EAI Endorsed Transactions
on Digital Transformation of Industrial Processes

Research Article **EAI.EU**

Bridging Epidemiological Modelling and Quality Engineering: Optimal Mitigation of Viral Contagion

Daniel Botelho*, M. Teresa Monteiro, Senhorinha Teixeira

ALGORITMI Centre / Department of Production and Systems, Braga, Portugal

Abstract

Industry 5.0 is based on the digital technologies of Industry 4.0. However, it shifts the main focus from pure efficiency to a sustainable, resilient, human-centred approach. Since quality control and waste reduction are essential to sustainability, this exploratory study proposed an innovative approach: the adaptation of epidemiological models to quality control. Specifically, this paper investigates how the classical Susceptible-Infected-Recovered (SIR) model, usually applied to biological disease and information transmission, can be also adapted to determine the optimal moment to apply decisions of quality control and, or, predictive maintenance interventions. The methodology formulated an Optimal Control (OC) problem which is first validated using an empirical case of disinformation spread and solved numerically through an indirect method (using the CasADi software). Results show that OC strategies were able to minimise the system's global cost as well as reduce the number of infected individuals by 8.46%, when comparing to a non controlled scenario. It is concluded that, by transferring this mathematical framework to physical manufacturing production lines, conceptualizing the propagation of defects in processes or components as an infectious phenomenon, engineers and managers can be equipped with a quantitative, data-driven tool. This allows the optimisation of timely interception of defects at their source, ensuring a sustainable reduction of industrial waste as well as the value of operators' decisions within production centres in the context of Industry 5.0.

Received on 13 May 2026; accepted on 27 May 2026; published on 28 May 2026

Keywords: Optimal Control, Epidemiological Modelling, Quality Engineering, Industry 5.0

doi:10.4108/dtip.13017

1. Introduction

Industry 4.0 (I4.0), also known as the Fourth Industrial Revolution, represents the transformative shift in manufacturing and industrial processes through the integration of advanced digital technologies. It enables more efficient, flexible and interconnected production environments [1]. It means that factories, products, people and supply chains are digitally connected, using data and smart technologies to make production more flexible, faster, and customised. However, with technological advancements a new industrial paradigm emerges. Industry 5.0 (I5.0) builds on I4.0 technologies but changes the goal from pure efficiency and profit to human centred, sustainable and resilience [2]. In the end, I4.0 and I5.0 are overlapping paradigms in modern manufacturing. I5.0 should be seen as a continuation or supplement to I4.0 and not a replacement since it addresses I4.0's weaknesses (e.g., dehumanization and social inequality) [3].

Both I4.0 and I5.0 heavily rely on data treatment collection, storage and analysis to support quality control and assurance. However, data handling challenges that were observed in I4.0 no longer apply. Edge-computing models optimise data volumes (99.7% reduction [4]) without losing key information, cutting energy and storage while maintaining real-time quality decisions. In fact, the way data are collected, processed and analysed determines how well defects are prevented, detected and controlled. Good data processing produces trustworthy signals that lead to reliable decisions. When data are noisy, incomplete or poorly analysed, quality decisions become inconsistent. However, when reliable decisions are applied in a timely manner they can reduce defective items, processes or, parts before the next step. This prevents downstream failures and

*Corresponding author. Email: daniel.botelho@dps.uminho.pt

reduces rework and scrap leading to one of the fundamental topics of I5.0: sustainability [5].

Historically, epidemiological models were designed to capture the dynamics of biological disease transmission. However, their underlying mechanics are widely adaptable which leads to their extensive application in the study of information spread [6]. The classic Susceptible-Infected-Recovered (SIR) framework, represented in Figure 1, perfectly illustrates this cross-disciplinary adaptation.

