



Why the system performance curve may fall short: a demonstration using linear time invariant feedback control

Tobias Demmer¹ · Daniel Eisenberg² · Riccardo Patriarca³

Received: 28 February 2026 / Accepted: 1 September 2026
© The Author(s) 2026

Abstract

The most common model for engineers to explain, measure and quantify resilience is a system performance curve that tracks system function over time in response to a failure event (also frequently called a resilience triangle, resilience curve, or rebound curve among others). While it is well recognized that resilience is more than just recovery from stress or functioning under pressure, the curve remains highly used. By discussing five cases using a generic flow-shop infrastructure example, this work highlights what may be overlooked when focusing solely on performance. The findings demonstrate that while there are benefits and some comparative strengths, a more nuanced understanding of the system performance curve is necessary, as many design and operational factors might be easily misinterpreted or missed.

1 Introduction

Resilience is broadly understood as the ability of a system to respond to change and plays a key role in the management of critical infrastructure systems. Still, its clear definition and application for infrastructure management remains elusive. This is partly because the term resilience is very popular across research streams (Woods 2015) and often acts as an “umbrella term” (McCubbin 2001) comprising many different infrastructure system capacities or dimensions (Mentges et al. 2023). Some authors emphasize the ability to recover (e.g. Teodorescu 2015), some the ability to absorb (e.g. Vugrin et al. 2011), some the ability to learn (e.g. Norris et al. 2008). It has been also interpreted as the envelope of four abilities, i.e. monitoring, learning, responding, anticipating

(Hollnagel 2018). Yet, resilience is often defined via different phases in infrastructure failure and disaster management, including the preparation, absorption, recovery, and adaptation to unforeseen events (Cutter et al. 2013; Klimek et al. 2019).

More formally, Woods (2015) outlines at least four differing perspectives on resilience, including concepts of graceful extensibility for systems to stretch capacities and sustained adaptability to adapt over long time scales, both of which are rarely considered in infrastructure settings (Sharkey et al. 2021).

Despite all these nuanced definitions, one very common approach to describe and assess resilience is a curve which tracks infrastructure system performance over time. We refer to this as a system performance curve, but it goes by many names in the literature, including the resilience triangle (Bruneau et al. 2003), resilience curve (Poulin and Kane 2021), and rebound curve (Eisenberg et al. 2025) among others. The primary use of the system performance curve is to describe, measure, and compare system response and recovery to stress (see for a review: Poulin and Kane 2021). The curve aligns well with resilience definitions based on the four phases of disaster response, including preparation, absorption, recovery, and adaptation (see Mentges et al. 2023; Fig. 1). Accordingly, the curve is often used to communicate and define what resilience is via these phases, further simplifying the explanation of related concepts such as robustness, vulnerability, and rapidity of recovery.

✉ Tobias Demmer
tobias.demmer@dlr.de

Daniel Eisenberg
daniel.eisenberg@nps.edu

Riccardo Patriarca
riccardo.patriarca@uniroma1.it

¹ Institute for the Protection of Terrestrial Infrastructures, German Aerospace Center (DLR), Sankt Augustin, Germany

² Department of Operations Research, Naval Postgraduate School, Monterey, USA

³ Department of Mechanical and Aerospace Engineering, Sapienza University of Rome, Rome, Italy

Researchers also use the system performance curve to quantify resilience (see Yodo and Wang (2016) for an overview). Indeed the original work introducing the curve (Bruneau et al. 2003) did so in order to quantify resilience of infrastructure. Today it is used across application domains by measuring the difference between system function with and without disruptions, most often as also proposed by Bruneau et al. as the area under the curve (Hosseini et al. 2016). In this view, more resilient systems will suffer less performance loss and/or recover more quickly, producing a smaller difference in the area under the curve, while less resilient systems will suffer greater performance loss and/or take longer to recover. This measure then facilitates comparisons between systems responding to adverse events or between potential system designs or alternative investment decisions. Some researchers even use the curve to disaggregate the concepts of risk and resilience (Aven 2022; Linkov et al. 2014), as risk analysis measures the probability of failure and resilience is related to its change in function.

However, the efficacy of the performance curve as model and measure of resilience has been recently challenged. Eisenberg et al. (2025) outline shortcomings of the curve which stem from fundamental issues with its theoretical and practical application for infrastructure and community systems. Critiques span each element of the curve, including its representation of system function (vertical axis), time (horizontal axis), dynamics (the curve itself), and associated decisions (based on metrics like the area under the curve). Ultimately, the authors argue that results derived from the performance curve lack explanatory power for why some infrastructure systems are resilient and others are not, which makes it a flawed model.

Still, Eisenberg et al. (2025) only provide a conceptual, qualitative argument for the curve's shortcomings based on past studies and review, and do not provide quantitative models or case studies that demonstrate the reported limitations.

This work addresses this gap by showing issues with the performance curve and associated resilience measures via a quantitative model. We develop a model that produces performance curves similar to those found across the literature. We then study a series of relevant case studies for infrastructure systems. Our results are mixed, showing some positive and negative elements of the model. Results demonstrate:

- A measure of resilience as the difference in loss produced by a system performance curve does provide some comparative power when recovery dynamics are stable and reach intended performance goals. However, recommendations may be idiosyncratic to the mathematical model. For example, our work demonstrates that common resilience metrics preference early recovery

(even if slow) over faster recovery that may take longer to initiate.

- The curve does provide a way to compare recovery dynamics given differences in system structure and information flows, where poor choice of sensing produces worse outcomes.

We also find the following issues with performance curve models:

- The preference for early recovery may be idiosyncratic to our choice of failure and recovery dynamics, such that different performance curve models may recommend the opposite.
- Measuring resilience using a performance curve does not make sense when recovery dynamics are not well-behaved. For example, many systems surge or overshoot performance goals that confuse calculation and interpretation of most resilience measures based on performance curves.
- Measuring resilience for the final use of a critical service also hides relevant dynamics and decisions made up-stream by interdependent subsystems (e.g., service production and distribution).

Overall, we conclude that the benefits of performance curve are highly sensitive to its use. When used as an aggregate measure (as discussed, for example, in Ayyub 2015; Linkov and Trump 2019), sub-system performance dynamics and causal events that may inform resilience are hidden and limit its efficacy. Moreover, dynamic slack resources that provide complementary forms of resilience confound performance curves as metric. Hence, our quantitative results corroborate the qualitative argument by Eisenberg et al. (2025): *the system performance curve may appear as an effective quantitative representation of system resilience but tends to fall short of offering a comprehensive tool for improving systems*. At minimum, the performance curve must be augmented with additional information to be useful, either at a multi-scale level that includes system and sub-system performance or via the tracking of entirely different resources and capabilities.

In the worst case, the performance curve hides and does not help address the root causes of system vulnerability or provide pathways for adaptation. Resilience, in its real-world application, requires a dynamic and iterative process, involving learning, foresight, and adaptation to changing conditions. The curve, by focusing primarily on the idea of a system returning to its previous state, neglects the need for systems to evolve and transform in response to ongoing challenges or to build long-term adaptive capacity. As a result, the curve inadvertently limits the scope of interventions.

2 Background and methods

Infrastructure system performance curves which were originally described in a phenomenological manner without a mathematical description. Thierney and Bruneau (2007) suggested the simplifying assumption that the performance drop can be approximated by a triangular-shaped curve (the “resilience triangle”) that is defined by the performance minimum, the beginning of the disruption and the end of the performance loss. Numerous rebound-based measures of performance change and resilience were derived from this original description. Many of these measures remain phenomenological, focusing on observed or estimated system performance rather than on mathematical, equation-based models capable of estimating and predicting rebound dynamics (Poulin and Kane 2021).

More recent efforts produce system performance curves using mathematical expressions that split the curve into a disruption and a recovery phase for resilience quantification and analysis. For example, Sharma et al. (2018) discussed which function fits best for the description of the recovery phase. Cimellaro et al. (2010) analyzed different mathematical approaches such as linear, trigonometric and exponential expressions and conclude that the type of the function depends on the system and society preparedness response. Recent works that try to describe the resilience performance curve using mathematical descriptions are Demmer et al. (2022) and Stolz et al. (2024). However, most performance curve models remain generic, rather than fit to data (e.g., historical events), suggesting there is no consensus on which mathematical model best captures infrastructure dynamics, and any model is equally good.

2.1 A system performance curve model

In this work, we use a chain of PT_n -elements (n -th order plus dead time) to model an infrastructure during a disturbance. These linear time-invariant (LTI) systems consist of a gain factor that scales the input signal (P), and a series of n first-order time delay elements (T), each introducing

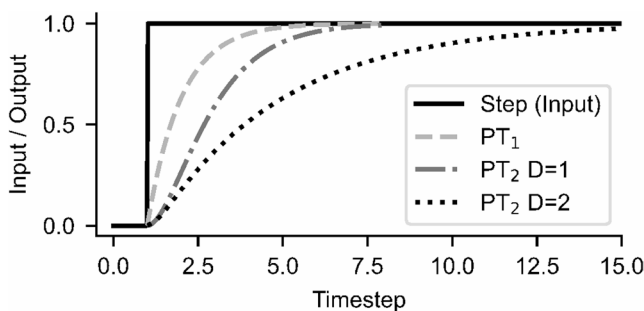


Fig. 1 Input step function and system reaction with various PT reactions

a certain inertia or lag in the system response. The concept of PT reactions is rooted in the fundamental principles of control theory and system dynamics. When describing a system as LTI, this means that the output depends only on its current input and internal state variables without any memory of past inputs or outputs. For example, a first-order lag system like PT_1 responds to a step change in input with an exponential rise time, reaches 63.2% (or $1/e$) of its final value within one time constant (τ) and decays exponentially back to zero after the step is reverted.

The transfer function of a PT_1 system (also known as first order system) is represented mathematically via Eq. (1) and PT_2 via Eq. (2) where $G(s)$ is the transfer function, K is the gain or proportionality constant, T represents the time delay or time constant, s is the complex frequency and D is the damping factor.

$$G(s) = \frac{K}{Ts + 1} \quad (1)$$

$$G(s) = \frac{K}{T^2s^2 + 2DTs + 1} \quad (2)$$

As already mentioned, the PT_1 system’s response to a step input is characterized by an exponential rise or decay. The PT_n response, that is a series connection of multiple PT_1 elements, approaches a logistic behavior. The delay is influenced by the damping factor D (see Fig. 1).

