working and certified security system for the HSL Zuid.
- Phase 4: Homologation (validation and certification). The process of homologation took place during the construction of the V250. The process was complex, because homologation of the train had to take place both in the Netherlands and in Belgium. Additionally, the European Technical Specifications for Interoperability (TSI) also had to be considered. The authorized body of the Nether-

lands did not inspect any physical trains and relied solely on the findings of the Notified Body which also did not inspect all trains (Parliamentary Inquiry Committee Fyra 2015). Furthermore, the terms and conditions under which rolling stock was transferred from contractor to railway operator were unclear due to different interpretations of the purchase agreement.

- Phase 5: V250 in commercial operations. Commercial operations of the V250 trains started on December 9, 2012 between Amsterdam and Brussels. On January 17, 2013, all V250 trains were removed from service due to an incident in which one of the V250 trains had lost a base plate due to ice formation and the continuing incidents with other V250 trains. The lack of timely communication of the introduction of a new train service to passengers and at the same time the cancellation of the existing Benelux line resulted in public outrage and high political pressure. Since unknown technical problems are one of the key characteristics of new rolling stock, reliable performance can never be guaranteed beforehand and a fallback scenario needs to be prepared. The so-called ‘teething problems’ can occur as a result of unexpected defects in the system in commercial operations due to technical, organizational, or human failures. As people, trains, and infrastructure are locally distinctive, testing or simulation may never prevent these (introduction) challenges completely.

#### 1.4 Case results

This section presents the results of the case study using pattern-matching analysis. Two assumptions from the theory of graceful extensibility state that resources are always finite and change is ongoing. As a result, both risk and uncertainty are always present (Woods 2018). This requires Units of Adaptive Behavior (UABs) at multiple nested scales (e.g., processes, individuals, organizations, teams, and networks). The pattern-matching analysis in this study was performed on an organizational level. The unit of analysis was the V250 introduction, consisting of several UABs (operator, infrastructure manager, suppliers etc.) with different accountabilities and responsibilities,

but all with the same final objective, safe, and reliable passenger railway transport. As defined in Sect. 3, surprise events are those events that fall near and/or outside the boundaries of the adaptive capacity of a system (Woods 2018). Figure 4 illustrates the operationalization of surprise events in the context of the V250 introduction. Surprise events which fall near the boundaries of the adaptive capacity of a UAB occur (Fig. 4: X). Other surprise events which fall outside the boundaries of a UAB occur and require extended adaptive capacity from that same UAB (Fig. 4: Y). Further surprise events may occur which fall outside the boundaries of a UAB, and these cannot be addressed by that same UAB and require extended adaptive capacity from a second UAB (Fig. 4: Z). Additionally, surprise events that fall outside the boundaries of the introduction system can occur (Fig. 4: Q). The case showed several surprise events on all nested scales. Some examples near the boundaries (X) of the NS were the daily disruptions which could be solved by operations themselves using standardized scripts. A typical example outside the boundary (Y) was the additional capacity for train drivers to ensure availability in case necessary. Examples of surprise events outside the boundary (Z) were the disruptions caused by failures in the railway tracks besides failures in rolling stock. This requires adaptive capacity from the infrastructure manager. Surprise events outside the boundaries of the introduction system (Q) were, e.g., the changing political agreements of the Dutch and Belgian governments or the changing legislation with regards to ERTMS. If the Capacity for Maneuver (CfM) is limited, the train system becomes brittle and performance decreases.

In Sect. 5.1, the proposition is outlined. Following this, by comparing the theoretical outcome patterns, as put forward by the theory of graceful extensibility, to the empirical

outcome patterns from the V250 case, (mis)matches were identified and will be presented in Sect. 5.2.

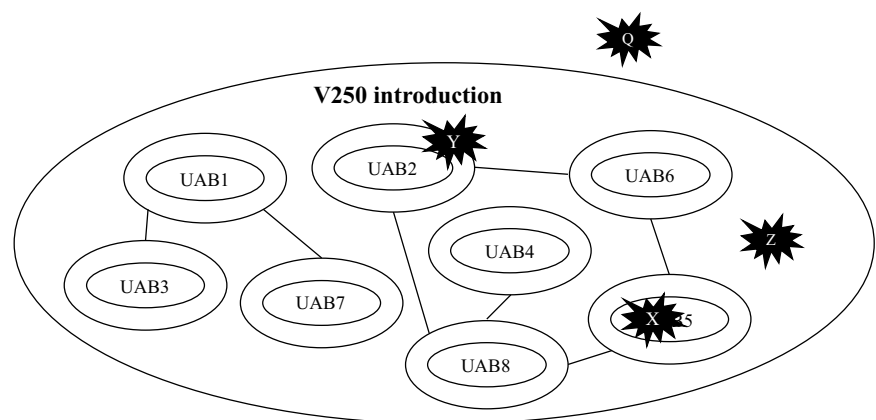
#### 1.4.1 Preposition

The main proposition in this study based on earlier research of Woods (2018) was: complex rolling stock introductions can benefit from graceful extensibility to sustain adaptability as demands change as a result of surprise events challenging the boundaries of the system. If the assumption of the authors is correct, and similar patterns are found, the theory of graceful extensibility might also be applicable in other long-term complex critical asset introductions.

#### 1.4.2 Pattern matching results

The theory of graceful extensibility entails the ability of a system to continuously extend its capacity to adapt when surprise events challenge its boundaries. It consists of 10 proto-theorems (S1–S10), categorized into three subsets as reported by Woods (2018). Following the research design as introduced in Sect. 3, the theoretical patterns of each proto-theorem were compared to empirically found patterns regarding sustained adaptability based on published and unpublished reports from the evaluation of the V250 introduction. The following three subsections summarize the qualitative results of the analysis using pattern-matching analysis. Detailed results have been included in the Appendices. Figure 5 shows an example of how the matching process was performed. The second column represents the theoretical sub-patterns and the fifth column shows the findings (coded items) including references to Atlas.ti. The third column translates the findings into an empirical pattern, which enables the match with the theoretical pattern. Based on this matching process, the fourth column states whether or not a match was observed.

