Proceedings of the 33rd European Safety and Reliability Conference (ESREL 2023)
Edited by Mário P. Brito, Terje Aven, Piero Baraldi, Marko Čepin and Enrico Zio
 Published by Research Publishing, Singapore.
doi: 10.3850/978-981-18-8071-1_P434-cd

On the Use of Control Theory to Enhance Systems Towards Resilience

Tobias Demmer

Institute for the Protection of Terrestrial Infrastructures, German Aerospace Center (DLR), Germany. E-mail: tobias.demmer@dlr.de

Jens Kahlen

Institute for the Protection of Terrestrial Infrastructures, German Aerospace Center (DLR), Germany. E-mail: jens.kahlen@dlr.de

Daniel Lichte

Institute for the Protection of Terrestrial Infrastructures, German Aerospace Center (DLR), Germany. E-mail: daniel.lichte@dlr.de

Kai-Dietrich Wolf

Institute for Security Systems, University of Wuppertal, Germany. E-mail: wolf@iss.uni-wuppertal.de

This contribution explores the potential of control theory for improving system resilience. It is essential that critical systems are able to withstand adversarial attacks and other forms of disruption. We discuss how this can be achieved through the use of control theory to allocate resources. In this work, control theory – as an established mathematical framework – is used to analyse the behaviour of a generic system in order to ensure resilience. Finally, this contribution provides an example of a resilient system design that uses control theory and we discuss the advantages and disadvantages of the approach, and how it may be implemented to achieve optimal system resilience.

Keywords: Resilience Quantification, Control Theory, Systems Engineering.

1. Introduction

In recent years, the concept of resilience has become increasingly important within the context of protecting critical infrastructures. To improve the resilience of a system, it is essential to develop methods to accurately quantify relevant resilience metrics (Häring et al., 2016). This will enable decision makers to identify resilience abilities, to assess and compare the various resilience enhancement options and decide on the best course of action. To this end, resilience can be be quantified through an evaluation of the system's performance during a disruption.

In this work we want to view the resilience of a system from the control theory perspective. From that perspective the objective of resilience theory is to bring a time-variable system to a certain state, called a resilient state. Here, a challenge is the imprecise knowledge about the system and various influencing variables. A technical solution to this challenge is to control such system by feedback. This does not necessarily require an accurate model of a system, but the measurability of as many influencing parameters as possible. The system variables of interest are measured and fed back to report the current system state, thus creating a closed-loop control. To illustrate the application of our findings, we use a first order differential equation model, describing a generic public infrastructure framework (Muneepeerakul 2017). Using this model, we show the application of a PID-controller to enhance the resilience under different conditions.

2. Background

The basis of resilient design requires consideration of all threats and actions that may occur during

operation. At the core, critical infrastructure operation can be viewed as a large control system, where the objective is to make the infrastructures more resilient against a variety of stressors. O'Connor et al. (2006) states that "Resilient control systems are those that tolerate fluctuations via their structure, design parameters, control structure and control parameters".

Following this, a resilient control system is one that maintains state awareness and an accepted level of system operationality in response to disturbances. This includes threats of an unexpected and malicious nature. (Rieger et al., 2009)

2.1. Resilience and its Terminology

From a systemic viewpoint, the notions of adaptation, learning, and feedback are of interest. Though there is not a consensus in the resilience literature on tight definitions for these terms (Mottahedi et al., 2021). In general, there is still a lack of comprehensive, cross-disciplinary conceptual treatment of the resilience concept. A first approach to unify the terminology is done in Mentges et al. (2023).

Resilience engineering uses many ideas that are also found in systems optimization literature. (Mayar et al. 2022) Especially in the calculus of extrema and nonlinear programming related to local and global optima and starting points for searches of resilient system states. But engineering resilience also borrows from the systems stability literature. (Wied, 2019) It concentrates on stability near an equilibrium steady state, where resistance to disturbance and speed of return to the equilibrium are used to measure resilience (Holling, 1996).