PT elements are often used in control theory to model any kind of delayed processes. The step function describes the presence of a disturbance and following that a PT element could be used to describe the delayed system response. If resilience is understood as a dampening factor to some disturbance, it is suitable to model it using a PT element, viewing the disturbance as a step function that initiates a decline in system performance. Indeed, the resulting performance curve from a PT-model has the shape of a S described in resilience theory. The PT-model can produce diverse failure and recovery dynamics that may be idiosyncratic to the system under analysis. For our example PT reactions give enough flexibility to model a resilience scenario.

Additionally, this approach respects the linkage between the intensity of a disturbance and the intensity of its effects. The maximum performance loss, sometimes called vulnerability (Häring et al. 2016), is defined by the gain K . Hence, the loss of resilience, cf. Equation (3), depends on the intensity of the disturbance.

The decay and recovery in performance occur in Fig. 2 with the same function. However, in most real cases, the decay happens much faster than the repair process. While the breakdown occurs quickly and may be best modelled with using a PT_1 reaction, repairing such complex equipment takes considerably longer. This may be better expressed

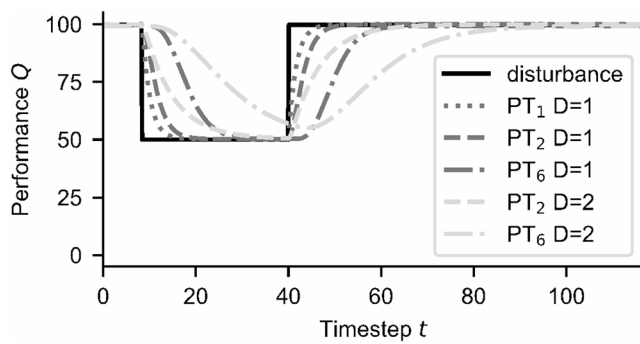


Fig. 2 Arbitrary resilience process modelled using various PT reactions to a sudden disturbance. The gain parameter K is 1 for every reaction depicted here

with a higher order like PT_6 or a higher damping factor D . Therefore, the system reaction should be modelled with respect to the direction.

Another use of PT behavior is the activation of a backup system, as redundancy and slacks can help increasing resilience (Jackson and Ferris 2013). When the main system collapses, the overall system performance depends on how fast the backup can be activated. In that case the down-time is only depended on the time the backup is activated and the time it takes to activate the system.

The objective here is not to construct a model that achieves realism by replicating the full complexity of real-world infrastructure networks. Rather, by using a well-understood, relatively tractable model class, which offer reasonable flexibility to cover various systemic operating scenarios, we demonstrate that the rebound curve and its associated resilience metrics may fall short even under favorable circumstances.

The significance of this exercise lies in demonstrating that even within this controlled setting, the theoretical framework produces dynamics that fundamentally confound its intended application. If such shortcomings emerge under these simplified conditions, it necessarily follows that more complex, nonlinear, and human-influenced real-world systems would expose them with even greater intensity, with yet more opaque and dangerous effects.

2.2 Resilience quantification

In addition to the mathematical model required for the generation of a performance curve, it is necessary to have a resilience measure for quantitative comparison of failure dynamics and system design. One of the first and most popular ways to quantify resilience was proposed by Bruneau et al. (2003). It uses the performance (or quality) $Q(t)$ to describe the loss of performance between t_0 and t_1 and calls it resilience loss:

$$RL = \int_{t_0}^{t_1} [100 - Q(t)] dt \quad (3)$$

There are many other definitions for resilience quantification metrics. Structured literature reviews regarding this topic are provided by Yodo and Wang (2016), Poulin and Kane (2021) and Tang et al. (2023).

Together, our mathematical model and measure of resilience allow us to determine whether the performance curve is an effective way to assess the resilience of a system. We can define situations when the curve “works as intended” and provides informative comparisons among system operations and designs, or when it “fails” by obfuscating information that is relevant for operational and design decisions. Specifically, the curve works as intended when three key conditions are met:

- i. Recovery dynamics are monotonic and return to a well-defined target state.
- ii. The system exhibits a simple, single-path structure without dynamic slack or conditional switching.
- iii. Performance is measured at the relevant level of resolution.

In contrast, the model does not work and fails if:

- i. It does not model recovery dynamics effectively.
- ii. There is complex behavior that confuses results when dynamic slack or switching activities are introduced.
- iii. Performance measured via Eq. (3) does not inform resilience or hides relevant operational and design information.

We demonstrate resilience loss and quantification with Eq. (3) “works as intended” given our model of generic infrastructure system and an outage of one subsystem. The infrastructure is modelled using various PT behaviors, which is commonly used to describe delayed systems (Lunze 1992). The resilience performance curve needs a target level (Ayyub 2014), mostly (e.g. Bruneau et al. 2003; Cimellaro et al. 2010) considered as 100% performance level (see Fig. 3). At some point the infrastructure fails which results in a performance drop, and later it is repaired so that the performance resumes at the target level.

At such a simple level, the resilience performance curve fulfils its needs. We see that the performance decreases and raises after a disruption. If we consider a single system, this chart shows when the output decreases and it is possible to repair or replace it. This is very straightforward and is, independently of resilience, common practice for decades. Additionally, this presents only the overall output. Therefore, it is still to check which one of the subsystems failed. As long

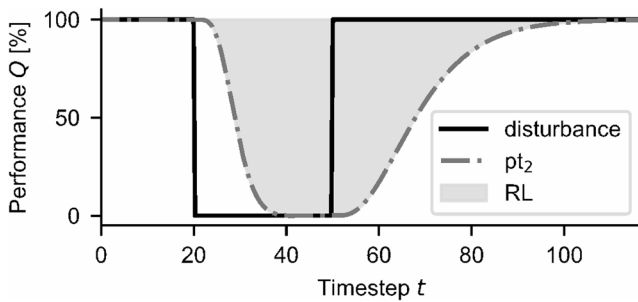


Fig. 3 Performance chart of an infrastructure. Whenever a failure occurs, the output drops slowly to zero. After repairs are completed, production gradually resumes until it reaches its previous performance level. The area between the curve and the target level marks the loss of resilience (RL)

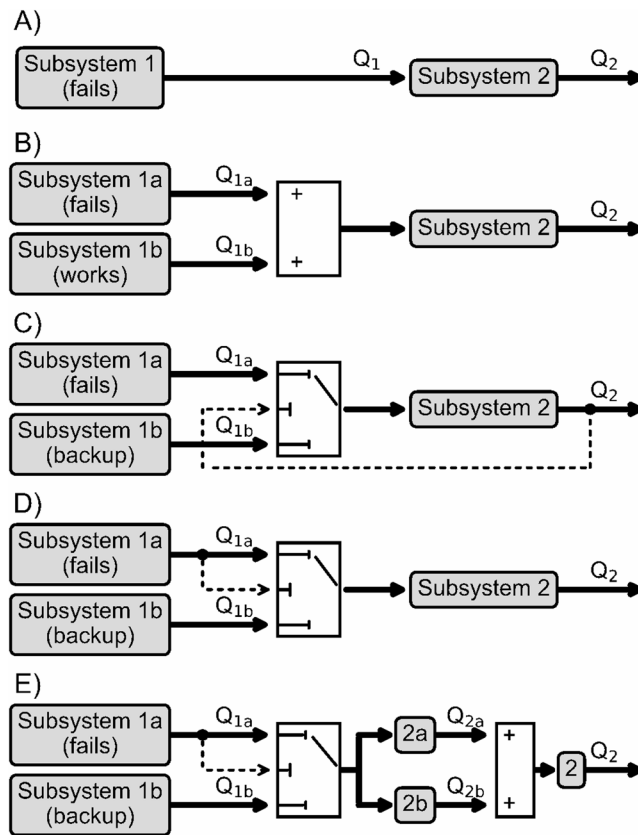


Fig. 4 The considered Cases. (A): Two-stage flow-shop. The failure of the first element simply cascades to the second and third. (B): Again, a two-stage flow-shop. Now Subsystem 1a works in parallel to Subsystem 1b, the output is summed up and equally distributed to the following Subsystem 2. (C): Subsystem 1b is activated and used only if Subsystem 2's output (Q_2) is below 90%. (D): Similar configuration as in (C), but the output is measured at Subsystem 1. (E): Same configuration as in D, but Subsystem 2 is now further decomposed into Sub-Subsystems

as the system is not too large, investigable and not a “black box” this might be manageable, but it would be favorable to have this chart for every subsystem instead of the whole infrastructure, so that it is easier to locate the failure.

3 Case studies

We consider five cases of infrastructure failure and recovery with a generic two-stage flow-shop layout (see Fig. 4) via our PT system performance curve model. The selected flow-shop layout, purposefully simple, involves series or parallel subsystems with dependent inputs and outputs.

Many critical infrastructure systems can be modelled this way, including delivery of energy products, critical chemicals production, healthcare treatments, Oil and Gas plants, and government services. A flow-shop model can also represent large-scale, time delayed assets working to provide a single integrated service, such as dam and waterway management. Although the case studies are intended to be broadly applicable, we illustrate them as a water-treatment plant to improve readability.

The benefit of adopting a flow-shop layout is to show the interplay between overall system resilience with subsystem performance, dynamics, and resilience metrics. In our setup, the output of every subsystem (n) is measured using its performance quality Q_n , which can be considered a different critical service necessary for meeting a final demand. Hence, the output of Subsystem 2 (Q_2) is equivalent to the overall system output representing final supply-demand matching. In many studies, only Subsystem 2 would be presented for assessing system resilience. We develop five case studies that highlight different aspects that this narrow view on Q_2 fails to capture. The four cases are as follows (Fig. 4).

In Case A) the failure of the first subsystem cascades directly to the subsequent stages, disrupting the entire flow. This is comparable to failure at a single water treatment plant, where the downstream distribution network loses supply with a delay as storage tanks drain. After repairs, pressure restores gradually.

In Case B), two subsystems (Subsystem 1a and Subsystem 1b) work in parallel at the beginning, with their combined output being distributed evenly to the next stage. This parallel structure enhances flexibility and throughput. This is similar to two treatment plants which supply the network in parallel. One fails, but the other keeps running, so end users see a reduced but not collapsed supply.

In Case C), Subsystem 1b only comes into play if Subsystem 2's output falls below a certain threshold. This changes Subsystem 1b to act as a backup system providing dynamic slack to critical services and ensures that resources are used efficiently while maintaining a continuous flow of production. In practice, this might be standby groundwater wells that remain idle until network pressure sensors register a drop below 90%. By the time the alarm is triggered and the wellfield comes online, storage reserves may already be substantially depleted.

In Case D), Subsystem 1b remains a backup but the trigger for its activation is based on Subsystem 1a, rather than the overall system performance metric Q_2 from Subsystem 2. This case shows how changes in sensing among interacting subsystems improves output resilience. This is like the backup wellfield is triggered by a flow sensor directly on the failing plant rather than waiting for pressure loss to reach the network. The wells activate almost immediately, and end-user supply barely dips.