**Fig. 4** Surprise events in the context of the V250 introduction



S1: The adaptive capacity of any unit at any scale is finite, therefore, all units have bounds on their range of adaptive behavior, or capacity for maneuver.				
#	Theoretical pattern	Empirical patterns	Pattern match	Findings from the case
S1.1	The location of boundaries to the ability to meet demands is uncertain. There is a boundary on any unit's adaptive capacity or the ability to be in-control or stay in-control as variations, disruptions and change occur.	Results showed that a new rolling stock introduction involves stakeholders from different units which all (need to) contribute to a reliable railway system. Boundaries on adaptive capacity were technical, cultural, political and organizational. A typical example was cross-border failures, which can not be solved by an UAB alone.	Observed	Findings showed that for example cross border challenges when (unexpected) failures occur cannot be controlled or solved by a single unit (10:60). Therefore, the adaptive capacity of a single unit has certain boundaries.
S1.2	There is a general parameter - Capacity for Maneuver (CfM), which specifies how much of the range the unit has used and what remains to handle upcoming demands.	The concept of CfM as such was not recognised. Patterns from the case showed the tendency to embrace the "belief" that everything becomes fluid under pressure, resulting in a too optimistic perspective in managing future (technical ) failures.	Not observed	Findings showed that, although the term CfM was not explicitly recognised or referred to, the ability to respond to upcoming events was overestimated which resulted in the inability to handle future (technical) failures.
S1.2a	All UAB's risk saturation, that is, running out of CfM as upcoming events present increasing challenges or demands. Managing the risk of saturation becomes the definition of what it means to be in-control.	Results showed the traditional approach to be in-control in case of increasing challenges in demand, which was illustrated by classic risk management practices, although not always executed correctly.	Observed	Findings showed that the implemented measures (lower frequency of train services and higher number of rolling stock reserves) to mitigate the increase in technical failures when commercially operated were insufficient to cope with the new situation (10:96). The lack of a fallback scenario was the result of the choice to primarily focus on repairing the V250 (26:1). Findings showed that at important moments in the introduction (tender phase and production phase), risks and the possible impact of these risks were not identified correctly and thoroughly (7:14). This pattern can be classified as "traditional" risk management theory and does not reflect the risk of saturation as proposed in the theory of graceful extensibility.

Fig. 5 Example of pattern-matching analysis for theoretical pattern S1

1.4.3 Subset A: managing risk of saturation (S1–S3)

Based on the assumptions that resources are finite and change is ongoing, the adaptive capacity of any unit at any scale is finite. Therefore, all units have bounds on their range of adaptive behavior. This is referred to as Capacity for Maneuver (CfM) (S1). Events which fall outside the bounds will always occur and demand response. Otherwise, the unit is brittle and performance may decrease (S2). As all UABs risk saturation of their adaptive capacity, they require some means to modify or extend their adaptive capacity when demands threaten their base range of adaptive behavior (S3). Based on focused coding of the dataset, several patterns were identified for sustained adaptability. These patterns were mapped to the first subset of proto-theorems of graceful extensibility, which consist of three proto-theorems and underlying patterns (refer to Appendix A for detailed results of the mapping).

- S1 All units have bounds on their adaptive capacity: Results show that the V250 introduction involved different UABs (e.g., infrastructure manager, operator, supplier, supervisor, consultant, and governments) which all (need to) contribute to ensure a reliable railway system. Boundaries on adaptive capacity were identified on different levels: technical, cultural, political, and inter- and intra-organizational. A typical example was cross-border failures, involving train and track systems from two different countries (The Netherlands and Belgium), where close cooperation is required to quickly address technical failures. This cannot be addressed by one UAB alone. As such, the concept of CfM was not observed. Patterns from the case showed the tendency to embrace the prevailing assumption that everything becomes fluid under pressure, resulting in an (overly) optimistic perspective in managing future (technical) failures. Results showed traditional risk management practices to be in-control.

- S2 Events will occur outside the bounds and demand response: The V250 shows that new rolling stock introductions are often characterized by ‘teething problem’ challenges. Therefore, reliability remained unpredictable as surprise events challenged the boundaries of the system. The CfM decreased as a result of ‘teething problems’ in both technical and organizational systems. The (fragile) interfaces between track and train increased the risk of brittleness when the system operated near its boundaries. Results reflected the attempt to gradually increase complexity during trial operations. Nevertheless, as previous research also states (Woods 2016), trial operations can never completely simulate commercial operations. Specific issues can only be identified after intensive use of the equipment in operations. An example of this is the TRAXX Amsterdam-Breda (April 2011), where failures suddenly emerged after a week in operation. V250 train sets had different failure modes, so each train can be considered unique. Just like the V250, new rolling stock will always deviate from existing rolling stock in both operations and maintenance and demands appropriate responses. As one of the engineers stated: “If we were stuck with current technologies, we would still use steam locomotives.”
- S3 Units modify and extend adaptive capacity: V250 results indicated an increase in effort and resources when the CfM decreased. The need for extended adaptive behavior was partly acknowledged by introducing a helpdesk for train drivers, additional support on the platforms, more capacity in tracks, and increase in turning points. A typical example was the absence of alternative fallback options in case of unreliable performance of the V250. Results showed the need for extended adaptive behavior, but this was often restricted by fixed strategies and plans. Furthermore, results showed patterns that indicated a slower pace of finding, deciding on, and implementing solutions than was required to meet

demand when disruption increased. Eventually, this resulted in a cancelation of all V250 train services as negative public opinion and political pressures mounted and the capacity to adapt was decreasing. Subcontractors were not involved in an early stage of the introduction process, which increased the risk of saturating CfM in a later stage. Furthermore, responses to standard failures depended on the train drivers involved and were mitigated by educating train drivers using standard solutions for standard failures.

In summary, due to the nature of the railway system, UABs depend on each other when adapting to surprising events. In managing the risk of saturation of adaptive capacity, UABs in the V250 introduction modified their base adaptive capacity, but did not (fully) utilize the network for enabling extended adaptive capacity. Main inhibitors were the fixed strategy and associated implementation plans formalized in legal agreements. Due to a large number of (international) stakeholders and the legalistic approach, clear and timely agreements were lacking, which caused ambiguity and uncertainty among stakeholders during the introduction.

Subset B: Networks of adaptive units (S4–S6).

As shown in statements S1–S3, graceful extensibility depends on how one UAB interacts with neighboring units in a network of interdependent units (subset B, refer to Fig. 1). No single unit can have sufficient range of adaptive behavior to manage the risk of saturation by itself. Therefore, synchronization across multiple UABs in a network is necessary (S4). Units in a network can monitor and influence other units in the network. Therefore, the risk of saturation can be shared within the network (S5). While independent units pursue their own goals and objectives, UABs generate points of pressure on other UABs which causes UABs to search for better operating points (S6). [Appendix B](#) includes the detailed mapping of theoretical and empirical patterns based on the case results.