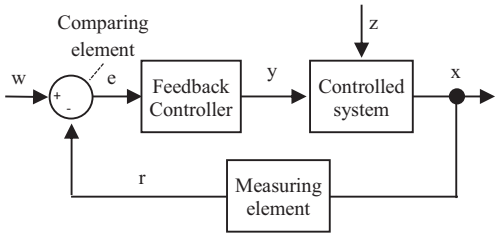
Resilience may thus be visualized, analogous with global and local minima in systems optimization, in terms of a landscape with a single valley (local optimum, mostly in engineering resilience) or multiple valleys (global optimum, mostly in socio-ecological resilience), and movement between states, equivalent to locations on the topography. Some publications describe a ball moving over the topography, where the ball location corresponds to the system state. (Walker et al., 2004)

Resilient systems have a system inherent internal control and an external management, especially in systems with resilience-based design, for example described in Cimellaro et al. (2014). This is one of the first works to extend the performance-based design of systems with a resilience framework.

2.2. Control Theory and its Terminology

In control theory, all external factors influencing the system are introduced in the form of inputs which are classified under two categories—those that can be influenced by the engineer and those that are not controllable. Contrary to the system state, system output indicates the system's external behavior, such as performance, and is normally a direct and observable measure of interest to the engineer. (Unbehauen, 2000)

A basic differentiation is made between open-loop and closed-loop controls. Open-loop, also known as feedforward control or passive control is, where the control action is independent of the system state/output and is selected upfront. Closed-loop (see Fig. 1), also known as active feedback, is selected based on the monitoring of the system state or output and its subsequent comparison with a target (reference/ equilibrium/ steady-state) with the help of a control law or objective.



w Reference variable
e Control difference variable
y Manipulated variable
z Disturbance variable
x Controlled variable
r Feedback variable

Fig. 1 Basic Principle of Closed-Loop Control (DIN IEC 60050-351).

Closed-loop control can be further divided into the three broad categories of optimal, robust, and adaptive controls. Optimal control ensures a system optimization around a reference point or path. In robust control, the control law does not change over time for a certain range of parameter uncertainties of the model and is designed to optimize stability within a particular domain. With

adaptive control, the control law changes over time for the system parameter uncertainties of the model and is designed to optimize stability for a certain criterion (Åström et al., 1987). This is important for different resilience strategies. In this work, optimal control is used as it is mostly about the general idea of system control. The fundamental variable of the system which is related to the uncertainty in the system environment is the disturbance variable.

2.3. Resilience and Control Theory

We dare to integrate these different viewpoints and use system thinking, by applying control systems theory to enhance the inherent resilience of a system. While the two disciplines have different objectives, they both aim to ensure that a system is reliable and safe. Furthermore, both offer the ability to adapt. This means that the system has the ability to change itself or its state in accordance with defined objective functions. And this adaption is better to handle when involving direct feedback. A resilient system also needs the ability to return to its original state (or another suitable system state to be benefitial for the users).

In resilience engineering, the system models are often continuous or discrete linear models (including locally linearized nonlinear models). Therefore, control theory can be implemented into these models. When focusing on resilience theory, the aim is to control the system, to return to its original state or another suitable state (e.g., maximum constant performance). This can be achieved by a passive feedback loop within the system or by active feedback, that is achieved in the form of a closed-loop control action. Here the control law does not change, and therefore it is in the field of optimal or robust control.

Following Holling (1996), we define resilience in this work as an adaption process where the system has the ability to respond to various stressors and change through passive and active feedback structures.

Thus, the system state is changed during a perturbation and returns to a starting position afterwards or transitions to another suitable (stable) state or form.

3. Application

To demonstrate the application of control theory, we use a model based on one first order differential equation, adapted from Muneepeerakul (2017). The original framework focuses on different classes of public infrastructure that affect how utilities interact with a natural resource. Here, the model is stripped down to an interplay between infrastructure providers and the state of public infrastructure, as visualized in Fig. 2. It should be noted, that the general perspective of this model is not restricted to any type of critical infrastructure or provider.

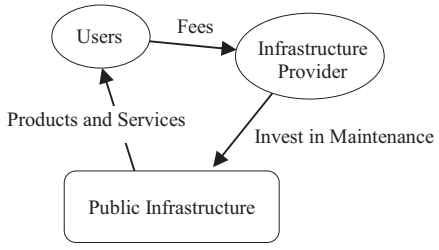


Fig. 2 Schematic Diagram of the Infrastructure-Provider-Model.