In Case E), the general system is equal to Case D). Subsystem 2 is viewed in more detail, meaning Sub-Subsystems 2a and 2b, lead to the performance in Subsystem 2. This case shows that the more interconnected a system is, the more difficult it is to track down the failure and that it is important to zoom in and out into components. Here, the distribution network is split into its real components: a treatment and storage stage, and a pressurized pipe network. Both experience internal stress during the failure event, but the final supply curve barely moves.

All cases produce several performance curves, one for each subsystem, that can inform loss or gain of resilience according to Eq. 3. We also consider the effect of early or fast repair in all cases to show whether these common resilience decisions produce improved recovery. The parameter choices are to some extent arbitrary, serving primarily to highlight certain effects among the curves. The cases serve as conceptual illustrations of different behaviors in response to a disturbance.

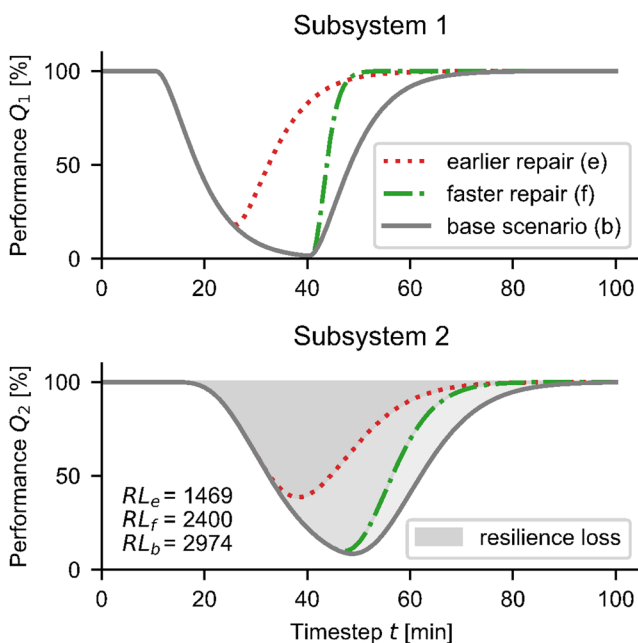


Fig. 5 Performance charts for Case A). At timestep $t=15$, Subsystem 1 experiences a failure. Subsystem 2 depends on the output of Subsystem 1, therefore the disruption propagates to Subsystem 2 with a delay. After Subsystem 1 is repaired, Subsystem 2 restarts production and resumes to its initial performance level

3.1 Case A)

In Case A), all subsystems are described as PT_2 and $T = 1$ within the LTI system. We modelled Subsystem 1 with $K = 1$, $D = 2.5$ and Subsystem 2 with $K = 1$, $D = 1.7$. The input of Subsystem 1 is blocked at $t = 15$, hence only the remaining service flows, and this propagates to Subsystem 2 (Fig. 5). In the base scenario, as well as in scenario “faster repair”, the issue is fixed at $t = 40$ and the system restarts. Within scenario “earlier repair” the issue is already fixed at $t = 30$. For the scenario “faster repair”, Subsystem 1 is modelled direction-dependent. The performance decline follows the behavior described before and the gain is modelled with $K = 1$, $D = 0.9$.

This initial case best relates to common uses of performance curves for measuring resilience. Based on these results, the system would recover performance faster given earlier repair vs. faster repair. This is captured in the resilience metric (Eq. 3), as the area under the curve is smaller for earlier repair. A key element of this case is that each subsystem experiences nearly the same dynamics in terms of performance loss and gain with and without repair strategies. Calculating the resilience metric for any subsystem would produce the same comparison between early and fast repair. Thus, focusing decisions only on output variable Q_2 appears sufficient for decision-making, i.e., is equivalent to using any subsystem for decision-making without loss of generality.

3.2 Case B)

Unlike Case A), most critical infrastructure systems are more complex with multiple service providers within subsystem layers prior to final demand.

Here, all subsystems are again described as PT_2 with $T = 1$. Subsystem 1a is equivalent to Subsystem 1 from Case A). The only difference for this case is that Subsystem 1b works in parallel and is independent of disturbance in Subsystem 1a. The sum of the output is distributed equally to Subsystem 2, which is again modelled with $K = 1$, $D = 1.7$. Like Case A), the input of Subsystem 1 is blocked at $t = 15$ and fixed by $t = 40$. Repair in Subsystem 1a is modelled the same as Case A) (Fig. 6).

The functional redundancy in Subsystem 1 in this case appears to make the system more resilient, as commonly theorized within the socioecological systems literature (Biggs et al. 2020; Yu et al. 2020) which highlight the importance of functional diversity. Indeed, the bathtub formed under the performance curve for Subsystem 2 is smaller than in Case A. However, it takes considerably longer to return to full functionality.

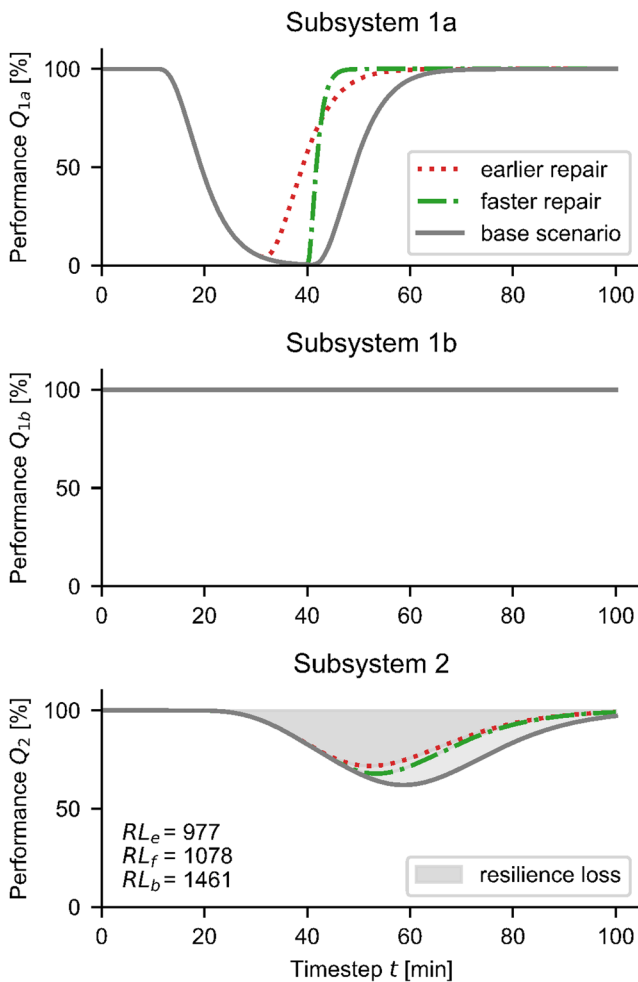


Fig. 6 Performance charts B) of a three-step infrastructure system, where Subsystem 1 and 1b provide parallel inputs to Subsystem 2. At timestep $t=15$, Subsystem 1 fails and the disruption propagates to Subsystem 2, which is the overall factory output

3.3 Case C)

Case C) retains much the same structure and parametrization of Case B), but now Subsystem 1b is a backup for Subsystem 1a. This backup is designed to activate only when the output of Subsystem 2 drops below 90% performance. All other subsystems are the same. Hence, this case changes the view of resilience from functional redundancy to dynamic slack via backups, which is more aligned with resilience design from the resilience engineering literature (Saurin 2015; Saurin and Werle 2017).

The disturbance happens at $t=15$, the alarm triggers at $t=25$ and the backup system starts to work. Subsystems 2 and 3 continue to experience performance drops even after $t=25$ due to the time it takes for the backup to reach its full performance level. Once the backup is fully functional, all systems begin to recover. Subsystem 1a repairs at $t=30$ or $t=40$ depending on strategy. However, there is

negligible effect on overall system performance because the backup is in place anyways. Faster repair may even be slightly favored for this specific case, but from a performance view resilience is worse than in Case B).

This case introduces additional confounding elements that make performance curves lose salience for resilience. First, Subsystem 1b cannot have a normal resilience measure, as its performance only increases in response to failure. This captures the notion of dynamic slack from resilience engineering literature but suggests resilience metrics, cf. Equation (3), do not help with its analysis. For example, if the backup failed, it would not be captured in the performance curve for Subsystem 1b.

An important dimension for comparing Case B) and Case C) is linked to the overall system operations and cost, which are not given within the curves. While resilience may be improved with functional redundancy, it also implies that backup systems are on and working when failure occurs. In practice, critical infrastructure systems are not incentivized to operate when not needed, as this incurs significant cost and resource use that is not necessary during normal operations. Moreover, failure diagnosis is easier in Case C) over Case B), as Subsystem 1b is a backup meaning failure must have happened in Subsystem 1a, which is not obvious in Case B). This suggests that remedial actions and potential system upgrades would be easier to implement in the design for Case C), even if the resilience of the system is reduced. Hence, any effort after the event, either to operate, diagnose, or upgrade systems, might be easier and better in the less resilient system.

3.4 Case D)

Case D) uses the same system structure as Case C), but the trigger for backup activation is informed by Subsystem 1a, rather than overall system performance Q_2 . Hence, the backup activates several timesteps earlier, almost immediately ($t=16$), resulting in a better system performance and less loss of resilience (Fig. 7).

In this case, the same issues as Case C) arise, but they are even more pronounced. From a performance curve and resilience metric perspective, Case D) appears more resilient than previous cases – the bathtub formed for Subsystem 2 is much more pronounced than in Case C) and B). Moreover, the backup system is used for a shorter time (until $t=60$) compared to Case C) (Fig. 7), which might result in lower operating and maintenance costs. Yet, the instability produced in Subsystem 2b is more pronounced and the difference between repair strategies is further confused. All repair strategies experience longer instability and multiple performance drops.