- S4 Synchronization of UABs: The complexity of train and track, multitude of stakeholders, safety and security issues, political interests, large investments, major risks, and fragmented factual expertise required alignment and coordination during the V250 introduction. Findings illustrated the complex interdependencies between infrastructure and operator (Train Track Integration), which showed the need for alignment and coordination. Part of the problems with the V250 introduction were related to the availability of conventional and high-speed track, communications between train and track, and the response time in case of major disruptions. Findings also showed the possible limitations caused by the liberalization of the railway system, reflected in a lack of shared interests, willingness to share Capacity for Maneuver,

and an increase in formalized interfaces, often confirmed by legal agreements.

- S5 Risk of saturation can be shared: Findings show underlying patterns of collecting and sharing monitoring data for optimizing the system. Data were mainly collected and analyzed locally, increasing the risk of misalignment in the railway network. The ERTMS system required strong alignment and integration of operational systems of train and track for reliable communications. Fallback scenarios were not aligned among stakeholders, and the main contractor did not share all information regarding defects and failures to facilitate problem-solving. Findings show incompatible modes of operation among stakeholders during the homologation process. Case results also show the need for railway operators to involve subcontractors early in the introduction process for better alignment during and after introduction.
- S6 UABs search for better operating points under pressure: Findings show the underlying patterns of network pressures on UABs. Pressures from commercial interests and the media caused by the incident in which one of the V250 trains had lost a base plate due to ice formation and the continuing disruptions of other V250 trains, led to the full cancelation of the V250 train services on January 17, 2013 (Parliamentary Inquiry Committee Fyra 2015). From the beginning of the introduction, stakeholders (public and private) did not align their interests (financial, competitive, and political), (in)formal agreements were lacking and pressure mounted continuously. Chosen design principles for rolling stock were effective for one stakeholder (the contractor), but ineffective for other stakeholders further down the introduction chain (maintenance and operations). Conflicts of interest existed as the government was the sole shareholder of the privatized railway operator (NS), but simultaneously promoted liberalization among railway operators due to the liberalization of the European Railway industry. Therefore, one could argue that the architectural principles of the railway system did not fully support alignment and coordination of UABs responding to varying pressures on trade-offs. For example, the Dutch government pushed for high financial gains as a shareholder, but at the same time also demanded highly reliable train services as defined by punctuality requirements in the concession agreement. The NS was focused on maintaining their strategic position in a competitive market as the main railway operator in the Netherlands (Parliamentary Inquiry Committee Fyra 2015), which contributed to the optimistic views in the business case to win the concession.

Although results from the case show a lack of alignment and coordination among UABs, the observed patterns



correspond with similar ones found in other networks of adaptive units. As the results indicate, the main inhibitor for synchronization and sharing the risk of saturation among UABs was the lack of an integrated (holistic) perspective on the railway system (train and track) and the strict formal (legal and political) agreements between stakeholders (e.g., the turnkey approach in the purchase agreement as briefly described in Sect. 4).

Subset C: Outmaneuvering constraints (S7–S10).

Given the proto-theorems of networks of adaptive units, statements S7–S10 propose general constraints on the Capacity for Maneuver (CfM). There are two fundamental forms of adaptive capacity which allow for UABs to be viable: base- and extended adaptive capacity. Both are necessary, but inter-constrained (S7). UABs are local, have a certain position relative to the world and other units in the network: therefore, there is no best location in the network (S8). Furthermore, individual UABs each have their own perspective which is enriched by shifting and contrasting over multiple perspectives (S9). There are limits on models of adaptive capacity: therefore, mis-calibration is the norm and requires ongoing efforts from UABs to match actual capability (S10). [Appendix C](#) includes the detailed mapping of theoretical and empirical patterns based on the case results.

- S7 Base and extended adaptive capacity: Findings show that a distinction between base- and extended capacity was not taken into consideration by all stakeholders or not synchronized across the network. For example, the Belgian railway operator mainly focused on reducing costs and eliminating non-profitable activities, even when performance was near saturation. Monitoring of redundant systems was often not implemented. This increased the risk of saturation as train drivers and operators were unaware of defects in primary systems. A more robust system anticipates the failures of components which may require adaptations from other components.
- S8 No best location in the network: Findings show that certain UABs in the railway system caused many conflicts on a network-wide level due to local goals and interests. For example, the contractor's main objective was to produce and deliver rolling stock, not solving issues. However, railway operators were more concerned with acquiring the support from the contractor when issues arose. Cultural differences also complicated the relative positions of UABs in the V250 introduction, and their respective goals. Recommendations from the evaluation showed a strong preference for installing a central command to be in control, a so-called system integrator. The responsibilities of the inspectorate did not include ensuring that the entire railway system was able to provide reliable performance to railway passengers.
- S9 Shifting perspectives: Findings show a lack of mutual understanding among UABs caused by different perspectives on several matters. A typical example from a technological point of view was the various interpretations of the ERTMS standards by contractors, which resulted in poor interfaces and communications between systems. Results show that it was impossible to implement ERTMS in the track without specification of the requirements of rolling stock systems to ensure compatibility and interoperability. Findings also show the involvement of a multicultural group of stakeholders consisting of public and private companies from at least four different countries and the need to identify the 'DNA' of involved stakeholders upfront for a better understanding. For example, the difficult collaboration between Dutch and Belgian operators and infrastructure managers with respect to solving technical failures in the test process was partly a result of different attitudes regarding anticipating, or reacting to failures when they occur. Findings also show the need to involve train drivers, train managers, cleaning staff, and mechanics early in the introduction process to develop knowledge and expertise to ensure reliability and usability when commercially operated. As the main conclusion of the parliamentary inquiry shows (Parliamentary Inquiry Committee Fyra 2015, p. 4), the perspective of railway passengers was overlooked, while other interests prevailed.
- S10 Mis-calibration is the norm: Findings show patterns of over-optimism during all phases of the introduction. In hindsight, the call to start train services in December 2012, despite technical failures in trial operations and the winter season (risk of environmental influences), was too optimistic and reliability was at stake. Results show a pattern of strong pressures to start commercial operations, even if the train sets were not yet reliable. Insufficient awareness of train-track integration resulted in misalignment and a low rate of technical failures being resolved (e.g., ERTMS) in train and track, whereas a multidisciplinary approach to technical issues was required. Findings show reduced effort in exploring alternatives as fallback options in case of canceled train services due to constant rolling stock failures and increasing pressures. Workarounds were implemented to overcome system design failures. Data also show that these workarounds were not managed well and failures popped up periodically.