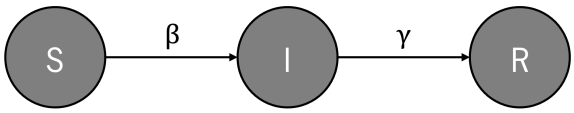

Figure 1. SIR model (S: Susceptible, I: Infected, R: Recovered, β: Infection rate, γ: Recovery rate)

In epidemiology the population is divided in three compartments: Susceptible (healthy, but vulnerable to the infection), Infected (contagious, i.e., actively spreading the disease), and Recovered (immune and no longer contagious). When used in information dynamics this taxonomy undergoes a semantic shift. The Susceptible compartment consists of individuals yet to be exposed to the information, the Infected compartments comprises of spreaders who actively spread the information and the Recovered compartment includes the individuals who have stopped spreading the information.

More recently, this information spread model has been employed to study the spread of fake news (i.e., false information) [6]. The objective is to formulate an Optimal Control (OC) problem solved via an optimisation algorithm to determine the precise optimal timing for implementing control interventions.

In OC theory, numerical solution techniques are generally classified into two main categories: indirect and direct methods. Indirect methods rely on Pontryagin's Maximum Principle (PMP) to derive the necessary conditions for optimality, effectively transforming the OC problem into a two-point boundary value problem that can be solved using the Forward-Backward Sweep Method [7]. While mathematically rigorous, indirect methods are often highly sensitive to the initial guesses of the adjoint variables and may struggle with complex state constraints. Consequently, this study adopts a direct method approach, which discretises the control and state variables to transcribe the OC problem into a Nonlinear Programming (NLP) problem.

This exploratory study presents the methodology to investigate how control optimisation, traditionally applied to biological disease transmission to information spread, can be adapted, and now applied to a new industrial context within the I5.0 era. Specifically, the aim is to address two key questions: when should a quality control decision be implemented to minimise defects, and when is the optimal time to apply predictive maintenance? Furthermore, since these decisions ultimately rely on human judgement, this research aligns with another core pillar of I5.0: the human centred approach.

2. Methodology

This section begins by presenting the empirical data derived from a real-world fake news spreading event. Then, it introduces the formulation of the OC problem based on the classic SIR model.

Empirical Data Acquisition. Following a severe earthquake on January 1^{st}, 2024, rumours quickly emerged on social media suggesting that the seismic event was artificially induced. As reported by [8], the X platform (former Twitter) recorded approximately 370,000 posts containing the keyword "artificial earthquake" between January 1^{st} and January 3^{rd}.

The hourly distribution of these posts is illustrated in Figure 2. The data highlights an immediate spike in activity right after the event, reaching a peak of nearly 30,000 posts per hour. This explosive growth is followed by a rapid decline, with a smaller secondary wave, before stabilizing into a low activity baseline. These data captures typical information dynamics during a crisis: an initial, emotionally driven surge followed by network saturation and eventual decay.

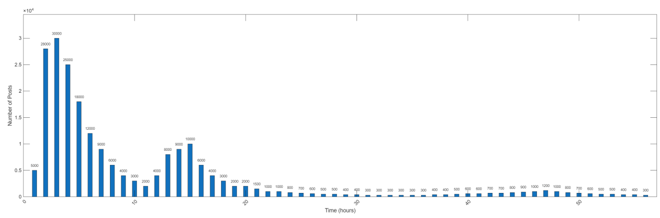

Figure 2. Number of posts between January 1^{st} and January 3^{rd} [8]

To reconstruct this numerical dataset, an Artificial Intelligence (AI) assisted data extraction tool was applied to the original plot provided by [8]. While this approach introduces minor approximations due to image resolution limits, the slight variations do not compromise the overall statistical validity or the underlying trends of the distribution.

Ultimately, these empirical data offer valuable insights into how speculation spreads during real-world emergencies. The initial spike mirrors the public's shock and trauma, whereas the subsequent ripples reflect ongoing debates and content resurfacing. The velocity and short lived intensity of this viral spread emphasize the urgency for mathematically sound control frameworks capable of mitigating such outbreaks before they amplify public panic and confusion ultimately endangering people's lives.