In this abstraction, users pay the public infrastructure provider based on their revenue. The infrastructure provider has a maintenance budget based on the revenue and the state of the infrastructure depends on the invested money and leads to a productivity of the users. We extend this model with a varying natural deprecation rate to simulate stressors.

To represent this behavior we use:

$$\frac{dI}{dt} = \mu y C p H(I) - I\delta \tag{1}$$

With

$$H(I) = \frac{100}{1 + e^{-2(I-1.65)}} \tag{2}$$

As in the original paper, I is the state of the infrastructure [Unit I], μ is the maintenance effectiveness [I/$], C is the fraction of user revenue contributed to maintenance [-], y is the fraction of C which infrastructure providers spend on maintenance [-] and δ is the deprecation rate of the

infrastructure [1/T]. Compared to the model in Muneepeerakul 2017, the following parameters had to be adapted: p as the revenue [$/Product] and $H(I)$ maps I to the productivity [Product/T]. This adaption was necessary to satisfy the unit balance. δ is modelled for different values as

$$\delta(t) = \delta_0 + \frac{L}{1 + e^{-1(t-t_1)}} - \frac{L}{1 + e^{-0.1(t-t_2)}} \quad (3)$$

Here L refers to the increment of the disruption, t_1 refers to the beginning of the disruption and t_2 refers to the end of the disruption. The authors assume there is a natural, permanent deprecation rate δ_0 during normal operation, this might be due to wear and tear or natural processes. We further assume that various stressors that can affect the system lead to an increase in delta. Thus, while the disturbance continues, the deprecation rate is permanently at a higher level. All activities that can be summarized under the "restoration" of the system subsequently lead to the fact that the delta can be reduced to its original value δ_0 again after a certain time.

3.1. Initial Situation

The system reacts differently to stress, here implemented by a changing amount of δ. Fig. 3 shows the exemplary situation where we look at different reactions of the system to stress. The initial situation is: $\mu = 0.001$; $C = 0.6$; $y = 0.6$, $I = 3.5$; $p = 10$ and $\delta_{max} = 0.10$ (same as in Muneepeerakul (2017)) We then model a stressor through the increase of the deprecation rate to $\delta_{max} = 0.11$. This small disruption leads to a reduced state of the infrastructure [I], while the system performance remains unchanged [H].

This is important to know, because many systems are judged by their performance (H) and not by their state (I). In this work, in contrast to the majority of approaches to resilience assessment, we differentiate between performance and the system state.

A slightly larger disruption with a further increased $\delta_{max} = 0.12$ leads to a system breakdown, if no action is taken quickly enough to lower the deprecation rate. A longer disturbance of this magnitude leads to system failure. This shows, that the resilience of the system likely depends on the reaction time of the system, i.e. for reinstating the steady state of the system, measures need to be implemented quickly.

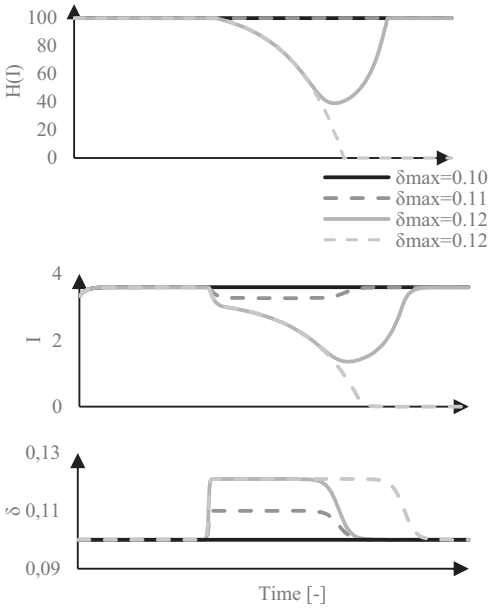


Fig. 3 System performance $H(I)$ and system state I for selected disturbances modelled through an increasing δ.

A further increase of δ leads to an immediate collapse of the system, without the possibility to take appropriate actions.