In summary, case results show patterns matching the theoretical patterns of S8–S10, except for the recognition of base and extended adaptive capacity (S7). The (formal) handovers from contractor to trial operations and from trial operations to commercial operations are also the appropriate

moments to reflect on the balance between base- and adaptive capacity as an increase in unexpected failures is likely to occur. Overlooking the perspective of the railway passenger and their interests was a typical example of S9 and one of the key constraints for the CfM in this case.

### 1.5 Confirming patterns and alternative explanations

This section discusses the confirmed patterns and explores alternative explanations, supported by relevant concepts from literature. By comparing the data of the V250 case to the 24 sub-patterns of the 10 proto-theorems, matches and mismatches were identified (refer to Appendices). Figure 4 illustrates the groundedness of each proto-theorem based on the V250 dataset. The groundedness indicates the degree of correspondence of the proto-theorems with the dataset. For instance, Fig. 6 shows that the reflection of the statements S1 and S8 in the dataset was limited (2%). On the contrary, the reflection of S6 (UABs search for better operating points under pressure) in the dataset was high. Statistics show that 31% of the coded items were related to subset A (S1–S3), managing risk of saturation, 46% were related to subset B (S4–S6), network of adaptive units and 23% were related to the subset C (S7–S10), outmaneuvering constraints. Although no conclusions can be drawn from these statistics, they show a broad reflection of the theoretical patterns in the case and the distribution among the three subsets.

In general, case patterns show high resemblance with the 10 proto-theorems of the theory of GE, resulting in 21 matching sub-patterns and three sub-patterns that were not fully observed in the case (details are included in the Appendices). Sub-patterns are marked with a second (number) or third (letter) suffix.

- a) The parameter Capacity for Maneuver (CfM), which specifies how much of the range the unit has used and what remains to handle upcoming demands (S1.2) was not recognized as such in the case, which is understandable as this (new) parameter currently lacks measurability.
- b) Risk becomes operationalized as some dynamic function of how CfM is being used and what remains compared to ongoing and possible future demands (S3. 3a). Case results show the need to optimize (traditional) control practices (e.g., timely identification of (shared) risks), but also the limitations of control and planning.
- c) The theory explicitly recognizes that there are two basic kinds of adaptive value, one far from saturation (base adaptive capacity) and another that operates near saturation (extended adaptive capacity) (S7. 1). Case results did not provide evidence for this distinction. However, saturation in complex rolling stock introductions differ from saturation in, e.g., commercial flights, where the scope for action in case of surprise events is limited. In the case study, the system became brittle, and eventually broke down, when the willingness to extend the CfM, for example by repairing the trains, was not broadly supported by the stakeholders involved in the network.

Case results (e.g., lack of integrated risk assessments, need for a system integrator, and more supervision) supported the need to optimize current control practices, but also outlined downsides of control, illustrated by, for example, many legal agreements and revisions of plans. This may imply the need for a more indeterministic perspective to avoid the ‘illusion of control’ (Langer 1975). The illusion of control refers to the notion that organizations are under the impression that they know more or less what is going to happen next. The focus on order

**Fig. 6** Degree of correspondance of GE patterns in V250 dataset (% of total number of 176 coded items)



and control is also reflected in most of organizational theory. Even before the work of Taylor (1914), management tended to assume that order is generally good, something to strive for, and that deviations from order, or disorder, are generally bad, and to be avoided (Shenhav 1995).

The Law of Requisite Variety (Ashby 1957) states that a controlling system can only control a system if it can generate the requisite variety to equal the variety generated by the system to control. This was restated by Beer (1985, p. 30): “only variety can absorb variety”. In other words, effective management control is only achieved when there is a balance between the control system, the controlled system, and the environmental system. However, management is caught between the desire to limit the variety of the organization (so as to control it) and the risk of limiting the variety of the organization to the extent that it cannot control its environment. Introna (1997) terms this the management control paradox. More control by the control system will limit the controlled system and thus may result in an inability to adapt to internal and external changes. One possible solution to the management control paradox is to locate control in the system, and hence, the system must control itself. In order for the organization to be structurally coupled with the environment, the concept of the manager as an ‘external controller’ must be eliminated (Introna 1997, p. 95). By shifting the perspective on planning from the observer—to the involved perspective—planning becomes crafting, as Mintzberg’s terms it (1994), or tinkering, as put forward by Ciborra (1996). Planning shifts from trying to find the rationally best alternative to negotiating meanings, translating actions, building alliances, and fixing obligatory passage points.

Another solution is accepting the fact that no accurate predictions are possible for every state of the system. Although this is considered psychologically disturbing, as it shows a lack of control over future outcomes, it results in increased benefits as the illusion of control will be avoided (Makridakis and Taleb 2009, p. 842). The concept of antifragility (Taleb 2012) may offer new insights into preparing for an uncertain future by embracing disorder. Taleb refers to fragility as the way in which a system suffers from the variability of its environment beyond a certain pre-set threshold, while antifragility refers to when it benefits from this variability (Taleb and Douady 2013). Furthermore, Taleb argues that we have been ‘fragilizing’ our systems by denying those stressors and disorder, making them vulnerable to surprise events. Nevertheless, most contemporary organizations do not like volatility, randomness, uncertainty, disorder, errors, stressors, or chaos. Yet, as the case introduction shows, disruption and randomness are increasing, and new approaches are required, as also observed by Martinetti et al. (2018).

## 1.6 Guidelines for adopting graceful extensibility in complex systems requiring sustained adaptability

The theory of GE (Woods 2018) is still in its infancy. Nevertheless, as it is based on empirical findings from former research and supported by the V250 case, it might already be valuable for organizations managing complex cyber-physical systems and striving for sustained adaptability. As with all new theories, operationalizing this theory to be applied to daily work is not an easy task. This section proposes guidelines for adopting graceful extensibility. Guidelines were identified and validated by key members of the case organization, based on the results of the pattern-matching analysis. These should be considered a starting point for new complex systems seeking sustained adaptability:

- Increase awareness of unexpected surprises in the network using historical (complex) projects and indicate the limitations of traditional risk management approaches;
- Assess the need for (base and extended) adaptive capacity for each unit in the network, and the network as a whole, based on the complexity involved over the lifetime of the introduction;
- Introduce the Capacity for Maneuver (CfM) as a parameter or key performance indicator for regulating the risks of saturation from an integrated network perspective. Compare integrated risk assessments to the assessment of the required adaptive capacity. This should lead to an initial understanding and shared agenda for action among UABs;
- Periodically challenge the ability of units in the system to extend capacity to adapt when surprise events challenge its boundaries and mitigate risks if necessary.

