An Intelligent and Simulation-Based Stress-Testing Framework for Resilience Assessment on Manufacturing Shop Floors

Tanel Aruväli[1[0000-0003-2077-6642]], Daniel Gravino[1[0000-0002-9535-4834]] and Emmanuel Francalanza[1[0000-0003-0317-5909]]

[1] University of Malta, Msida MSD 2800, Malta

tanel.aruvali@um.edu.mt

Abstract. Volatility in manufacturing companies, driven by multiple severe disruptions originating from external environments, has emerged as a critical challenge affecting manufacturing shop floors' operational stability, strategic planning, and investment management. Existing resilience assessment methods mainly remain fragmented, generic, and retrospective in nature, offering limited predictive insights into future resilience dynamics. To address these limitations and enhance resilience assessment in volatile industrial contexts, this study proposes a novel framework that integrates the best practice as implemented in the financial sector, together with realistic disruption scenarios, and simulation of shop floor processes and recovery actions. The study was inspired by the well-established stress testing in the banking sector that provides a structured framework for systemic resilience and risk assessment. A predictive component is incorporated into the framework through the contextualization of a Large Language Model that will retrieve and interpret relevant information from news databases. The transformation of potential external disruptions into more tangible dynamic disruption effects on the shop floors enables manufacturing company managers and stakeholders to analyze the shop floor's longer-term investment strategy and strategic planning. Investors and financial institutions can benefit from deeper insights into the upcoming periods' resilience assessment before making their investment decisions. The proposed framework serves as the foundation for developing a simulation-based resilience assessment model that will facilitate the creation of a decision-support tool to enhance manufacturing competitiveness by increasing shop floors' operational flexibility.

Keywords: Shop Floor Resilience, Operational Resilience Assessment, Disruption Scenario Generation, LLM Training, Stress-Testing Framework.

1 Introduction

Manufacturing companies have recently suffered under several adverse events and disruptions, including increasing energy prices, the pandemic, sharply increased export tariffs, and policy-driven shifts in strategic industrial priorities. The current geopolitical

landscape is further increasing the unpredictability of the external environment, significantly destabilizing the overall manufacturing sector. External factors, though beyond manufacturers' direct control, pose significant challenges to shop floor operational stability, strategic planning, and investment management. Thus, companies have begun to complement their long-term operating strategy with resilience. Resilience can be considered a new KPI that characterizes a system's ability to withstand high-impact disruptions and to mitigate or absorb their effects, enabling a faster recovery [1].

The aim of the research is to develop a scalable stress-testing framework to enhance the assessment of operational resilience on manufacturing shop floors across various resilience categories using a predictive approach. The proposed solution focuses solely on shop-floor resilience and does not address the manufacturing company as a whole.

2 Background

2.1 Resilience Assessment Frameworks

For manufacturing, several models have been proposed for quantifying resilience analytically [2], however, the proposed methods vary significantly, and no industry standard or best practice for shop floors is recognized [3]. Furthermore, the frameworks are predominantly explicit about specific parts of shop floor processes, such as supply chain resilience [4], and are often limited to predefined disruptions. Digital twin-based resilient manufacturing system frameworks provide predictive simulation capacity [5]. However, these contain customized, expensive sensor-based self-optimization tools [6], and are therefore not scalable.

In the banking sector, the Basel Framework [7] sets guidelines and standards for banking regulators to conduct stress tests to assess the banks' resilience. The current Basel III framework was developed as a response to the 2008 financial crisis. The crisis highlighted many weaknesses in the prior Basel II framework: a lack of flexibility to respond quickly as the crisis evolved, reliance on historical statistical relationships, failure to cover extreme scenarios, and reliance on single-event scenarios. The reality revealed that many adverse events occurred simultaneously, and historical data alone was insufficient to design adverse scenarios. Instead, emerging risks should be included, and plausible forward-looking multi-event scenarios developed [8]. For instance, in Europe, the European Systemic Risk Board applies a number of models to develop a narrative-based macro-financial stress scenario, including a shock definition with calibration and a system response, to assess its impact on the resilience of banks [9].

2.2 Large Language Models' Predictive Capacity

The capabilities of Large Language Models (LLMs) have significantly advanced over the last few years, enabling access to massive amounts of information and the ability to reason. LLMs can process large amounts of unstructured data, generate original text, and summarize content. Pre-trained foundation models such as Gemini 3, GPT-5, and Claude Opus 4.5 have been trained on broad information and can be applied across a

wide range of use cases. The foundation models include a pre-training stage as a base and a fine-tuning layer that trains the model on additional data to improve performance. The final quality of LLM output depends on the prompt engineering techniques used to formulate natural-language queries [10].