Optimal Control Formulation. A control variable $u(t)$ that represents the mitigation efforts (e.g., fact checking or algorithmic reach reduction) is now introduced. In a real-world context, this variable translates into specific countermeasures. Low intensity controls might correspond to soft moderation techniques whereas high intensity controls could represent the mass suspension of accounts, blocking specific keywords or even temporarily disabling the sharing function. A real example is the restriction on message forwarding implemented by Whatsapp. In response to the abusive use of the platform during elections, the company drastically reduced the forwarding limit to a maximum of five contacts at a time, effectively acting as a control to restrain the spread of false information [9].

In order to study the application of different strategies, an Objective Functional (J) is defined (Equation 1). In this case, the aim is to penalise the disinformation spread (i.e., cumulative number of Infected) as well as the cost of intervention implementation. The objective is to minimize J and determine the control function $u(t)$ that minimises both the number of individuals spreading misinformation and the cost of applying the control. The smaller the value of J, the better equilibrium between effectiveness and cost

Our OC problem is defined in Equation 1:

$$min\, J = I(T) + R(T) + \int_0^T \left[B\, u(t)^2\right] dt, \qquad (1)$$

which is subject to the SIR model with control:

$$\begin{cases} \dot{S}(t) = -\beta S(t)I(t), \\ \dot{I}(t) = \beta S(t)I(t) - \gamma I(t) - u(t)I(t), \\ \dot{R}(t) = \gamma I(t) + u(t)I(t), \end{cases}$$

with initial state conditions:

$$S(0) = \frac{207200}{N}, \qquad I(0) = \frac{5000}{N}, \qquad R(0) = 0,$$

and the control constraints:

$$0 \leq u(t) \leq 1, \quad \forall t \in [0, T].$$

The initial state conditions for the OC problem were derived directly from the empirical dataset. At the start of the observation period ($t = 0$), the dataset recorded an initial active spreading population of 5000 individuals ($I(0) = 5000$).

Assuming a closed system where no individuals have recovered at the beginning ($R(0) = 0$), and considering a total effective population $N = 212200$, the absolute number of susceptible individuals is determined by the conservation of the population:

$$S(0) = N - I(0) - R(0) \qquad (2)$$

To formulate the Optimal Control problem using standard proportions, ensuring numerical stability within the solver, these absolute values are normalized by the total population N. Consequently, the state variables reflect the fraction of the population in each compartment, leading to the initial conditions ($S_0 = S(0)/N$, $I_0 = I(0)/N$, $R_0 = R(0)/N$). This normalisation allows all state variables to remain within the interval $[0, 1]$, making the results easier to interpret.

The term $I(T) + R(T)$ represents the cumulative proportion of individuals who have spread misinformation up to the last time T. $B > 0$ is a weighting parameter that penalises the control effort and can be seen as the cost of applying the control measures. Regarding the system dynamics, S, I and R denote the state variables, $u(t)$ the control variable and, β and γ represent the infection and recovery rates, respectively.

There is also the need to assign values for each parameter that describe the transition rates between the compartments of the SIR model. These parameters are essential to characterise the spread and adjust the model to the empirical data being used. Table 1 shows the values of the different parameters that were obtained using the Least Squares parameter estimation technique [6].

Table 1. Optimal parameters for the SIR model.

Parameter	Optimal value
β	2.8785
γ	1.4841

Numerical Resolution Strategy. To numerically solve the formulated OC problem, presented in the previous section, via the direct method, CasADi [10] software was used. It is an open-source software framework for nonlinear optimisation and algorithmic differentiation.

The numerical resolution procedure can be summarised as follows:

- Time Discretisation: The total time interval $[0, T]$ is divided into equal N subintervals of length Δt;
- State and Control Parametrisation: the state variables (S, I, R) and the control variable $u(t)$ are treated as discrete decision variables at each node of the time grid;
- System Dynamics Constraint Formulation: the differential equations, governing the SIR model, are integrated numerically using the Euler or Fourth-Order Runge-Kutta scheme (RK4). Although the RK4 method offers superior accuracy and lower numerical error, the simpler Euler scheme provides adequate precision for this framework, provided that a sufficiently small step size is employed [11];
- Boundary and Path Constraints: initial conditions (S_0, I_0, R_0) and numerical bounds (e.g., $0 \leq u(t) \leq 1$) are applied as direct constraints on the decision variables;
- NLP Resolution: the resulting NLP problem, comprising the objective function and the set of algebraic constraints, is solved using the Interior Point Optimiser (IPOPT). This specific solver is a robust software package used for large scale nonlinear optimisation and is widely adopted in this field;
- IPOPT iteratively finds the OC sequence $u^*(t)$ and the corresponding state trajectories that minimise the objective function.