## 2 Conclusions and future research directions

This paper contributes to the field by assessing the explanatory power of the theory of graceful extensibility (GE) in a historical case study and provides guidelines for the operationalization of the theory in practice. Case results indicate that the majority of the theoretical and empirical patterns match, which provides evidence that the proposition is largely recognized. The proposition was defined as: “Complex rolling stock introductions can benefit from graceful extensibility to sustain adaptability as demands change as a result of surprise events challenging the boundaries of the system”. However, the parameter Capacity for Maneuver (CfM) was not recognized, which is required to manage the risk of saturation both at the level of an adaptive unit and

at the level of a network (Woods 2018). If organizations are ‘infected’ by the illusion of control, and assume high levels of predictability, surprise events are almost not considered and the organizations think that there is no need to explore CfM. Traditional control mechanisms are insufficient to deal with the increased complexity effectively. Therefore, the authors propose to adopt a more indeterministic approach, besides GE.

The usefulness of pattern matching in this study lies in supporting the authors’ assumptions that graceful extensibility can support future complex introduction of cyber-physical systems for sustained adaptability. However, as briefly touched on in Sect. 3, a research design based on pattern matching needs to consider several issues (Trochim 1989). Although most issues have been addressed by the researchers, two issues which require further explanation remain. The conceptualization of the theory was based on a single theory. As such, case results can only provide a (mis)match with the theory of GE: other theories on sustained adaptability were excluded. A second factor potentially limiting the

accuracy of the procedures required to provide evidence for a match lies in possible confirmation bias as both open and focused coding were performed by one researcher. Coding was evaluated by the case organization, but not by a second independent researcher.

This study can be considered a first attempt to empirically evaluate the applicability of the new theory on graceful extensibility. The scope of this study was limited to a single in-depth case study. Further empirical research is required to (dis)confirm the proto-theorems of GE. The convergence between information and operational technologies is expected to further increase complexity of the railway system, resulting in more surprise events which need to be managed. This will require human–technical systems that are able to continuously adapt to new (technical) challenges and demands. Although the V250 case can be considered an unsuccessful case of sustained adaptability, it may serve future rolling stock introductions and other complex cyber-physical asset introductions in their push for sustained adaptability when dealing with surprise events.

## Appendix A: Pattern matching analysis subset A: managing risk of saturation

S1: The adaptive capacity of any unit at any scale is finite; therefore, all units have bounds on their range of adaptive behavior, or capacity for maneuver

#	Theoretical pattern	Empirical patterns	Pattern match
S1.1	The location of boundaries to the ability to meet demands is uncertain. There is a boundary on any unit’s adaptive capacity or the ability to be in-control or stay in-control as variations, disruptions, and change occur	Results showed that a new rolling stock introduction involves stakeholders from different units which all (need to) contribute to a reliable railway system. Boundaries on adaptive capacity were technical, cultural, political and organizational. A typical example was cross-border failures, which cannot be solved by an UAB alone	Observed
S1.2	There is a general parameter—Capacity for Maneuver (CfM), which specifies how much of the range the unit has used and what remains to handle upcoming demands	The concept of CfM as such was not recognized. Patterns from the case showed the tendency to embrace the “belief” that everything becomes fluid under pressure, resulting in a too optimistic perspective in managing future (technical) failures	Not observed
S1.2a	All UAB’s risk saturation, that is, running out of CfM as upcoming events present increasing challenges or demands. Managing the risk of saturation becomes the definition of what it means to be in-control	Results showed the traditional approach to be in-control in case of increasing challenges in demand, which was illustrated by classic risk management practices, although not always executed correctly	Observed
S2: Events will occur outside the bounds and will challenge the adaptive capacity of any unit, therefore, surprise continues to occur and demands response, otherwise the unit is brittle and subject to collapse in performance			
#	Theoretical pattern	Empirical patterns	Pattern match
S2.1	There are recurring patterns that characterize model surprise—how events challenge boundaries	New rolling stock introductions are characterized by teething problems both technical as organizational. The reliability of new rolling stock systems are unpredictable as surprise events will challenge the boundaries of the system	Observed
S2.1a	Events will occur at some rate and of some size and of some kind that increase the risk of saturation—exhausting the remaining CfM	The capacity to maneuver was exhausted by teething problems in processes, technology, infrastructure, and organizations	Observed

S2.1b	Brittleness is how rapidly a unit's performance declines when it nears and reaches its boundaries (S1). Brittleness describes how a UAB performs near, at and beyond its boundaries, separate from how well it performs when operating far from its boundaries	The (fragile) interface between track and rolling stock causes brittleness when operating near its boundaries. Complexity is gradually introduced during trial operations but teething problems in commercial operations can only partly be transferred to trial operations, never completely	Observed
S2.1c	The range of adaptive behavior of a UAB is a model of fitness; that model has boundaries (S1) and events occur which fall outside that boundary → model surprise	Results showed the need for adaptive behavior. Trial operations can never completely simulate commercial operations. Therefore, adaptive behavior is required to deal with surprises but has a certain range.	Observed
S2.1d	Events that occur near or outside a UAB's boundary increases the risk saturation, and this occurs independent of how well that UAB matches responses to demands (the degree of fit) well within its range of adaptive behavior (or competence envelope)	The high density of the railway network, which will further increase due to an increasing passenger demand, new technologies and the increase in high-frequency train services, will further challenge the adaptive behavior of new introductions	Observed
S3: All units risk saturation of their adaptive capacity, therefore, units require some means to modify or extend their adaptive capacity to manage the risk of saturation when demands threaten to exhaust their base range of adaptive behaviour			
#	Theoretical pattern	Empirical patterns	
S3.1	The work (effort/energy/resources) required to adapt and handle changing demands increases as CfM decreases, i.e., there is some function relating effort to be in-control to the risk of saturating CfM	Case results indicated an increase in efforts and resources when optionality or the CfM decreased. Onboarding of subcontractors in an early stage of the introduction process decreases the risk of saturating CfM. Although the lead time of introductions is on average 6 years, preparations for new rolling stock should start in time	Observed
S3.2	As risk of saturation increases and CfM approaches exhaustion, UABs need to adapt to stretch or extend their base range of adaptive behavior to accommodate surprises. This extended form of adaptive capacity is graceful extensibility: how to deploy, mobilize, or generate capacity for maneuver when risk of saturation is increasing or high	Observed patterns showed the need for extended adaptive behavior, but was restricted by fixed strategies and plans in advance. The need for adaptive behavior when introducing the V250 was acknowledged by introducing a helpdesk for train drivers, additional support on the platforms, more capacity in tracks, and turning points. The capacity to maneuver in case of these measures did work was lacking. A typical example was the absence of alternative options in case of unreliable performance of the V250	Observed
S3.3	The risk of saturating controls as demands grow and cascade creates systematic patterns in how adaptive systems break down. The first systematic pattern is decompensation, which is, exhausting the capacity to adapt as disturbances/challenges grow and cascade faster than responses can be decided on and deployed to effect	Results showed patterns which indicated a slower pace of finding, deciding on, and implementing solutions than demands required when disturbances grew. Eventually this resulted in a cancelation of all V250 train services as public opinion and political pressures were cascading faster and the capacity to adapt was decreasing	Observed
S3.3a	All UABs have some potential for adaptive response when information varies, conditions change, or when new kinds of events occur, any of which challenge the viability of previous adaptations, models, plans, or assumptions. Concepts about varieties of adaptive capacity can be integrated around the single parameter of Capacity for Maneuver (CfM) and how UABs adjust/regulate their adaptive capacities relative to the risk of saturating CfM as they respond to future challenges and opportunities. The struggle for fitness in the face of changing demands is ongoing and requires the potential to adjust adaptive capacities. This leads to a new operational and actionable definition of brittleness as the risk of saturating CfM and to the concept of graceful extensibility as the opposite of brittleness. Risk here becomes operationalized as some dynamic function of how CfM is being used and what remains relative to ongoing and possible future demands	The terms of CfM, brittleness and graceful extensibility as defined are partly observed in practice. For example, responses to standard failures depended on the train drivers involved and were mitigated by educating train drivers using standard solutions for standard failures. This decreased the risk of saturating CfM. However, the operationalization of risk as a dynamic function of how CfM was used was not observed	Not observed