3.2. Open-Loop Controller

The solution proposed within safety related resilience theory is to increase the margin to failure. This perspective is close to an open-loop controller. Such a controller reacts to pre-defined conditions, but is not able to estimate the time or magnitude of a disturbance. This leads immediately to the question for the amount of margin and of course any invest into such resources can be seen as waste, as long as nothing happens.

The situation where the infrastructure provider spend various fractions of their revenue on maintenance is visualized in Fig. 4, we assume y=1.0 is the absolute possible maximum and everything between 0.6 and 1.0 is one possible approach to the open-loop control. With a larger invest in maintenance, the system is safe to these disruptions. This seems to enhance the resilience, but is not economically smart and additionally a

bigger disruption would still lead to a system collapse.

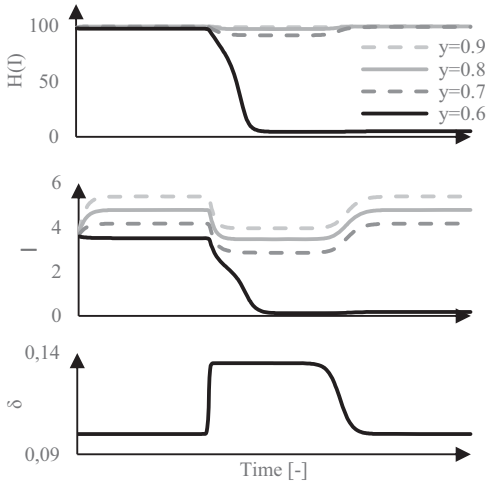


Fig. 4 System performance (*H(I)*) and system state *I* for $\delta_{max}=0.13$ with various fractions which infrastructure providers spend on maintenance (*y*).

The behavior of this open-loop solution is visualized in Fig. 5. It represents *dI/dt* as a function of *I*. This representation is useful because stable and semi-stable states of the system, in which *dI/dt=0* must hold, can be identified very easily. For the initial situation ($\delta_{max} = 0.10, y = 0.6$), there are three points where this applies to. The point at $I\approx1.7$ is semi-stable, because any parameter variation will push system away from this state. $I\approx0$ and $I\approx3.6$ are stable states, because even after a variation the system will bounce back to this state. Note that in the chart of *H*, there is nearly no difference between y=0.8 and y=0.9, such changes are only visible in the chart of *I*.

The increase of *y* from 0.6 to 1.0 leads the system from $I \approx 3.6$ (Fig. 3) to a new steady state at $I \approx 6$, because at this point $dI/dt = 0$. At that point is the stable equilibrium between $\mu yCpH(I)$ and $I\delta$. Now when the disturbance happens, the system only moves to $I\approx5$, but there is still a steady state. Without this open-loop solution the system collapses, because the only steady state is at $I = 0$ ($\delta_{max} = 0.125, y = 0.6$). Note that all of this action can only be seen when looking at the charts of *I*, the charts of *H(I)* remain at 100.

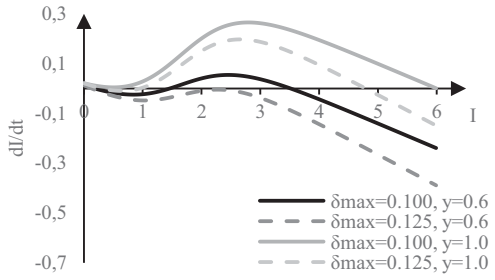


Fig. 5 Visualization of *dI/dt* as a function of *I* (eq. (1)) to identify stable and meta-stable states.

It is important to keep the system in a state where $dI/dt(I) = 0$ exists not only for $I = 0$, because once the system state is zero, it is impossible to recover from such a stressor. The space of possible solutions is pictured in Fig. 6. The figure visualizes the highest stable state of the parameter combinations *y* and *δ*. This can be derived from picturing Fig. 5 for every parameter combination. From an economical perspective, it would be best to optimize the system along the red line, because the costs for the infrastructure providers would be as low as possible. Also, in such a system the smallest disruption would cause a collapse. To prevent this, implementing a margin to failure would lead to an increase of *y* in a distance to this line.