3. Results and Discussion

In this section, the main results obtained from the empirical analysis of the Fake Earthquake dataset are presented and discussed. First, the baseline scenario in which no control strategy is implemented, is introduced. Subsequently, the OC problem, previously formulated, is solved. Finally, a comparative analysis between the uncontrolled baseline scenario and the OC based scenario is conducted. Finally, an adaptation of this framework to industrial quality control is proposed. Given this new context, the same underlying SIR dynamics are used with new taxonomy.

3.1. Baseline Epidemic Dynamics

The initial step in assessing the behaviour of the application of control strategies is to establish the evolution of the epidemic in the absence of any intervention. This baseline scenario provides a reference trajectory against which the performance of OC policies can be quantitatively evaluated. In this case, the control input is zero ($u(t) = 0$), while the associated impact or damage is maximal. The resulting uncontrolled dynamics of the SIR model are depicted in Figure 3. Note that two distinct scales are employed on the y-axis to ensure that the trajectory of the infected population remains clearly visible.

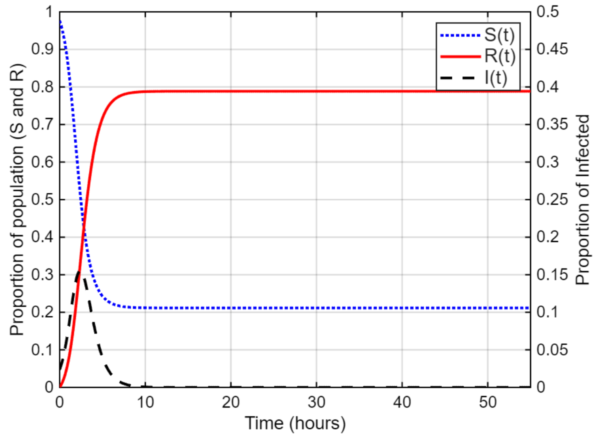

Figure 3. Baseline SIR model dynamics for the Fake Earthquake event without control interventions ($u = 0$)

3.2. Optimal Intervention Strategy

The OC trajectory $u^*(t)$, illustrated in Figure 4, exhibits a decisive and localised intervention strategy. The control effort initiates immediately at $t = 0$ and rapidly escalates to its peak intensity within the first 5 hours of the event. This increase in mitigation effort is strategically aligned with the exponential growth phase of the infected population, as seen in the baseline scenario. By concentrating the available resources precisely during the peak of the information cascade, the controller maximises the recovery rate when the spreader density is at its highest. Following the containment of the initial surge (approximately at $t = 10$), the control effort gradually decrease, as further intervention would yield diminishing returns in terms of cost-effectiveness (J).

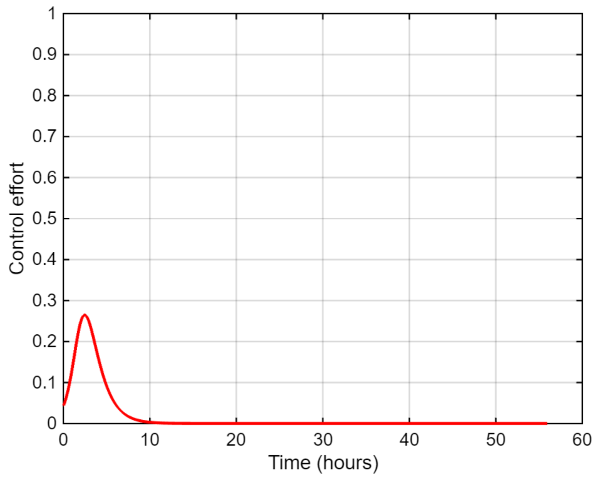

Figure 4. Control effort

Figure 5 illustrates the numerical evolution of the SIR compartments under the OC strategy. A critical observation is the suppression of the infectious peak $I(t)$ compared to the uncontrolled baseline. While the epidemic still reaches a significant portion of the population due to the high transmission rate ($\beta = 2.8785$), the dynamic intervention succeeds in flattening the curve, delaying the peak and reducing its absolute magnitude.