## Appendix B Pattern matching analysis subset B: networks of adaptive units

S4: No single unit, regardless of level or scope, can have sufficient range of adaptive behavior to manage the risk of saturation alone; therefore, alignment and coordination are needed across multiple interdependent units in a network

#	Theoretical pattern	Empirical patterns	
S4.1	UABs exist in and are defined relative to a network of interacting and interdependent UABs at multiple scales → networks with multiple roles, multiple echelons.	Observed outcomes indicated that decision-making on investments and priorities in train and track are strongly interdependent, also called train-track integration (TTI). If you are unaware of the technical specifications of new rolling stock, it is difficult to decide on the right investments in infrastructure. The complexity of train and track, multitude of stakeholders, safety and security issues, political interests, major investments, major risks, and fragmented factual expertise requires alignment and coordination.	Observed
S4.1a	As risk of saturating the base adaptive capacity grows, additional adaptive capacity must be brought to bear, and this requires invoking other UAB that extend CfM beyond the remaining capacity of the unit at risk of saturation. To bring additional adaptive capacity to bear, requires alignment, coordination, and synchronization across multiple units and echelons.	Findings illustrate the complex interdependencies between infrastructure and operators (train and track), which shows the need for alignment and coordination. Part of the v250 problems were related to the availability of conventional and high speeds tracks, communication between train and track and the response time in case of major disruptions. Findings showed the limitations of the liberalization of the railway system, reflected in a lack of shared interests and willingness to share capacity for maneuver. Liberalization caused an increase in complexities due to a more formalized interfaces.	Observed

S5: Neighboring units in a network can monitor and influence—constrict or extend—the capacity of other units to manage their risk of saturation, therefore, the effective range of any set of units depends on how neighbors influence others as the risk of saturation increases somewhere in that neighborhood of the network

#	Theoretical pattern	Empirical patterns	
S5.1	Misalignment and mis-coordination across UABs increases the risk of saturating control as demands grow and cascade. This creates a second form of adaptive system breakdown—working at cross-purposes where one UAB responds to demands by managing its CfM in ways that reduce the CfM of UABs nearby or at a larger or finer scales. When this occurs it reveals a general pattern of responses that are locally adaptive (from one perspective), but globally maladaptive (from a different perspective). On the other hand, some UABs monitor the risk of saturating CfM in another UAB by monitoring signals associated with the increasing effort to stay in-control. When they recognize that the risk of saturating the CfM of the other unit is becoming too high, they respond in ways that have the effect of extending the capacity and behavior of the UAB at risk.	Results showed underlying patterns of collecting and sharing the right monitoring data for global performance of the system. Data was collected locally, but resulted in misalignment in the network. The ERTMS system required a strong alignment and integration of operational systems of train and track for reliable communications. Fallback scenarios were not aligned among stakeholders and the main contractor did not share all observed defects and failures. Findings showed incompatible modes of operation among stakeholders. Case results showed the need for the railway operator to involve subcontractors early in the introduction program for better alignment.	Observed

S6: As other interdependent units pursue their goals, they modify the pressures experienced by a UAB of interest. In response to changing experienced pressures, a UAB searches for better operating points in a multidimensional trade space

#	Theoretical pattern	Empirical patterns	
S6a	In pursuing their goals, a Unit of Adaptive Behavior (UAB) generates pressure on neighboring UABs. As a result, the goals UABs pursue or prioritize are changed relative to the pressures they experience and the conflicts these pressures exacerbate or generate. As the pressures generated by other interdependent units change, the trade-offs a unit faces change. The pressures experienced influence the search for how to balance or prioritize across basic trade-offs, especially when trade-offs' intensify (Woods 2006). This constraint poses the research question—what architectural properties of the network influence the way units in a network respond to varying pressures on trade-offs?	Results showed the underlying patterns of network pressures on UABs. Specifically, the commercial and societal pressures eventually led to the cancelation of the V250 train services. Railway passengers are part of the logistic production system. Trade-offs were observed in the resource planning of train managers and train drivers on other tracks. From the start of the project, stakeholders (public and private) did not align their interests (financial, competitive, political), (inf)formal agreements were lacking and pressures build up continuously. Chosen design principles for rolling stock were effective for one stakeholder (the contractor), but ineffective for other stakeholders further down the introduction system (maintenance and operations). Conflict of interests exists as the government is sole shareholder of the privatized NS company, but at the same time promotes liberalization among railway operators. Therefore, one could argue that the architectural principles of the current railway system do not support alignment and coordination of UABs to respond to varying pressures on trade-offs.	Observed

## Appendix C Pattern matching analysis subset C: outmaneuvering constraints

S7: Performance of any unit as it approaches saturation is different from the performance of that unit when it operates far from saturation, therefore there are two fundamental forms of adaptive capacity for units to be viable—base and extended, both necessary but inter-constrained.