Tao et al [11] concluded that although LLMs have the ability for event reasoning, they remain unsatisfactory and unbalanced in their reasoning and relation creation. While Shi et al. in the LaMP project [12] received more successful results in LLM-based reasoning in event prediction. The most relevant prediction was obtained by applying abductive reasoning: a contextualized LLM suggested causes for events, which were pattern-matched against actual previous events. This model was trained with real-world news datasets. Another study has reported up to 28% improved prediction accuracy using LLMs crowd compared to median human predictions [13]. Hence, LLMs have been tested across various reasoning and event-prediction applications, demonstrating moderate to high potential.

Whilst Artificial Intelligence (AI) approaches are emerging as enablers for resilience in manufacturing [14], there is a lack of solutions that can support companies in determining or forecasting their performance under the impact of external disruptions. Although manufacturing processes scheduling and performance can be successfully computationally simulated [15] [16] based on mathematical models, no LLM-based predictive resilience assessment frameworks exist for shop floors.

3 Stress Testing Framework on Manufacturing Shop Floors

To address concerns about the practical applicability of current models for shop floor resilience assessment, this research proposes an intelligent and flexible shop floor resilience stress-testing (RESISTTE) framework. The proposed RESISTTE framework (Fig. 1) addresses industrial needs by integrating LLM capabilities for plausible disruption scenario generation as well as simulation to support dynamic resilience assessment.

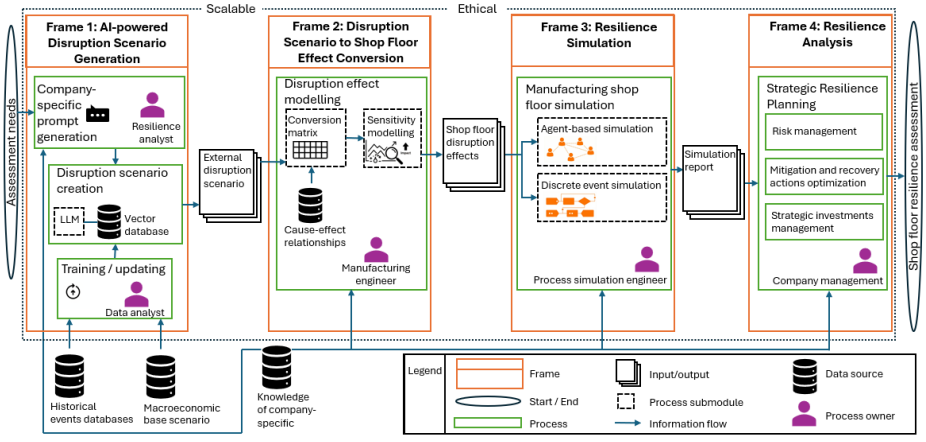


Fig. 1. RESISTTE framework for resilience assessment.

3.1 Framework Description

The framework comprises four frames, supported by external data sources. We propose adopting a predictive feature to identify the most likely upcoming disruptions related to a specific manufacturing shop floor over the coming periods. Simulation capability enables the assessment of resilience dynamics over time and guides the planning of resources and strategic investments of shop floors. Scalability of the framework for broad applicability requires modular manufacturing process models that enable flexible analysis of various types of manufacturing system modifications. An ethical approach to AI-supported processes requires adherence to the Ethics by Design principles, with a focus on human agency and oversight, privacy and data governance, transparency, fairness, and non-discrimination.

AI-powered Disruption Scenario Generation. The aim of this frame is to generate plausible disruption scenarios for the manufacturing companies using event reasoning. For contextualization, Retrieval-Augmented Generation (RAG) can be used by converting retrieved data into numerical representations and storing them in a vector database. This requires input from external databases to deploy an event-detection framework [17].

Disruption Scenario to Shop Floor Effect Conversion. Each disruption has one or more effects on the shop floor, which we call disruption effects. This frame translates external disruption scenarios into their disruption effects on manufacturing shop floors. Some disruptions may have multiple effects on the shop floors. For instance, changes in incoming orders, social distancing, and workers' absences can all be effects of pandemics. Some disruption effects can be caused by various external disruptions. For instance, supply chain disruptions can result from geopolitical tensions and natural disasters. In addition, the frame applies a company-based sensitivity model to predict the severity level of each disruption effect. This can be applied using conversion matrices.

Resilience Simulation. This frame aims to simulate changes on the shop floor resulting from predictive disruption effects and from data of the assessed shop floor. The shop-floor historical datasets and risk management with potential recovery actions constitute the base scenario, which is subsequently complemented by the most probable disruption effects to form the disruption scenario. For scalability, the computational models of the manufacturing system must be modular, enabling the selection of the best-matching set of models for each shop floor. For simulation, AnyLogic provides multimethod capabilities: agent-based simulation for modelling systems and discrete-event simulation for modelling processes.