Furthermore, the Susceptible population $S(t)$ stabilises at an earlier t than in the uncontrolled case, confirming that the optimised effort $u^*(t)$ effectively disrupted the transmission chains to prevent a segment of the population from ever contracting the misinformation. This visual representation validates the trade-off handled by the CasADi solver, where the infectious spread is mitigated while respecting the operational cost constraints.

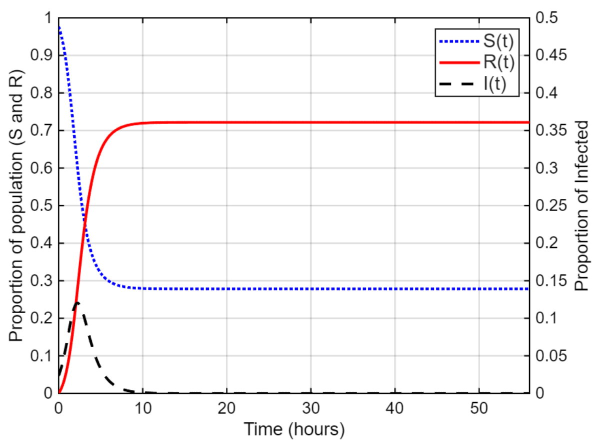

Figure 5. Evolution of the SIR compartments under the OC strategy.

3.3. Comparative Analysis

Figure 6 shows the comparison of both scenarios on each compartment. The visual gap between the solid and dashed curves across all three panels captures the overall effect of the intervention. In the top panel, the solid blue curve ($S(t)$) reaches a higher asymptotic equilibrium than the uncontrolled baseline, visually quantifying the proportion of the population that was successfully shielded from exposure. The middle panel illustrates the curve-flattening effect: the optimal intervention reduces the infectious peak, reducing the maximum instantaneous proportion of spreaders from approximately 15% down to 12.5%. However, the close proximity of the curves in the bottom panel ($R(t)$) also provides a visual understanding to the velocity of this viral event ($\beta = 2.8785$). The explosive initial contagion leaves a very narrow temporal window for the OC to operate, which explains why the gap between the final cumulative number of affected individuals, while mathematically optimal, is constrained to a reduction of only 8.46%.

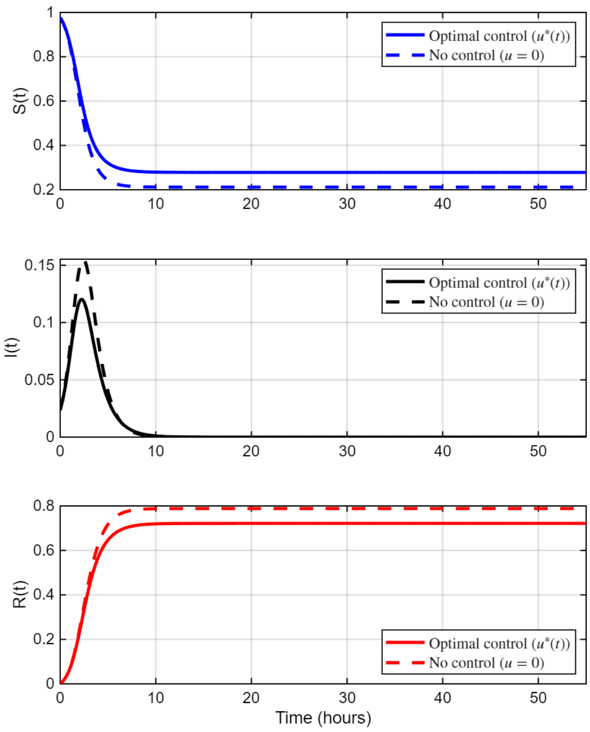

Figure 6. Comparison of scenarios

Table 2 summarises the performance in both scenarios and also presents the cost functional (J) of both scenarios.