#	Theoretical pattern	Empirical patterns	
S7.1	To extend, adaptive capacity requires mechanisms that consume resources; investing in the resources that provide the extended adaptive capacity negatively impacts on base adaptive capacity. And the reverse holds—improving base adaptive capacity in isolation reduces the resources that underpin the capacity to extend response capability when risk of saturation is high. Net adaptive value, as a sense of fitness, includes both. Adaptive value is a term often used in models of how biological and neurobiological systems increase their fitness to a changing environment (e.g., Bialek et al. 2007). The ‘value’ refers to the advantage in fitness gained for the unit in question when it adapts. The theory builds on this tradition and recognizes explicitly that there are two basic kinds of adaptive value—one far from saturation and another that operates near saturation. Operating far from saturation, when criteria are oriented toward optimality (that is, pressures for adding value to base adaptive capacity), gains come from achieving a reference level of performance from a reduction in resources (more efficiency or productivity). For graceful extensibility needed near saturation, adding adaptive value comes from expanding the performance possible from a reference level of resources. This leverages the adaptive value from a set of available resources to produce and sustain graceful extensibility.	Results showed that extended adaptive capacity was not considered by all stakeholders. For example, the main focus of Belgian operators was on reducing costs and non-profitable activities. The introduction of new rolling stock in an existing railway infrastructure requires much effort of the organization in their base adaptive capacity and extended adaptive capacity as a result of (un) expected teething problems. Do the upfront investments outweigh the decreased risk of saturation? Monitoring of redundant systems is often not implemented. This increases the risk of saturation as train drivers and operators are unaware of the defect. A more robust system already anticipates on the failures of components which may require adaptations from other components. This needs to be considered in the specification/requirements of the design.	Not observed

S8: All adaptive units are local—constrained based on their position relative to the world and relative to other units in the network, therefore there is no best or omniscient location in the network.

#	Theoretical pattern	Empirical patterns	
S8.1	A UAB is embedded in a place relative to an environment and a set of relationships across a network of UABs. A UAB is responsible for goals relative to its local position in the network—responsible in the sense that that the UAB experiences that consequences that result from achieving or failing to achieve its goals. Different UABs in the network are differentially responsible for different subsets of goals that can interact and conflict.	Results showed that due to the many stakeholders/UABs in the railway system goals and interests caused many conflicts. For example, the contractor’s main objective is to produce and deliver rolling stock. Railway operators are more concerned with following-up on issues with support of the contractor. Cultural differences also confirmed the relative position of UABs in a network and their goals.	Observed

S9: There are bounds on the perspective of any unit—the view from any point of observation at any point in time simultaneously reveals and obscures properties of the environment—but this limit is overcome by shifting and contrasting over multiple perspectives.

#	Theoretical pattern	Empirical patterns	
S9.1	Each UAB in a network has a perspective where perspective consists of a point of observation (think of this as the position of a virtual camera) relative to a point of interest in a scene which defines a view direction and a field of view. The view from any point of observation simultaneously reveals and obscures properties of the environment. There is no best perspective. To see perspective requires another perspective (or a perspective shift).	Results showed the pattern of the need for perspective shifts. A typical example from a technology perspective was the different interpretations of contractors of the ERTMS standards which resulted in bad interfaces between systems. Complicating factor is the train-track integration. It is impossible to implement ERTMS in the track, if you do not specify the requirements of rolling stock systems to ensure compatibility. A second observation was the multicultural group of stakeholders consisting of public and private companies from at least four different countries. Lessons learned was the need to identify the DNA of involved stakeholders upfront to understand the entire system by exploring multiple perspectives. A third observation was the need to involve train drivers, train managers, cleaners and mechanics early in the introduction for knowledge transfer and developing expertise to ensure reliability and usability (RAMS).	Observed

S10: There are limits on how well a unit's model of its own and others' adaptive capacity can match actual capability, therefore, mis-calibration is the norm and ongoing efforts are required to improve the match and reduce mis-calibration.

#	Theoretical pattern	Empirical patterns	
S10.1	A UAB's model of itself and others will be mis-calibrated without mechanisms to shift and contrast perspectives. Mis-calibration risks include all of the parameters of networks of UABs defined previously (e.g., boundaries, risk of saturation, demands, perspective).	Results showed the pattern of strong commercial pressures to transfer trains to commercial operations, even if they were not completely reliable. This resulted in too much optimism in improving defect trains by the main contractor. The call to start train services in December 2012 despite technical failures in trial operations was too optimistic.	Observed
S10.2	Since risk of mis-calibration is omnipresent, effort must be invested to reduce risk of mis-calibration. In other words, since there is a bound on how well models of capability match actual capability, effort must be invested to improve the match.	Insufficient attention to train-track integration resulted in a mis-calibrated system and a low rate of resolving technical failures (mainly ERTMS) in train or track, whereas multi-disciplinary approach to technical issues is required.	Observed
S10.3	To fail to continue to check and adjust calibration means that learning will slow or stop. This learning breakdown defines the third basic form of maladaptive behavior: where models of adaptive capacity become stuck and outdated as a result of change. Given changes afoot, models of demands and models of effective responses to those demands, which had been adaptive in the past, become stale, are no longer effective and require revision.	Observations showed reduced effort in exploring alternatives for the backup or fall-back in case of canceling train services due to permanent rolling stock failures. As a result adaptive capacity became outdated and learning stopped. Workarounds were implemented to cover-up system design failures. Results showed that these workarounds were not managed well, as a result failures popped up from time to time and learning broke down.	Observed



S10.4	Boundary areas are discovered and known only through the experience of surprise and the experience of risk of saturation. Furthermore, changing to handle the risk of saturation produce change to the system adapting. These changes modify what is base adaptive capacity and modifies what and when and where surprise occurs.	Findings indicate that the warranty (and aftercare) process, that is aimed at removing technical and organizational imperfections of equipment and equipment-related transport processes after commercial operations, needs some improvements. Most teething problems occur within three years after commissioning. This is also the time to discover boundary areas through surprises.	Observed
S10.4a	A UAB has limits on its ability to model its own and other's ability to regulate CfM including the risk of saturating CfM. It tends to underestimate demands and how they change and to overestimate base adaptive capacity. When mis-calibrated, UABs are under-responsive to changes in demands and slow to learn and adopt new responses to handle the changes. As the location of boundaries are uncertain and dynamic, mis-calibration further limits a UAB's ability to explore boundary areas and update models. Thus, mis-calibrated UABs tend to act in ways that constrict the CfM of other units in the network.	Results showed several patterns of over-optimism during all phases of the introduction. It started with the business case, followed by the tender process, the production process, testing phase and finally the introduction process. By over-estimating their base adaptive capacity to handle upcoming changes, other UABs were constricted in their CfM. In the end, all "issues" ended up in commercial operations as a result of under-responsive UABs.	Observed