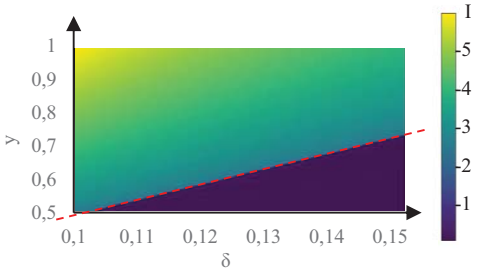


Fig. 6 System behaviour for a selected range of *y* (fraction the infrastructure provider spends on maintenance) and *δ* (deprecation rate of the infrastructure). Below the red line the system collapses.

3.3. Closed-Loop Controller

The results of the Open-Loop Controller show , that more sophisticated methods and multiple criteria should be considered in order to find economically viable ways to stabilise the system behaviour in anticipation of larger disruptions. In reference to the introduced ability of adaptation of

a resilient system (see Sec. 2.3), we strive to adapt *y* only when necessary.

To do that, we introduce a closed-loop controller. The most basic implementation is a PID-controller with the objective to keep the system performance *H(I)* (controlled variable) at 100 (reference For all disruptions. The controller adapts *y* (manipulated variable) depending on the error between *H* and 100 (control difference). The controller uses the following equation for that:

$$y_i = y_{i-1} + PID \tag{4}$$

With

$$PID = K_P e + K_I e_{sum} + K_D \Delta e \tag{5}$$

Where e is the current error, e_{sum} the sum of errors and Δe the change in error. K_P is the proportional, K_I the integral, and K_D derivative term coefficient. K_P, K_I, and K_D control how fast *y* is increased when a disruption happens and how fast it decreases in case of overshooting. This can be due to an overshoot in the control algorithm or due the end of a disruption.

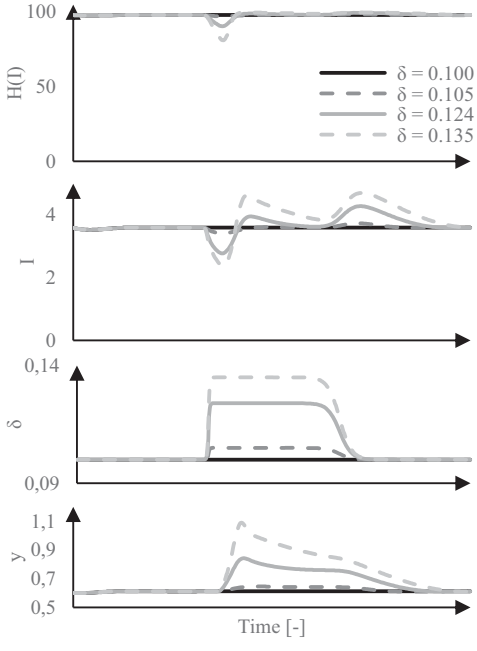


Fig. 7 System performance *H(I)* and system state *I* for disturbances δ of various severity. Performance control by a feedback PID-controller.

Using this controller, the fraction which infrastructure providers spend on maintenance (*y*) is only increased during the incident. The system recovers after the disturbance and maintenance efforts can be reduced afterwards.

A major drawback when looking at system performance in reality is the time delay between an incident and its consequence. We implemented exemplary a delay of 20 timesteps to the controller because in reality it is mostly not possible to make the system to immediately react to stressors. Due to this fact the charts in Fig. 7 are similar to the ones in Fig. 3 until the controller reacts. At this time the controller recognizes a change and adapts the system to the new conditions. This delay is another important aspect of the system resilience, as a higher delay makes it harder to control the system. This implies that a system with shorter dead time is more resilient than a system with slow reaction times.

When looking at the system performance *H*, the controller has no deviation to the reference, apart from the initial drop in performance due to the implemented delay of the controller. The *I* chart (Fig. 7) shows a slightly different picture. The system overshoots and is above the desired state of ~ 3.6 *I*. This happens at first because the controller has to correct the loss in performance and does not register the overshoot. The second overshoot happens when the stressor disappears. The system performance *H* is already at its maximum so the controller does not register the possibility to reduce *y* immediately. The workaround here is to set the reference value slightly below that threshold, so that an overperformance of the system can be detected. We used 99.9 as the reference value here, but this is completely up to the controller designer.