Table 2. Comparison of different control strategies: Effectiveness vs. Cost

Strategy	Total Infected	Reduction of Infected (%)*	Cost (J)
Uncontrolled ($u = 0$)	167 319	-	0.7885
Optimal Control**	**153 162**	**8.46%**	**0.7575**

* *Reduction percentage was calculated relative to the uncontrolled scenario*
** *With* $B = 0.2$.

As detailed in the Table 2, remaining inactive ($u(t) = 0$) during this event leads to massive and rapid propagation, affecting 167 319 individuals and resulting in a baseline system cost of $J \approx 0.7885$. By applying the OC strategy, the algorithm successfully fulfils the mathematical condition of optimality, achieving a lower overall cost ($J \approx 0.7575$) and saving 14,157 users from exposure.

However, the practical reduction in total infected individuals is strictly limited to 8.46% (dropping to 153 162). This value highlights a critical limitation of reactive mitigation policies, such as fact-checking or algorithmic debunking. When when the transmission rate (β) is overwhelmingly high, the rapid depletion of the susceptible population overwhelms the system's capacity to transition active spreaders to the recovered state. Consequently, while OC remains mathematically superior to total inaction, these results scientifically demonstrate that combating explosive, panic-driven information cascades requires proactive interventions that directly penalise the transmission rate (β), rather than relying exclusively on reactive recovery efforts.

3.4. Framework Adaptation to Quality Control

In order to broad the applicability of the epidemiological models to industry, more specifically, quality control, the SIR model can be adapted to the context of productive systems with defects propagation of processes or parts.

In this context, $S(t)$ denotes the fraction of components or operational equipment in the production system that are functioning properly, $I(t)$ corresponds to the fraction of failing or unavailable components, and $R(t)$ represents the fraction of recovered units that have returned to service following maintenance or repair. The parameters β and γ, which in epidemiology typically characterize the infection and recovery rates, respectively, assume a different interpretation here. Specifically, β can be regarded as the rate of failure propagation within the system (for example, through cascade effects or operational dependencies between pieces of equipment), whereas γ represents the recovery rate associated with maintenance or corrective interventions. The control variable $u(t)$ characterizes the intensity of preventive maintenance, technical interventions, or quality-related actions implemented to govern the transition of operational components into failed states.

The objective functional J formalises the trade-off between minimising the unavailability of the production system and minimising the costs associated with preventive maintenance, thereby capturing the typical balance, i.e., trade-off, encountered in systems engineering optimisation problems. One example of an objective functional is given by:

$$\min J = \int_0^T AI(t) + Bu(t)^2 dt, \tag{3}$$

Here, $A, B > 0$ denote weighting parameters that control the trade-off between competing criteria (e.g., cost versus reliability).

4. Conclusions

This study successfully formulated and solved an OC problem based in the SIR epidemic model. Although epidemiological models are traditionally employed in biological and public health contexts, the present work repurposes this framework to describe the spread of information, with the further intent of transferring it to industrial applications and quality control to reduce defects and waste. The main objective is to enable the (human) controller to make informed, evidence-based decisions. This human-centred decision-making is fundamental to the paradigm of I5.0.

The numerical results obtained using the CasADi solver indicate that the OC strategy $u^*(t)$ effectively minimised the overall system cost J. In particular, implementing intervention strategies at optimally determined times led to an 8.46% reduction in the number of infected individuals and prevented exposure for more than 14 000 users when comparing to the uncontrolled baseline scenario. Despite the high transmission rate of the considered event ($\beta = 2.8785$), which constrains the responsiveness of any mitigation policy, the dynamic, optimally controlled approach demonstrated superior performance under operational resource limitations.