This case further emphasizes that a focus on final system performance is inadequate for resilience. As this case is

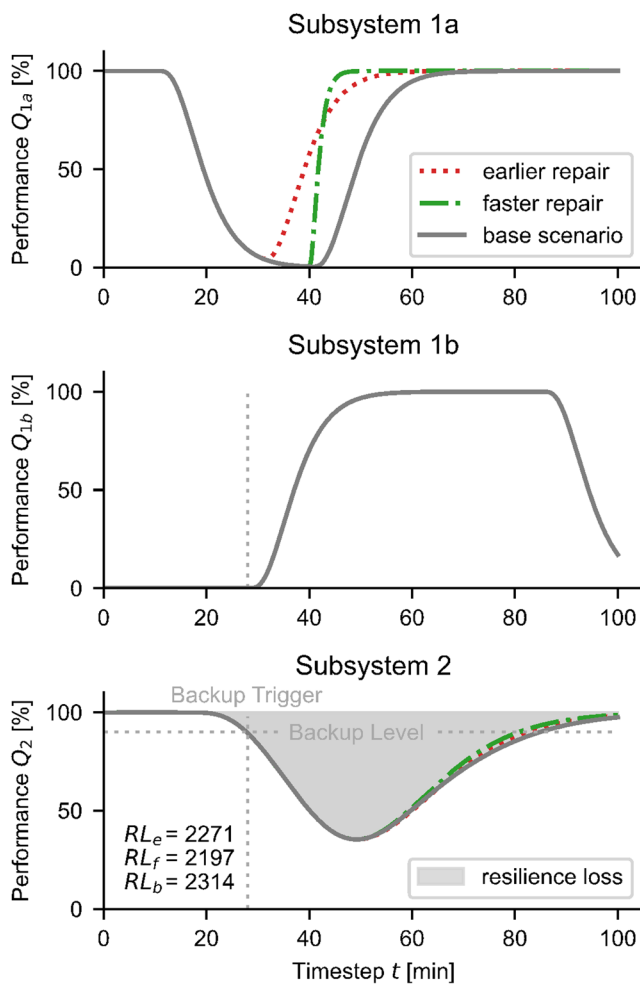


Fig. 7 Case C): Activation of a backup system when the overall system performance (Subsystem 2) drops below 90%. Subsystems 2 takes the input of 1a if it is above 90%, otherwise 1b

most resilient compared to previous, it shows that appropriate sensing influences the effectiveness of dynamic slack. Sensing performance from the overall system perspective (Case C)) introduces too large of a delay in response, despite this being the most likely scenario from a critical infrastructure system perspective. In contrast, linking backups to their direct performance creates more resilient systems, but hides how difficult upstream Subsystems may be operating to ensure final demand is met.

3.5 Case E_i)

Case E) uses the same system structure as Case D), but Subsystem 2 is now split up into the Sub-Subsystems 2a and 2b. However, this small change shows how difficult it is to diagnose the error by looking only at the overall system output Q_2 . For example, the performance of Subsystem 2a and 2b both experience service disruption, which hides where system failure occurred from the overall perspective

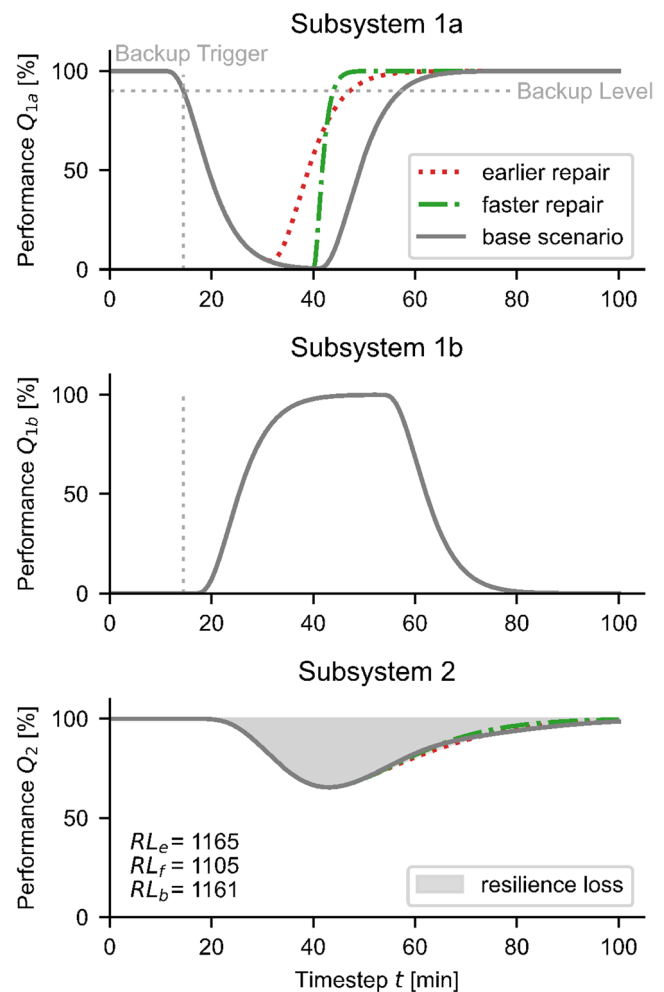


Fig. 8 Case D): Activation of a backup system when the Subsystem 1's performance drops below 80%

of Subsystem 2. It also hides which repair strategy may have been better – the difference in resilience between repair strategies is negligible for Q_2 . Hence, efforts to make the system more resilient are successful from a metric perspective, but introduce significant confounding elements that make future response to failure much more difficult.

An important consequence of hiding the failure and its effects, is that individual subsystems appear less important. This is, of course, not correct, as the total service provision for the full flow-shop requires all subsystems to function. Yet, as “improved” resilience ensures milder consequences for Subsystem 2, then Subsystem 1 may change its status from “critical” (in Case A) to less important. In practice, this change in perspective on asset importance is often influenced by faster, better, and cheaper pressures (Woods 2018a) that eventually lead to underfunding and poor maintenance for previously well-maintained systems. These pressures are common among many critical infrastructure systems that are optimized to remove unnecessary backups

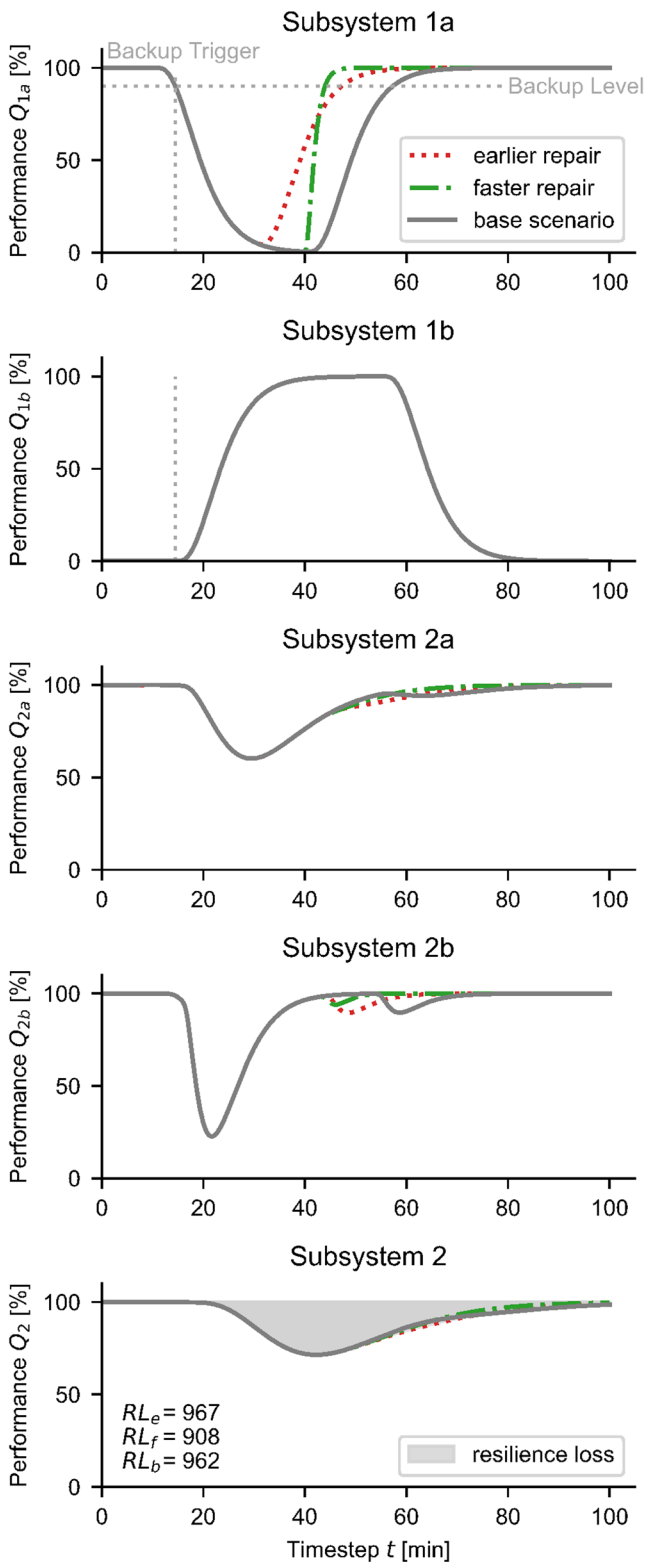


Fig. 9 Extended performance charts compared to Fig. 8. Subsystem 2 consist of Sub-Subsystems 2a and 2b

and to decrease any delays within a system. This leads to

more severe and immediate reactions to the whole system when subsystems fail.

3.6 Case E₂)

In case E₂) we consider the effect of multiple disturbances with the backup system structure from Case E₁). The purpose of this case is to provide even more realistic estimate of resilience given that most system failures occur due to delayed and compounding threats, rather a single event with a single repair process.

In this Case, the subsystems at the first stage are now prone to additional external disturbances and the backup is now targeted to provide 120% system performance. The backup is triggered already at $t=5$, even though this results only from a small system fluctuation. However, coincidental it becomes useful at $t=25$ when the real disturbance happens. As a result of the overuse of system performance the backup system reduces to a 60% performance level at $t=35$ and switches of once the disturbance is past ($t=58$) (Fig. 10).

All this results in a less constant system performance within the second stage systems. However, Subsystem 2 shows an almost constant system output:

This more realistic case introduces many confounding elements that question the usefulness of performance curves and the resilience metric (Eq. (3)). First, the resilience measure for this case is the smallest for all, suggesting this design is most resilient. This is in lieu of the significant instability in the system due to dynamic slack and its design as its activation point from Q_{1a} . Because the backup activates for a minor disturbance, rather than the major failure event, all downstream subsystems become unstable and performance fluctuates significantly. Moreover, we see performance for Subsystem 2 exceed 100%, which further confounds measuring resilience via Eq. 3. In other words, without detailed analysis of the performance curves themselves, the common resilience measure for critical infrastructure systems would prioritize designs that increase subsystem instability. Practically, this may be the worst scenario from a critical infrastructure systems operations and management perspective, even if the demand appears resilient.

3.7 Sensitivity analysis

The resilience loss RL is often the result of the overall system assessment, which means this value depends on model parameters across subsystems that may be uncertain or different from those considered here. Hence, we perform sensitivity analysis to view the impact of model parametrization on system dynamics.