To optimize the described reaction of the system, an important part in control-theory comes to the tuning of the PID-controller. In Fig. 7 the configuration is $K_p = 1\times10^{-3}$, $K_i = 0$ and $K_d = 1\times10^{-2}$. We applied a random number sampling method to tune the PID-controller. The tuning objective was to minimize the overall difference between the measured value for *H(I)* and the target state.

$$\min\left(\int |H(I) - 100|\, dt\right) \tag{6}$$

For this example of a first order differential equation, the only challenge is the deadtime. For this task Eq. (6). is sufficient. For a real system, a more sophisticated approach to tune the controller may be used. In the end, it is very system depended how a controller needs to behave. There are also many more sophisticated objective functions, but this is out of scope here.

When tuning a controller it is important to respect system borders. A bad tuning can lead to an unwanted system behavior, for example overshooting, as visualized in Fig. 8. Such oscillation is in most industrial processes or organizational systems not desirable. The cause for this behavior here is that the selected K_d component is too large and the system starts oscillating even with no stressor.

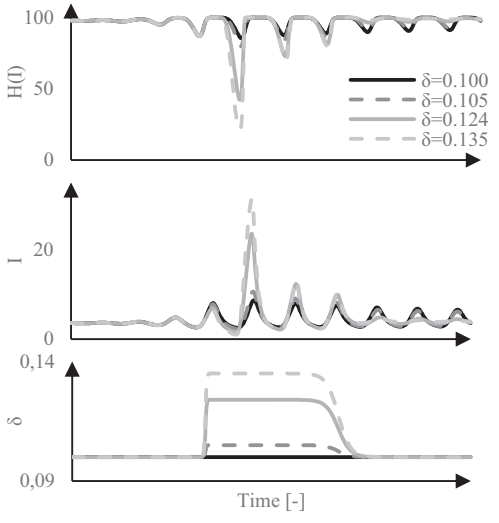


Fig. 8 Misconfigured Feedback Controller H with Kp=0.0001 and Kd=0.1.

3.4. Multiple Stressors

To improve the system behavior described in Fig. 7, a feasible way is to control the infrastructure state *I* instead of the system performance *H(I)*.

Fig. 4 already showed that there are processes in this system, that are only visible in the chart of *I*, and this might be the same in reality: It is easier to measure the performance of a system than the state, but it gives much more insights and possibilities to control the system.

Fig. 9 visualizes a scenario where two independent stressors affect the system subsequently. Additionally to the control of *H(I)*, described in Section 3.3, the control of variable *I* is pictured in this figure. Again, the controller adapts *y* depending on the control difference between *I* and 3.6. The main advantage is that a control of the infrastructure state has no dead time in this model, so it is possible to keep the system at a constant state and following that no drop in performance is visible.

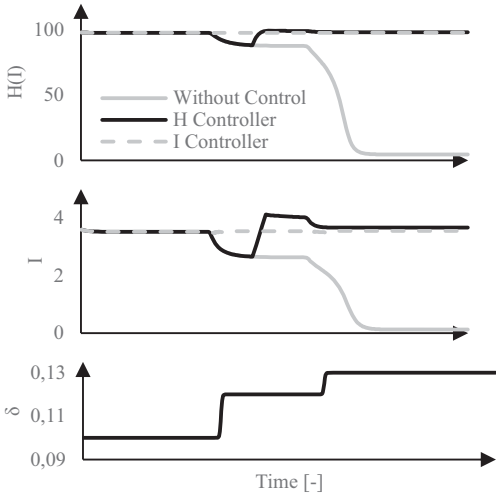


Fig. 9 Sequence of two Stressors.

Finally it is to say that system performance is mostly a product of the system state and that different performance indicators may be more or less sensitive regarding the system state. Therefore it is very important to assess resilience not on system performance but on the system state (*I*). And it is very important, especially when no controller is used, to register every stressor. An example for that is given in Fig. 9. The first stressor increases δ by 0.02 but the system performance decreases only to 88. The second stressor increases δ by an additional 0.01 to 0.13, but this time the system collapses.