Having established and validated the validity of this mathematical framework in a simulated environment, the broader aim of this research is its translation into industrial settings, specifically within the domain of quality engineering. Consequently, the immediate future work will focus on conducting an empirical case study within a real manufacturing environment to experimentally validate the proposed concepts. By conceptualising defect propagation along an assembly line as an "infectious" process, historical time-series data on non-conformities can be analysed to derive optimal schedules and intensities for quality control inspections. The application of OC to production lines is intended to intercept and mitigate defects as close as possible to their point of origin, thus directly reinforcing the core pillars of Industry 5.0.

Ultimately, this framework provides a quantitative basis for improving sustainability by systematically reducing scrap and material waste while maintaining the human operator at the centre of the manufacturing system. In doing so, it equips managers and engineers with a decision-support tool for improving planning and prioritising intervention efforts on the basis of data-driven insights.

Acknowledgement. This work has been supported by FCT – Fundação para a Ciência e Tecnologia within the R&D Unit Project Scope UID/00319/2025 - Centro ALGORITMI (ALGORITMI/UM).

References

[1] Gupta, M.K., Gupta, T., Mangal, D., Thapliyal, P. and Biswas, D. (2023) *Study and Analysis of IoT (Industry 4.0) A Review*. doi:10.1201/9781003333760-2.

[2] Dacre, N., Yan, J., Frei, R., Al-Mhdawi, M. and Dong, H. (2024) Advancing sustainable manufacturing: a systematic exploration of industry 5.0 supply chains for sustainability, human-centricity, and resilience. *Production Planning & Control* **36**: 1499 – 1528. doi:10.1080/09537287.2024.2380361.

[3] Ghobakhloo, M., Mahdiraji, H.A., Iranmanesh, M. and Jafari-Sadeghi, V. (2024) From industry 4.0 digital manufacturing to industry 5.0 digital society: a roadmap toward human-centric, sustainable, and resilient production. *Information Systems Frontiers* doi:10.1007/s10796-024-10476-z.

[4] Bajić, B., Suzić, N., Morača, S., Stefanovic, M., Jovicic, M. and Rikalović, A. (2023) Edge computing data optimization for smart quality management: Industry 5.0 perspective. *Sustainability* doi:10.3390/su15076032.

[5] Weiss, E., Caplan, S., Horn, K. and Sharabi, M. (2024) Real-time defect detection in electronic components during assembly through deep learning. *Electronics* doi:10.3390/electronics13081551.

[6] Botelho, D., Monteiro, M.T. and Teixeira, S. (2024) The infodemic issue: Numerical modelling. In Pereira, A.I., Fernandes, F.P., Coelho, J.P., Teixeira, J.P., Lima, J., Pacheco, M.F., Lopes, R.P. *et al.* [eds.] *Optimization, Learning Algorithms and Applications* (Cham: Springer Nature Switzerland): 313–325.

[7] Lenhart, S. and Workman, J.T. (2007) *Optimal Control Applied to Biological Models* (Boca Raton: Chapman & Hall/CRC).

[8] Ushio, A. and Mizuno, A. (2024), Scourge of "fake news" spreads after killer quake in ishikawa: The asahi shimbun: Breaking news, japan news and analysis. URL `https://www.asahi.com/ajw/articles/15102962`.

[9] de Freitas Melo, P., Vieira, C.C., Garimella, K., de Melo, P.O.S.V. and Benevenuto, F. (2019) Can WhatsApp counter misinformation by limiting message forwarding? In *Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*: 108–115. doi:10.48550/arXiv.1909.08740.

[10] Andersson, J.A.E., Gillis, J., Horn, G., Rawlings, J.B. and Diehl, M. (2019) CasADi – A software framework for nonlinear optimization and optimal control. *Mathematical Programming Computation* **11**(1): 1–36. doi:10.1007/s12532-018-0139-4.

[11] Rodrigues, H.S., Monteiro, M.T.T. and Torres, D.F.M. (2009) Optimization of dengue epidemics: A test case with different discretization schemes. In *AIP Conference Proceedings* (AIP Publishing), **1168**: 1385–1388.