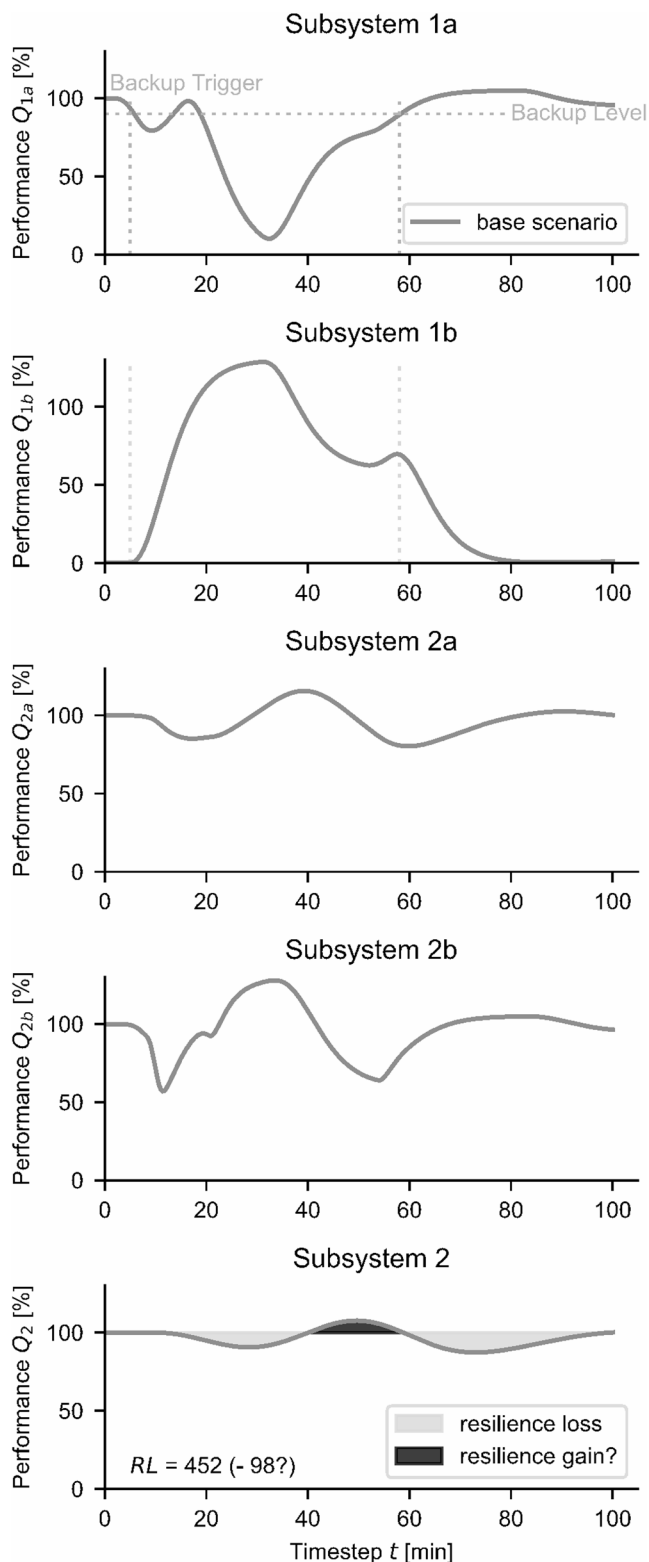


Fig. 10 More realistic performance charts (Case E_2). The structure remains the same as in Fig. 9, but the Subsystems 1a and 1b deviate slightly due to other external factors. This results in an almost constant performance in Subsystem 2

In general, across all cases, we find conclusions robust to

diverse parameter combinations. For example, conclusions regarding RL, earlier repair, or faster repair or for Case A remain similar across tested scenarios.

We present sensitivity analysis for our most complex case, E_2 to demonstrate this fact. Here, we adjust the transfer functions between systems and measure the overall system output (see Fig. 11). Specifically, we vary the transfer function parameters T_1 and D_1 of Subsystems 1a and 1b and T_2 and D_2 of Subsystems 2a and 2b by ± 0.2 (which is 20% for Subsystem 1 and 10% for Subsystem 2). This analysis results in 81 (3^4) simulation runs, of which the general dynamics, and resulting conclusions, remain the same.

However, choice of transfer function parameters does impact system performance measured. Figure 12 shows the quantified loss of resilience (RL) varies considerably between the simulation runs. The mean RL for all simulation runs is 357. The variance of G_1 is 5,233 (standard deviation of 72, relative standard deviation of 20%), whereas the variance of G_2 is 38 (standard deviation of 6, relative standard deviation of 1%). Although these standard deviations are usually classified as high, they can certainly occur in the context of resilience analyses due to the large number of uncertainties and influencing factors that are included in the assessment.

We further demonstrate that for specific choice of system parameters, the calculated RL can be different than that quantified in our given cases. For example, for Case D, we analyze the relationship between RL and choice of the backup trigger level (Fig. 11). As expected, the RL for the system decreases as the backup trigger level % increases. This reduces the effects of system delays in terms of utilizing dynamic slack to replace primary system function (Fig. 13).

Overall, we do not find that these variations output (e.g., total performance loss measured) to change resulting dynamics or interpretation of the system performance function. Accordingly, our results should produce robust conclusions as to when the model “works” as intended,” or, “does not work” as intended.

4 Discussion

Overall, our PT_n approach is representative of many found in the literature (see e.g. Ayyub 2014; Fig. 3; Mentges et al. 2023; Fig. 1; Poulin and Kane 2021), but shows when the curve and resilience measure hide the background processes that inform system resilience alongside situations where the method seems to work as intended. Only when the system structure is simple (i.e., no dynamic slack, sensing, or otherwise) did the resulting curves have some explanatory power for resilience. When the system had added complexity that

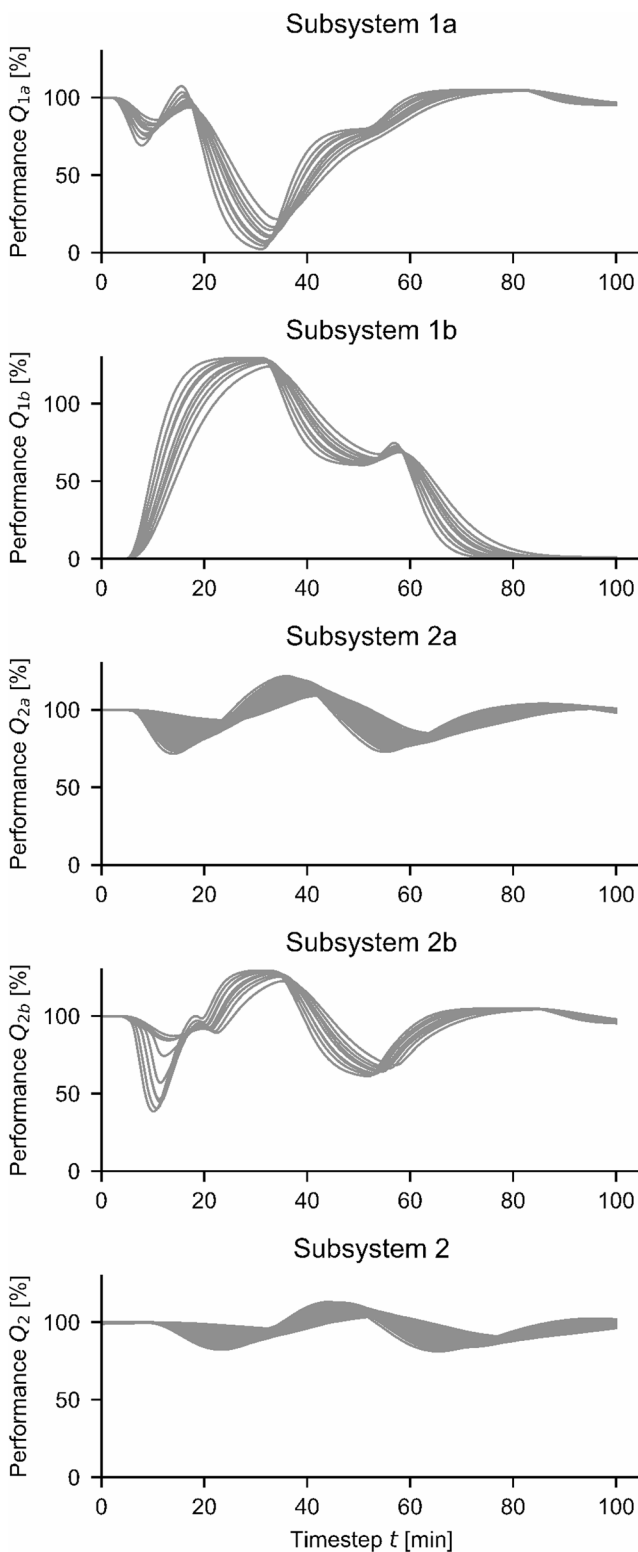


Fig. 11 Sensitivity analysis for Case E₂ (81 simulation runs)

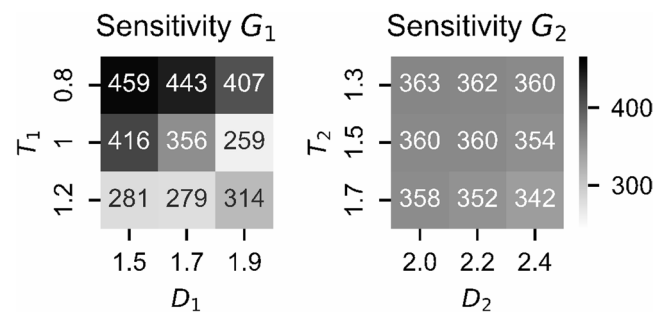


Fig. 12 Resilience Loss sensitivity analysis for varying transfer function parameters (+/- 0.2)

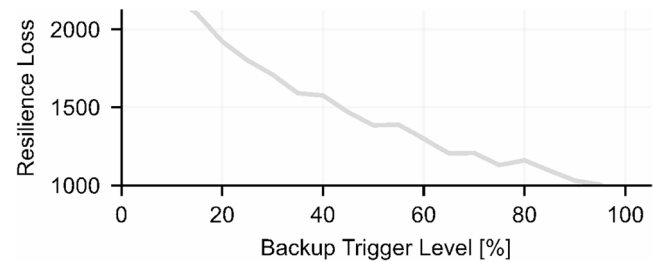


Fig. 13 Sensitivity analysis for the trigger threshold in Case D)

matches real-world infrastructure, both the resilience measure and dynamics confused interpretation of the event and/or any benefits of design.