**Acknowledgements** The authors would like to thank the case organization of NS for their cooperation and participation in this study, even though the interpretations or conclusions of the authors may differ from theirs.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Akanmu A, Anumba CJ (2015) Cyber-physical systems integration of building information models and the physical construction. *Eng Constr Arch Manag* 22(5):516–535. <https://doi.org/10.1108/ECAM-07-2014-0097>
- Andersson P, Mattsson L (2015) Service innovations enabled by the “internet of things.” *IMP J* 9(1):85–106. <https://doi.org/10.1108/IMP-01-2015-0002>
- Ashby WR (1957) *An introduction to cybernetics*. Chapman and Hall Ltd, London
- Baheti R, Gill H (2011) Cyber-physical systems. *Impact Control Technol* 12(1):161–166
- Beer S (1985) *Diagnosing the system for organizations*. John Wiley and Sons, US
- Carlson JM, Doyle J (2000) Highly optimized tolerance: robustness and design in complex systems. *Phys Rev Lett* 84(11):2529–2532. <https://doi.org/10.1103/PhysRevLett.84.2529>
- Ciborra C (1996) *The platform organization: recombining strategies, structures, and surprises* (Vol. 7)
- Hermann M, Pentek T, Otto B (2016) Design principles for industrie 4.0 scenarios. Paper presented at the system sciences (HICSS), 2016 49th Hawaii International Conference on
- Hollnagel E, Cacciabue PC (1999) Cognition, technology and work: an introduction. *Cogn Technol Work* 1(1):1–6. <https://doi.org/10.1007/s101110050006>
- Introna L (1997) *Management, information and power: a narrative of the involved manager*. Macmillan, US
- Johns G (2006) The essential impact of context on organizational behavior. *Acad Manag Rev* 31(2):386–408. <https://doi.org/10.5465/amr.2006.20208687>
- Kagermann H, Helbig J, Wahlster W (2013) Recommendations for implementing the strategic initiative INDUSTRIE 4.0: securing the future of German manufacturing industry; final report of the Industrie 4.0 Working Group: Forschungsunion
- Langer EJ (1975) The illusion of control. *J Pers Soc Psychol* 32(2):311
- Makridakis S, Taleb N (2009) Living in a world of low levels of predictability. *Int J Forecast* 25(4):840–844. <https://doi.org/10.1016/j.ijforecast.2009.05.008>
- Martinetti A, Moerman J, Van Dongen LAM (2018) Storytelling as a strategy in managing complex systems: using antifragility for handling an uncertain future in reliability. *Safety Reliability*. <https://doi.org/10.1080/09617353.2018.1507163>
- Mintzberg H (1994) The fall and rise of strategic planning. *Harvard Business Rev* 72(1):107–114
- Parliamentary Inquiry Committee Fyra (2015) *De reiziger in de kou* (33678) Retrieved from The Hague, Netherlands: [https://www.tweedekamer.nl/sites/default/files/atoms/files/rapport\\_dereiziger\\_indekou\\_enquetcocommissiefyra\\_kst-33678-11.pdf](https://www.tweedekamer.nl/sites/default/files/atoms/files/rapport_dereiziger_indekou_enquetcocommissiefyra_kst-33678-11.pdf)
- Perrow C (1984) *Normal accidents: living with high risk technologies*. Princeton University Press, Princeton
- Porter ME, Heppelmann JE (2014) How smart, connected products are transforming competition. *Harvard Business Rev* 92(11):64–88
- Ramasesh RV, Browning TR (2014) A conceptual framework for tackling knowable unknown unknowns in project management. *J Oper Manag* 32(4):190–204. <https://doi.org/10.1016/j.jom.2014.03.003>

- Saunders FC, Gale AW, Sherry AH (2016) Responding to project uncertainty: evidence for high reliability practices in large-scale safety-critical projects. *Int J Project Manage* 34(7):1252–1265. <https://doi.org/10.1016/j.ijproman.2016.06.008>
- Shenhav Y (1995) From chaos to systems: the Engineering Foundations of Organization Theory, 1879–1932. *Adm Sci Q* 40(4):557–585. <https://doi.org/10.2307/2393754>
- Silfhout M, Bergv A van den (2014) *De ontsporing* (V. 1 Ed.): Zilverster media
- Taleb NN (2012) *Antifragile: things that gain from disorder* (Vol. 3): Random House Incorporated
- Taleb NN, Douady R (2013) Mathematical definition, mapping, and detection of (anti) fragility. *Quant Finance* 13(11):1677–1689
- Taylor FW (1914) *The principles of scientific management*: Harper
- Trochim WMK (1989) Outcome pattern matching and program theory. *Eval Program Plann* 12(4):355–366. [https://doi.org/10.1016/0149-7189\(89\)90052-9](https://doi.org/10.1016/0149-7189(89)90052-9)
- Wears R, Perry S, Anders S, Woods D (2008) Resilience in the emergency department. *Resilience engineering: remaining open to the possibility of failure*. Ashgate studies in resilience engineering. Ashgate Publishing, UK
- Weick KE, Sutcliffe KM, Obstfeld D (2008) Organizing for high reliability: processes of collective mindfulness. *Crisis Manag* 3:81–123
- Woods DD (2015) Four concepts for resilience and the implications for the future of resilience engineering. *Reliab Eng Sys Saf* 141:5–9
- Woods DD (2016) The risks of autonomy: Doyle's catch. *J Cogn Eng Decis Mak* 10(2):131–133. <https://doi.org/10.1177/1555343416653562>
- Woods DD (2018) The theory of graceful extensibility: basic rules that govern adaptive systems. *Environ Sys Decis*. <https://doi.org/10.1007/s10669-018-9708-3>
- Woods D, Branlat M (2011) Basic patterns in how adaptive systems fail. In: Hollnagel E, Paries J, Woods D, Wreathall J (eds) *Resilience engineering in practice*. Ashgate, Farnham, pp 127–144
- Yin RK (2003) *Case study research: design and methods*: Sage publications
- Yoo Y, RichardBoland JJ, Lyytinen K, Majchrzak A (2012) Organizing for innovation in the digitized world. *Organ Sci* 23(5):1398–1408. <https://doi.org/10.1287/orsc.1120.0771>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.