4. Conclusion

We began to look at resilience from a control theoretic point of view, because many terms that are used in the resilience literature are borrowed from system theoretic thinking and from a very top-level, infrastructures can be seen as controlled

systems with the objective to maintain a minimum productivity during disruptions.

We presented an safety-related approach using an open-loop controller and a more resilience-related approach using a closed-loop control example at a first order differential equation model and evaluated its resilience enhancement.

To enhance the resilience of a system, one approach is the implementation of feedback control loops. Relatively simple approaches may already lead to a more resilient system behaviour. Our initial results for such approaches show, that it is important to consider damping-effects and dead-time of such feedback loops. However, it should be noted, that the implementation of control-loops is generally also feasible for more complex systems which are not described in differential-equation based models.It is to be discussed how sensitive such controllers need to be and this is very system dependent. For some systems it is necessary to act drastically (e.g. higher K_p or K_d) and other systems allow a more moderate (e.g. lower K_p or K_d) controller setting. A controller that is too sensitive tends to overshoot and oscillate, this is not desireable, but a more inert controller might be too slow to save the system from collapse.

The presented procedure shows potential for a possible adaptive control to enhance the resilient capabilities. An important aspect in particular is the observation of the infrastructure state instead of indirectly observing the performance. Further work may consider other, more complex systems and investigate methods for optimizing the adaptive control.

References

Åström, K. J., L. Neumann, and P. O. Gutman. 1987. A Comparison Between Robust and Adaptive Control of Uncertain Systems. IFAC Proceedings Volumes 20 (2): 43–48. doi: 10.1016/S1474-6670(17)55935-2.

Cimellaro, G. P., C. Renschler, and M. Bruneau. 2015. Introduction to Resilience-Based Design (RBD). In Computational Methods, Seismic Protection, Hybrid Testing and Resilience in Earthquake Engineering, 151–183. Springer, Cham.

Häring, I., S. Ebenhöch, and A. Stolz. 2016. Quantifying Resilience for Resilience Engineering of Socio Technical Systems. European Journal of Security Research 1 (1): 21–58. doi: 10.1007/s41125-015-0001-x.

Holling, C. S. 1996. Engineering Resilience versus Ecological Resilience: Engineering Within Ecological Constraints, 31–44. Washington, D.C: National Academies Press.

Mayar, K., D. G. Carmichael, and X. Shen. 2022. Resilience and Systems—A Review. 2071-1050 14 (14): 8327. doi: 10.3390/su14148327.

Mentges, A., L. Halekotte, M. Schneider, T. Demmer, and D. Lichte. 2023. A resilience glossary shaped by context: Reviewing resilience-related terms for critical infrastructures. doi: 10.48550/arXiv.2302.04524

Mottahedi, A., F. Sereshki, M. Ataei, A. Nouri Qarahasanlou, and A. Barabadi. 2021. The Resilience of Critical Infrastructure Systems: A Systematic Literature Review. Energies, 14(6). doi:10.3390/en14061571.

Muneepeerakul, R., and J. M. Anderies. 2017. Strategic behaviors and governance challenges in social-ecological systems. Earth's Future 5 (8): 865–876. doi: 10.1002/2017EF000562.

O'Connor, M. K., S. M. Mitchell, and M. Mannan. 2006. Designing Resilient Engineered Systems. Chemical Engineering Progress.

Rieger, C. G., D. I. Gertman, and M. A. McQueen. 2009. Resilient control systems: Next generation design research. 2009 2nd Conference on Human System Interactions. doi:10.1109/hsi.2009.5091051

Unbehauen, H. 2000. Regelungstechnik III, Identifikation, Adaption, Optimierung. doi: 10.1007/978-3-322-94391-0

Walker, B., C. S. Holling, S. S. Carpenter, and A. Kinzig. 2004. Resilience, Adaptability and Transformability in Social- ecological Systems. Ecology and Society 9 (2).

Wied, M., J. Oehmen and T. Welo. 2020. Conceptualizing resilience in engineering systems: An analysis of the literature. Systems Engineering, 23(1), 3–13. doi: 10.1002/sys.21491