Naturally, the PT_n approach is another, admittedly imperfect, form of abstraction. Its sole purpose is to provide a generic representation of behavior that aligns with the observed performance curves in the resilience literature.

4.1 Ambiguities and assumptions in the system performance curve

Table 1 summarizes key results from our case study and system performance curves. In general, we find some benefits of using performance curves in all cases. These benefits range from providing comparative power between different system designs (e.g., earlier/faster recovery, system sensing) and revealing the effects of decision delays.

We also find issues with the system performance curve in most cases. These issues range from an inability to diagnose when and how failures occur to an inability to handle irregular recovery dynamics. In general, we find that these issues arise from:

1. **Limitations of using aggregate system metrics for complex dynamic systems:** Simple performance curves use aggregate system outputs without subsystem-level diagnostics that mask variability among subsystems. Failures, bottlenecks, vulnerabilities, or compensatory mechanisms may remain hidden, reducing the ability to

Table 1 Positive and negative aspects of the performance curve for resilience analysis

Case	The curve allows...	The curve lacks...
A)	...defining when and where failure occur ...measurement of certain “regular” resilience patterns like earlier/ faster recovery or drop/ recovery intensity	...nothing
B)	...comparing the resilience of a system with functional redundancy (multiple sources), to one with a single source ...deciding where resilience investments might have the highest leverage	...ability to diagnose where upstream failures occurred and how that may impact final services ...meaningful comparison of reduced impacts vs. longer or delayed recovery
C)	...comparing the resilience of dynamic slack versus other methods (recovery, redundancy)	...relevant measures for sub-systems that surge or extend performance in response to failure (what if the backup fails?). ...consideration of increased costs for system operation given changes in upstream function (i.e., backups are expensive compared to normal operations).
D)	...comparison of system sensing – backups that trigger more quickly provide more resilience ...interplay between dynamic slack and repair strategies.	...ability to identify where the failure occurs and where performance is measured. Final system performance is entirely an outcome from upstream decisions.
E ₁)	...visualizing delays within the system or the activation of a backup source	...ability to disaggregate the effects of upstream subsystems and dynamics. ...processes that underlie resilience ...ability to make sense of non-regular dynamics that classical resilience measures cannot handle
E ₂)	...tracking the performance difference between similar functional structures	...mechanism to handle compounding events and/or subsystem failures that do not match regular recovery dynamics

diagnose why resilience was gained or lost and limiting the usefulness.

2. **Area-under-the-curve metrics:** Metrics based on integrating the area-under-the-curve assume regular dynamics. However, many infrastructure systems experience irregular dynamics, especially when utilizing dynamic backup systems or with multiple subsystems and performance measures.
3. **The performance curve itself:** Furthermore, using a single indicator like resilience loss (RL) potentially obscures temporal characteristics such as the timing,

duration, cost, severity, and shape of disruption and recovery. Very different performance progressions may lead to a very similar RL, making it difficult to distinguish between them.

Case A) covers two important dimensions of resilience: when and where (which subsystem) the failure occurs. These dimensions help answer a common question regarding infrastructure and community systems: “resilience of what, to what?” (Cutter 2016). Hence, our model has enough detail to examine how a flow-shop infrastructure system may deal with a real outage.

Case A) also shows that the resilience measure (Eq. 3) provides explanatory power when failure dynamics are “regular”, i.e., recovery only to previous performance levels. Results align with literature that recommend having early warning systems that enable quicker response for infrastructure resilience. Faster repair also increases the resilience compared to its original repair speed even with delayed onset. This recommendation may not hold for other similar mathematical models, nor when the resilience decision is not recovery-focused, e.g., efforts to avoid the failure altogether.

However, the resilience measure shows issues when recovery dynamics are not “regular.” The measure is based on the assumption that there is a “steady” or “target” state Q^* to which a system returns after a disturbance (Ayyub 2014). Many systems, especially social or ecological systems, have not a single, fixed state. Systems might be in a state of constant flux, which is recognized by the resilience landscape with multiple equilibria (Walker et al. 2004). To acknowledge this, it may be necessary to compare the current performance to a former performance that was considered to be “normal” (for different target performances see Poulin and Kane 2021). In our infrastructure example, the production could vary over time, not because a subsystem fails, but because of seasonality or other external effects. Therefore, it is difficult to define exactly what 100% performance is.

Comparing Cases C) and D), we see that the location and frequency of measurement is very important for the resilience of a system. When a measurement is only possible at the final subsystem, it might take considerably longer to implement countermeasures resulting in a higher loss of resilience.

It follows, that the idea of a resilience cycle is very ambiguous. In Case E₂) it shows that it can become very unclear where to set the boundaries that define one resilience cycle for the curve, not only in terms of the loss of resilience, but also the rapidity of recovery.

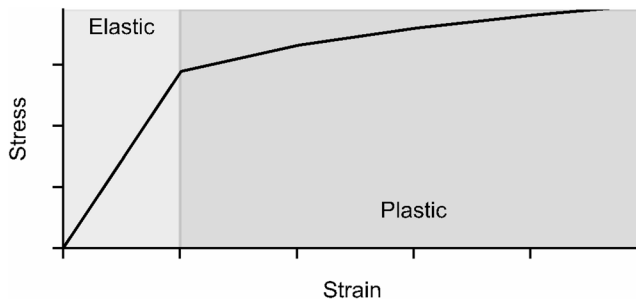


Fig. 14 The stress-strain analogy to resilience is an example where the disturbance intensity (the stress) is in relation to the effects (the strain) of it. Interestingly, at a certain stress level these curves become always non-linear. Therefore, any knowledge of the elastic behavior does not help to predict the plastic behavior

4.2 Why disturbance intensity and system functional structure matter

The performance curve is always depicting a reaction to a disturbance, but very often it does not relate the effects of a disruptive event to any of the events characteristics (Gama Dessavre et al. 2016). Therefore, a very light and a very strong attack to a system may have the same effect and become indistinguishable. For example, unplugging a machine cuts its power and shuts it down, but this can be easily reversed. Cutting the cable also shuts it down, but may be irreversible. This is not observable from the presented figures, and therefore out of scope of the resilience assessment. However, it should be important to look at the root-cause of incidents in order to learn and prevent them the next time.

Systems may in general respond differently to minor versus catastrophic disturbances, and the intensity can shape both the response and recovery trajectory. Small, frequent disturbances may be easily absorbed, but a sudden disruption may lead to a completely new alignment of the infrastructure. Intensity is also not only about the magnitude of one event but how disturbances accumulate over time. A series of small shocks (like machine part failures, or natural degradation) can erode resilience even if each shock is not intense on its own. Conversely, a single catastrophic event (like an electrical blackout) may expose underlying weaknesses but also offer an opportunity for systemic change and improvement. Understanding the disturbance intensity may help to identify thresholds beyond which a system may no longer recover in its original form (Connelly et al. 2017). Interventions need to account for varying intensities of disturbances and recognize that the system might need to transform rather than just recover after a high-intensity event (Walker et al. 2004). Additionally, the loss of performance in complex systems may very likely not be linear (Hollnagel et al. 2006). Just because a certain level of stress intensity

leads to a specific loss of resilience, it does not necessarily help drawing conclusions about the effects of higher or lower intensities. Actually, this is exactly what the stress-strain diagrams (Fig. 14), sometimes related to resilience (Woods and Wreathall 2008) state.

In that sense, it seems to be almost useless to monitor the performance curves of very different scenarios or between different systems because this would be only measuring one point at a certain stress level on the stress-strain diagram. Even if it is possible to measure the “strain” at multiple “stress” levels, this analogy assumes that systems always behave the same (time-invariant), which is certainly not the case for complex systems with humans involved. Moreover, the diagrams represent the material’s properties under the assumption that no fatigue processes have occurred. Repeated loading, material fatigue, or other factors can change the original mechanical properties, exactly what happens to organizational systems. The stress-strain diagrams do not provide information about the behavior under cyclic loading, which is why evaluating fatigue strength is crucial when selecting materials for long-term applications. Therefore, similar is necessary for a resilience assessment, even though we have to admit that it seems impossible to collect such data.

Besides looking at only a resulting performance, it is more important to understand the functional properties of a system, including its components, the relationships between them, and the processes that drive its behavior. By examining these elements, we gain insight into how different parts of the system interact, where inefficiencies or vulnerabilities might lie, and how changes in one area can impact the whole. We did this in the example by plotting the flows (Fig. 4). Furthermore, it is also important to understand the feedback loops and emergent behaviors that may not be immediately obvious. A deep knowledge of a system’s structure enables us to anticipate unintended consequences, improve resilience, and make informed decisions that lead to sustainable improvements. For that issue the System Thoretic Accident Model and Processes (Leveson 2004) or Functional Resonance Analysis Method (Hollnagel 2012), when exercised properly, give a good insight into the structure.

The example also examined that systems should be designed with buffering capacity (or delays) to absorb more intense shocks or adapt through redundancy, diversity, and flexibility. The delay may also act as a stabilizer that helps the system distinguish real signals from noise, thereby enabling graceful degradation (Woods 2018b), meaning slowing the pace of stress so that the system has time to adjust, repair, or activate backups. In these ways, well-placed delays help complex systems remain stable, adaptable, and robust under changing conditions.

It is important to remind the level of detail or the scale at which a system is observed or analyzed. The resolution of a model determines how fine-tuned the analysis is. The performance curve is too often depicted at a macroscopic level, which leads to fail to account for variations at a more granular level. This disconnect is critical because, as we showed with the example, resilience can manifest differently depending on the granularity of the analysis. A system might appear resilient when viewed at a high level, but its individual components may be fragile (Carlson and Doyle 2002). Understanding resilience requires a systemic view, considering both the aggregated behavior of the system and the resilience of its individual components. The second one is often harder to assess, but equally important. At a lower level, it should be easier to identify vulnerabilities and areas for improvement. Looking at finer scales can reveal where vulnerabilities exist that prevent adaptive capacity. This is especially the case in complex, non-linear systems, that are so often addressed within the field of resilience engineering.

When modelling systems, boundaries define the scope of the system and determine what is considered inside the system and what is external to it. Defining system boundaries is not always easy (Lowe et al. 2024) but crucial when discussing resilience because the interactions between a system and its environment are integral to its ability to absorb shocks and adapt (Hochrainer-Stigler et al. 2025). Resilience models that do not clearly define system boundaries often ignore how external factors affect the system. For example, an organization's resilience might not only depend on its internal processes (like the subsystems presented in the example) but also on external economic conditions, regulatory changes, and social movements that might be beyond the organization's control. Especially in social and ecological systems, boundaries are porous or flexible with interactions crossing over between different systems (e.g., between human and natural systems). A rigid boundary may limit a system's ability to adapt to changes that affect both the system and its external environment. The guidance is to define narrow boundaries first (for example the area of the infrastructure) to find interventions to the most impactful areas. By expanding the boundaries of analysis afterwards, external influences and the system's adaptability to changes that might not be internal, but are just as crucial to resilience.