Resilience Analysis. The simulation report presents the dynamics of various resilience categories on the shop floor over time. It helps visualize the evolution of the disruption effects, the impact of recovery actions, and recognize potential shortcomings. Thus, this document serves as input to strategic resilience planning, including the optimization of mitigation and recovery actions and investment management.

3.2 Framework Implementation and Validation

We propose adopting a Generative-AI-powered LLM foundation model to add capacity for rare-event prediction beyond what current data models, such as macroeconomic

base scenarios, support. The final quality of LLM output depends on the prompt engineering techniques used to formulate queries in natural language. Each prompt engineering technique and its combinations have their specific approach, thus offering distinct benefits. The zero-shot technique relies on pre-trained knowledge being the least complex. The few-shot technique includes providing examples, and Prompt Chaining helps break the complex tasks into manageable steps. In prompting, specific company-based variables, such as its geographical location and field of activity, must be included to identify relevant patterns and disruption scenarios.

For the data to be used for RAG, the framework adopts the approach introduced in [12], exploiting historical news databases that contain news about historical events and related preceding news, such as GDELT [18] and ICEWS to contextualize and train the model on data on historical disruptive events, their dynamics, and news-to-disruption patterns. The macroeconomic base scenario serves as a foundation for modelling the impact of disruptive events. This data can be derived for Europe from the European Economic Forecast [19], published twice a year, and for the world from the World Economic Outlook [20], published quarterly, as well as the same publishers' related datasets. Knowledge of company-specific information can be derived from company management systems, such as Enterprise Resource Planning, Manufacturing Execution System, and other historical and predictive company management tools.

Based on the implementation details, a prototype will be developed, tested, and validated. The validation comprises several aspects for evaluation, such as time to collect company-specific knowledge, base-scenario accuracy, and manufacturing process simulation in the prediction scenario, in terms of performance loss and recovery time.

4 Use Case Scenarios

The following two hypothetical narratives illustrate flexible multipurpose exploitation of the framework. The narratives demonstrate that, in addition to providing intelligent input to manufacturing company managers, the stress test can yield useful insights into shop floor resilience for investors and banks.

4.1 Scenario 1: Manufacturing Company Acquisition

Acryl SARL (Acryl) is a small French manufacturing company operating under a make-to-order manufacturing model. It specializes in customized furniture components made from polymethyl methacrylate (PMMA) sheets in various colors, thicknesses, and production methods, sourced from three suppliers. The wide variety of input materials complicates raw material inventory management. The company owner, preparing for company sales, commissioned an independent stress-test resilience assessment (Fig. 2) after potential investors raised concerns regarding insufficient process transparency.

The predictive module identified risks associated with methyl methacrylate (MMA) availability that is the main ingredient for PMMA. Europe’s third-largest MMA producer, Trinseo PLC, is planning a shutdown. Historical pattern analysis indicated similarities to the 2017 European MMA supply disruption. The conversion module

translated this upstream risk into projected shop floor impacts, indicating a potential PMMA shortage over a 4–9-month horizon. Although two of Acryl's three suppliers were not directly linked to Trinseo's supply chain, the simulation estimated a high probability of material shortage for approximately 6% of customer orders, peaking at 13% in month five before gradual recovery. Other resilience dimensions remained robust.

To mitigate this risk, Acryl engaged customers to obtain advance demand commitments and increased safety stock levels for its most frequently used materials by 50%. The resulting resilience report supported investor due diligence and contributed to a mutually acceptable valuation.

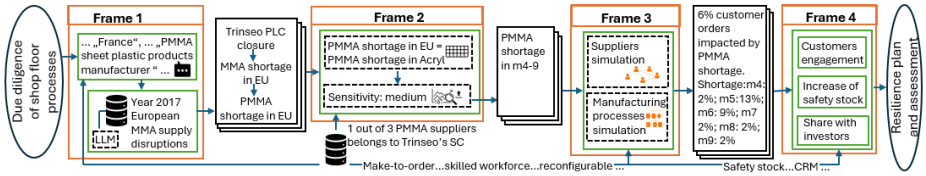


Fig. 2. RESISTTE scenario on manufacturing company acquisition

4.2 Scenario 2: Shop floor resilience mapping

Laptop Ltd. (Laptop) is a large-sized Maltese manufacturer of high-end personal computers. The company's shop floor is highly automated and employs advanced Industry 4.0 technologies for most of its manufacturing processes. Laptop uses last-generation microchips to boost the computational power of its most popular model. It exports 50% of this model to the USA, accounting for 20% of its total sales. The company holds a resilience stress-testing assessment software licence, and running the stress test is part of its monthly risk assessment process to maintain shop floor resilience.