Therefore, to become resilient backup solutions must be available at the level that you can measure (or repair): for example, if you can only measure the performance of the whole flow chain, you better have at least a second flow chain ready. If you can measure the performance of one subsystem, you should have one backup subsystem ready. If you can measure every screw within the subsystem, you might only need backup screws.

4.3 Why humans and decisions matter even more

We did not consider any humans in the example. However, they play a critical role in determining the resilience of any system, from organizational structures to ecological systems. In most systems, human actions are not just the cause of disruptions (e.g., handling error on a subsystem) but also a source of recovery (fixing simple errors or finding creative solutions). This makes the human element essential for both responding to and preventing future shocks. Moreover, the resilience of a company after a crisis might depend more on the leadership and decision-making processes than on any inherent qualities of the product or system itself. Therefore, engaging people in systems thinking and resilience-building is essential as it directly influences the system's ability to adapt. Training, policy interventions, and decision-making frameworks can help improve the resilience of human-influenced systems. However, it is hardly possible to describe the system dynamics as soon as humans are involved. The best we can do is to develop indicator-based assessments (Lee et al. 2013). Only if the system dynamics are clearly modelled, it is possible to use quantification metrics like (Bruneau et al. 2003; Ingrisich and Bahn 2018; Nan and Sansavini 2017) etc. The discussion on the use of such quantification metrics on top of the curve would deserve a dedicated analysis and a corresponding writing exercise. In this paper, we showed only roughly that a sensitivity analysis is necessary to provide a better insight into the system (Fig. 11).

Ultimately, resilience is also a trade-off between costs and benefits (Ale et al. 2021; Hanley 1998). For a private company, in terms of business continuity, it may be economically reasonable to accept the risk of failures. However, resilience is frequently of large interest within the context of critical infrastructures or safety-critical processes, where the potential costs of a failure exceed the costs of any preventive measures.

As modern technology evolves, it is becoming increasingly important to infer the performance and condition of individual subsystems from the behavior of the overarching system. A growing number of technologies are effectively turning into black boxes (Hill and Slater 2024). Devices are shrinking, components are tightly integrated, and embedded or certified electrical systems are often impossible to repair. Instead of being serviced, they are simply replaced. This trend is most visible in complex fields such as AI systems, where internal processes may be opaque by design, making direct inspection difficult (Sadeghi et al. 2024). As a result, understanding how a subsystem performs requires careful analysis of the system's overall outputs, interactions, and observable characteristics.

All of these factors (and maybe more) lead to a system reaction in the shape of the famous curves. We showed this

in various examples and cases. However, it does not work the other way around. The performance curve can hardly be used to compare the effects of different disruptive events and to compare different systems or different decisions within a system. The solution to this problem is about accepting an additional layer of complexity, with a rebuttal of simplification at all costs. Analysts need multiple perspectives that may or may not be aggregated into one metric. They need the flexibility to zoom in and out of a system and to view it from all of its sides. It is necessary to look at the infrastructure as a whole, but also at the subsystems.

Saying that does not mean that it is not important to measure and observe, it is more that it does not help directly just to measure. It is clearly an iterative process of measuring, changing and measuring again (for example as proposed by Simone et al. (2024), Fig. 1). A nice example for that is again the stress-strain plot analogy (Woods and Wreathall 2008). While it is definitely helpful and necessary to measure the material properties, the curves do not help to develop better (meaning more resilient) materials. Instead, the internal structure of a material must be considered and it requires creativity and exploration combined with knowledge in order to achieve a more suitable result. Projecting it back to organizational resilience, it is only possible after an event (or disturbance) has taken place to evaluate whether the preparations were helpful.

4.4 What should be done?

We recommend that resilience studies using performance curves should analyze results from both a system and subsystem level. Our case studies show repeatedly that only considering one of these perspectives conceals dynamics that are directly relevant to operational decisions.

Where backup systems or conditional switching are present, we recommend complementing the performance curve with a state diagram or simple event log that documents when and why each subsystem changed state. This makes the causal structure of recovery visible in a way the curve cannot.

For systems subject to compounding or repeated disturbances (as in Case E₂), we recommend that resilience assessment explicitly tracks remaining slack capacity (e.g., reserve storage, backup runtime) as a parallel indicator alongside the performance values. A system can appear highly resilient on the curve while exhausting the very resources that made it so.

5 Conclusion

The system performance curve is widely used model of infrastructure resilience despite its flaws. We developed our own using PT reactions and compared five cases that represent infrastructure systems and resilient designs as a flow-shop model. Our results are mixed, as there are benefits and some comparative power across the cases showing improvements and changes in system resilience that align with theory. However, if system dynamics were not regular for simple recovery and/ or system structure was more realistic with multiple upstream processes, switching, and backups, then the performance curve model and resilience measure hid salient information for decision-making.

These mixed results suggest a more nuanced understanding of the system performance curve as many design and operational factors are easy to overlook. Granularity, boundaries, human involvement, and disturbance intensity all bring their implications for improving system resilience. Additionally, in order to improve system dynamics, it is important to understand black-box systems, effectively turning them into grey, if not white, boxes. Incorporating these considerations into system models can lead to more robust, adaptive systems that are better equipped to handle both expected and unexpected challenges.

Resilience measurement therefore requires multiple method assessment approaches that capture perceptions, opinions, judgments and the nature of social interactions as well as the observable or easily measurable characteristics of sociotechnical systems. Measuring resilience means understanding the perspective of affected organizations and individuals, so analysis must include context-specific, qualitative and subjective information. As Lundberg et al. (2009) reminds us with the What-You-Look-For-Is-What-You-Find principle, an exclusive focus on performance curves inevitably constrains what can be observed and understood. Resilience cannot be reduced to a single metric or visual trajectory. It requires multi-scale, multi-method assessment that integrates structure, context, and human judgement. This is not confined to performance curves. Any modeling approach that attempts to quantify resilience will face similar challenges. The point is not to dismiss performance-curve analysis, it is more to emphasize that additional metrics and methods are required. Otherwise, we risk confusing observable recovery with genuine adaptive capacity.

Author contributions T.D. drafted the manuscript. D.E. and R.P. contributed to the interpretation of the results and critically revised the manuscript. All authors reviewed and approved the final manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL. T.D. was supported by the German Aerospace Center (DLR) through the internal project InterMoVe. D.A.E. was supported by the Defense Advanced Research Projects Agency (DARPA) Advanced

Research Concepts (ARC) Critical Orientation of Mathematics to Produce Advancements in Science and Security (COMPASS) program. This work represents the opinions of the authors and does not necessarily reflect the opinions of the U.S. Government and the U.S. Department of War or its components including DARPA, the U.S. Navy, and Naval Postgraduate School. R.P.'s work was partially supported by Sapienza University of Rome through the project HARMONISE (Human neurophysiological Assessment for Risk Modelling in complex socio-technical Systems) within "Bandi di Ateneo 2023" funding initiative.

Data availability No datasets were generated or analysed during the current study.

Declarations

Conflict of interest The authors declare no competing interests.

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

- Ale BJ, Hartford D, N. D., Slater DH (2021) Prevention, precaution and resilience: Are they worth the cost? *Saf Sci* 140:105271. <https://doi.org/10.1016/j.ssci.2021.105271>
- Aven T (2022) On Some foundational issues concerning the relationship between risk and resilience. *Risk Anal* 42(9):2062–2074. <https://doi.org/10.1111/risa.13848>
- Ayyub BM (2014) Systems resilience for multihazard environments: definition, metrics, and valuation for decision making. *Risk Analysis: Official Publication Soc Risk Anal* 34(2):340–355. <https://doi.org/10.1111/risa.12093>
- Ayyub BM (2015) Practical resilience metrics for planning, design, and decision making. *ASCE-ASME J Risk Uncertain Eng Syst Part a: Civil Eng* 1(3):Article04015008. <https://doi.org/10.1061/AJRUA6.0000826>
- Biggs CR, Yeager LA, Bolser DG, Bonsell C, Dichiera AM, Hou Z, Keyser SR, Khursigara AJ, Lu K, Muth AF, Negrete B, Erisman BE (2020) Does functional redundancy affect ecological stability and resilience? A review and meta-analysis. *Ecosphere* 11(7):e03184. <https://doi.org/10.1002/ecs2.3184>
- Bruneau M, Chang SE, Eguchi RT, Lee GC, O'Rourke TD, Reinhorn AM, Shinozuka M, Tierney K, Wallace WA, von Winterfeldt D (2003) A Framework to Quantitatively Assess and Enhance the Seismic Resilience of Communities. *Earthq Spectra* 19(4):733–752. <https://doi.org/10.1193/1.1623497>
- Carlson JM, Doyle J (2002) Complexity and robustness. *Proc Natl Acad Sci USA* 99(Suppl 1):2538–2545. <https://doi.org/10.1073/pnas.012582499>. *Suppl 1*
- Cimellaro GP, Reinhorn AM, Bruneau M (2010) Framework for analytical quantification of disaster resilience. *Eng Struct* 32(11):3639–3649. <https://doi.org/10.1016/j.engstruct.2010.08.008>
- Connelly EB, Allen CR, Hatfield K, Palma-Oliveira JM, Woods DD, Linkov I (2017) Features of resilience. 2194–5403 37(1):46–50. <https://doi.org/10.1007/s10669-017-9634-9>
- Cutter SL (2016) Resilience to What? Resilience for Whom? *Geographical J* 182(2):110–113. <https://doi.org/10.1111/geoj.12174>
- Cutter SL, Ahearn JA, Amadei B, Crawford P, Eide EA, Galloway GE, Goodchild MF, Kunreuther HC, Li-Vollmer M, Schoch-Spana M, Scrimshaw SC, Stanley EM, Whitney G, Zoback ML (2013) Disaster resilience: a national imperative. *Environment: Sci Policy Sustainable Dev* 55(2):25–29. <https://doi.org/10.1080/00139157.2013.768076>
- Demmer T, Kahlen J, Lichte D, Wolf K–D (2022) Towards the Prediction of Resilience; An Equation-based Resilience Representation. In M. C. Leva, E. Patelli, L. Podofillini, & S. Wilson (Eds.), *Book of Extended Abstracts for the 32nd European Safety and Reliability Conference* (pp. 1417–1424). Research Publishing Services. https://doi.org/10.3850/978-981-18-5183-4_R23-04-456-cd
- Eisenberg DA, Seager TP, Alderson DL (2025) The rebound curve is a poor model of resilience. *PNAS Nexus* 4(3):pgaf052. <https://doi.org/10.1093/pnasnexus/pgaf052>
- Gama Dessavre D, Ramirez-Marquez JE, Barker K (2016) Multi-dimensional approach to complex system resilience analysis. *Reliab Eng Syst Saf* 149:34–43. <https://doi.org/10.1016/j.res.2015.12.009>
- Hanley N (1998) Resilience in social and economic systems: a concept that fails the cost–benefit test? *Environ Dev Econ* 3(2):221–262. <https://doi.org/10.1017/S1355770X98250121>
- Häring I, Ebenhöch S, Stolz A (2016) Quantifying Resilience for Resilience Engineering of Socio Technical Systems. *Eur J Secur Res* 1(1):21–58. <https://doi.org/10.1007/s41125-015-0001-x>
- Hill R, Slater D (2024) Using a Metadata Approach to Extend the Functional Resonance Analysis Method to Model Quantitatively, Emergent Behaviours in Complex Systems. *Systems* 12(3):90. <https://doi.org/10.3390/systems12030090>
- Hochrainer-Stigler S, Mechler R, Roa H, Bachmann O, Šakić Trogrlić M, Handmer R, J., Dieckmann U (2025) Understanding multiple resilience dividends and system boundaries in disaster- and climate-risk management: a systems approach for enhanced decision-making. *Environ Res Lett* 20(4):44026. <https://doi.org/10.1088/1748-9326/adac7a>
- Hollnagel E (2012) *Fram: The Functional Resonance Analysis Method: Modelling complex socio-technical systems*. CRC. <https://www.taylorfrancis.com/books/mono/10.1201/9781315255071/fram-functional-resonance-analysis-method-erik-hollnagel> <https://doi.org/10.1201/9781315255071>
- Hollnagel E (2018) *Safety-II in practice: Developing the resilience potentials*. Routledge. <https://doi.org/10.4324/9781315201023>
- Hollnagel E, Woods DD, Leveson N (2006) *Resilience engineering: Concepts and precepts* (1st edition). <https://permalink.obvsg.at/>
- Hosseini S, Barker K, Ramirez-Marquez JE (2016) A review of definitions and measures of system resilience. *Reliab Eng Syst Saf* 145:47–61. <https://doi.org/10.1016/j.res.2015.08.006>
- Ingrisch J, Bahn M (2018) Towards a comparable quantification of resilience. *Trends Ecol Evol* 33(4):251–259. <https://doi.org/10.1016/j.tree.2018.01.013>
- Jackson S, Ferris TLJ (2013) Resilience principles for engineered systems. *Syst Eng* 16(2):152–164. <https://doi.org/10.1002/sys.21228>
- Klimek P, Varga J, Jovanovic AS, Székely Z (2019) Quantitative resilience assessment in emergency response reveals how organizations trade efficiency for redundancy. *Saf Sci* 113:404–414. <https://doi.org/10.1016/j.ssci.2018.12.017>
- Lee AV, Vargo J, Seville E (2013) Developing a tool to measure and compare organizations' resilience. *Nat Hazards Rev* 14(1):29–41. [https://doi.org/10.1061/\(ASCE\)NH.1527-6996.0000075](https://doi.org/10.1061/(ASCE)NH.1527-6996.0000075)

- Leveson NG (2004) A new accident model for engineering safer systems. *Saf Sci* 42(4):237–270. [https://doi.org/10.1016/S0925-7535\(03\)00047-X](https://doi.org/10.1016/S0925-7535(03)00047-X)
- Linkov I, Trump BD (2019) Resilience quantification and assessment. *Sci Pract Resil*. https://doi.org/10.1007/978-3-030-04565-4_6
- Linkov I, Bridges T, Creutzig F, Decker J, Fox-Lent C, Kröger W, Lambert JH, Levermann A, Montreuil B, Nathwani J, Nyer R, Renn O, Scharte B, Scheffler A, Schreurs M, Thiel-Clemen T (2014) Changing the resilience paradigm. *Nat Clim Change* 4(6):407–409. <https://doi.org/10.1038/nclimate2227>
- Lowe M, Bell S, Briggs J, McMillan E, Morley M, Grenfell M, Sweeting D, Whitten A, Jordan N (2024) A research-based, practice-relevant urban resilience framework for local government. *Local Environ* 29(7):886–901. <https://doi.org/10.1080/13549839.2024.2318571>
- Lundberg J, Rollenhagen C, Hollnagel E (2009) What-You-Look-For-Is-What-You-Find – The consequences of underlying accident models in eight accident investigation manuals. 09257535(10):1297–1311. <https://doi.org/10.1016/j.ssci.2009.01.004>
- Lunze J (1992) *Feedback control of large scale systems* (1. publ). Prentice Hall International series in systems and control engineering. Prentice Hall Internat
- McCubbin L (2001) *Challenges to the Definition of Resilience*. <http://files.eric.ed.gov/fulltext/ED458498.pdf>
- Mentges A, Halekotte L, Schneider M, Demmer T, Lichte D (2023) A resilience glossary shaped by context: Reviewing resilience-related terms for critical infrastructures. *Int J Disaster Risk Reduct* 96:103893. <https://doi.org/10.1016/j.ijdrr.2023.103893>
- Nan C, Sansavini G (2017) A quantitative method for assessing resilience of interdependent infrastructures. *Reliab Eng Syst Saf* 157:35–53. <https://doi.org/10.1016/j.res.2016.08.013>
- Norris FH, Stevens SP, Pfefferbaum B, Wyche KF, Pfefferbaum RL (2008) Community resilience as a metaphor, theory, set of capacities, and strategy for disaster readiness. *Am J Community Psychol* 41(1–2):127–150. <https://doi.org/10.1007/s10464-007-9156-6>
- Poulin C, Kane MB (2021) Infrastructure resilience curves: Performance measures and summary metrics. *Reliab Eng Syst Saf* 216:107926. <https://doi.org/10.1016/j.res.2021.107926>
- Sadeghi K, Ojha D, Kaur P, Mahto RV, Dhir A (2024) Explainable artificial intelligence and agile decision-making in supply chain cyber resilience. *Decis Support Syst* 180:114194. <https://doi.org/10.1016/j.dss.2024.114194>
- Saurin TA (2015) Classification and assessment of slack: implications for resilience. In *Proceedings of the 6th Resilience Engineering Association Symposium. Lisbon, Portugal: Resilience Engineering Association*
- Saurin TA, Werle NJB (2017) A framework for the analysis of slack in socio-technical systems. *Reliab Eng Syst Saf* 167:439–451. <https://doi.org/10.1016/j.res.2017.06.023>
- Sharkey TC, Pinkley N, Eisenberg SG, D. A., Alderson DL (2021) In search of network resilience: an optimization-based view. *Networks* 77(2):225–254. <https://doi.org/10.1002/net.21996>
- Sharma N, Tabandeh A, Gardoni P (2018) Resilience analysis: a mathematical formulation to model resilience of engineering systems. *Sustainable Resilient Infrastructure* 3(2):49–67. <https://doi.org/10.1080/23789689.2017.1345257>
- Simone F, Cilli C, Di Gravio G, Patriarca R (2024) A Data-Driven Cyber Resilience Assessment for Industrial Plants. In A. Rocha, H. Adeli, G. Dzemyda, F. Moreira, & V. Colla (Eds.), *Lecture Notes in Networks and Systems: Vol. 799, Information Systems and Technologies: Worldcist 2023, Volume 1* (1st ed. 2024, pp. 467–476). Springer Nature Switzerland; Imprint Springer. https://doi.org/10.1007/978-3-031-45642-8_46
- Stolz A, Tang JH, Fischer SG, Fischer K (2024) A rheological model analog for assessing the resilience of socio-technical systems across sectors. *Environ Syst Decisions* 44(4):922–945. <https://doi.org/10.1007/s10669-024-09983-w>
- Tang J, Han S, Wang J, He B, Peng J (2023) A comparative analysis of performance-based resilience metrics via a quantitative-qualitative combined approach: Are we measuring the same thing? *Int J Disaster Risk Sci* 14(5):736–750. <https://doi.org/10.1007/s13753-023-00519-5>
- Teodorescu H–NL (2015) Defining resilience using probabilistic event trees. *Environ Syst Decisions* 35(2):279–290. <https://doi.org/10.1007/s10669-015-9550-9>
- Thierney K, Bruneau M (2007) Conceptualizing and Measuring Resilience: A Key to Disaster Loss Reduction. *TR News* 250:14–17. https://onlinepubs.trb.org/onlinepubs/trnews/trnews250_p14-17.pdf
- Vugrin ED, Warren DE, Ehlen MA (2011) A resilience assessment framework for infrastructure and economic systems: Quantitative and qualitative resilience analysis of petrochemical supply chains to a hurricane. *Process Saf Prog* 30(3):280–290. <https://doi.org/10.1002/prs.10437>
- Walker B, Holling CS, Carpenter SR, Kinzig AP (2004) Resilience, Adaptability and Transformability in Social-ecological Systems. *Ecol Soc*. <https://doi.org/10.5751/ES-00650-090205>
- Woods DD (2015) Four concepts for resilience and the implications for the future of resilience engineering. *Reliab Eng Syst Saf* 141:5–9. <https://doi.org/10.1016/j.res.2015.03.018>
- Woods DD (2018a) Resilience is a verb. In: Trump BD, Florin M–V, Linkov I (eds) *IRGC Resource Guide on Resilience (Volume 2)*. EPFL International Risk, pp 167–172
- Woods DD (2018b) The theory of graceful extensibility: Basic rules that govern adaptive systems. *Environ Syst Decisions* 38(4):433–457. <https://doi.org/10.1007/s10669-018-9708-3>
- Woods DD, Wreathall J (2008) Stress-strain plots as a basis for assessing system resilience. In C. P. Nemeth & E. Hollnagel (Eds.), *Resilience Engineering Perspectives, Volume 1*. CRC Press. <https://doi.org/10.4324/9781315244396-17>
- Yodo N, Wang P (2016) Engineering resilience quantification and system design implications: a literature survey. *J Mech Des* 138(11):111408. <https://doi.org/10.1115/1.4034223>
- Yu DJ, Schoon ML, Hawes JK, Lee S, Park J, Rao PSC, Siebeneck LK, Ukkusuri SV (2020) Toward general principles for resilience engineering. 1539–6924 40(8):1509–1537. <https://doi.org/10.1111/risa.13494>

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