The predictive module identified multiple disruptions that may affect shop-floor processes over the next 18 months. The first potential disruption was political, causing legal changes. Pattern analysis of geopolitical protection mechanisms indicates that predominantly Greenland-related geopolitical tensions between the USA and Europe may prompt the USA to impose export controls on its last-generation microchips on Europe, prompting Europe to reciprocate with the same measures. The second disruption concerns a news pattern in which many large European companies are relocating production from Asia back to Europe, with many relocating to Malta, creating thousands of new jobs in the area. The third potential disruption is a storm period with floods next January. The conversion module translated the disruptions into shop floor effects: a reduced number of orders from the USA, a qualified workforce shortage, and multiple electricity and Internet outages. The simulation model accounts for the current manufacturing system's flexibility, historical customer orders, planned recovery actions, and mitigation measures. The qualified workforce shortage and storm-related outages are disruptions that the stress test tool has already identified over the last 6 months. Therefore, the company has planned and described mitigation measures and recovery actions to protect against these events. For instance, the shop floor is equipped with generators and fuel to power it for at least 7 days. Cloud resources for manufacturing management

systems, such as Enterprise Resource Planning and Manufacturing Execution Systems, are backed up daily to ensure continued operation in the event of an outage.

Export controls are a newly identified potential disruption that, according to the simulation, will cause the stock to reach its maximum capacity in month 16. Company management was informed of the potential risk, and within a week, the management council convened to develop an action plan to address this risk (continuous monitoring, mitigation measures, recovery plan) (Fig. 3).

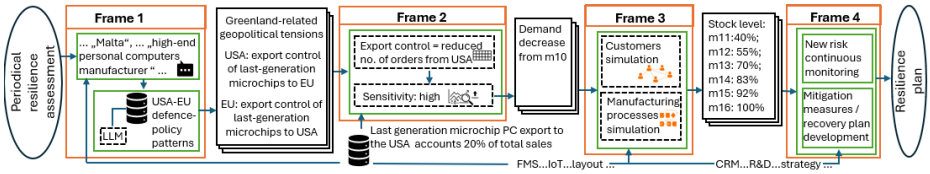


Fig. 3. RESISTTE scenario on shop floor resilience mapping

5 Conclusion and Future Work

The study introduces a RESISTTE framework for operational resilience assessment on manufacturing shop floors. It comprises LLM-based predictive capacity for plausible disruption scenario generation, a conversion module for quantifying disruption effects, and resilience simulation of the respective shop floor processes. Use case scenarios demonstrate the stress test's potential utility for decision-makers, including company managers and investors. However, this framework is a preliminary conceptual system. Thus, before the validation in real-world industrial settings, further empirical testing and iterative refinement within controlled laboratory environments are required.

To enhance flexible resilience assessment through simulation, the future work requires modelling the interconnections among manufacturing process computational models, related disruption effects, and recovery actions. In addition, future work requires the implicit design, testing, and evaluation of the LLM-powered disruption-scenario generator, including the specification of external datasets, LLM libraries, RAG model, and prompting method.

Acknowledgment

Funded by the European Union

This work was funded by the European Union under the Horizon Europe grant number 101207066. The authors have applied a CC BY public copyright licence to the Author Accepted Manuscript of this submission.

References

1. Gu, X., Jin, X., Ni, J., Koren, Y.: Manufacturing system design for resilience. Procedia CIRP. 36, 135–140 (2015).

2. Aruväli, T, De Marchi, M., Rauch, E., Matt. D.: Design decomposition for cyber resiliency in cyber-physical production systems. LNNS. 849, 3-14 (2023).
3. Aruväli, T., Rauch, E., Matt, D.T., Francalanza, E.: SMEs' requirements for resilient cyber-physical production system, Procedia Computer Science, vol. 253, pp. 1611-1620 (2025).
4. Shishodia, A., Sharma, R., Rajesh, R., Munim, Z.H.: Supply chain resilience: A review, conceptual framework and future research. Int J. of Logist. Manag., 34(4), 879–908 (2023).
5. Saporiti, N., Cannas, V.G., Pozzi, R., Rossi, T.: Challenges and countermeasures for digital twin implementation in manufacturing plants: A Delphi study. Int J Prod Econ. 261, 108888 (2023).
6. Wang, H., Peng, T., Nassehi, A., Tang, R.: A data-driven simulation-optimization framework for generating priority dispatching rules in dynamic job shop scheduling with uncertainties. J. Manuf. Syst. 70, 288–308 (2023).
7. Basel Committee on Banking Supervision: The Basel Framework, https://www.bis.org/baselframework/BaselFramework.pdf, (2026).
8. Basel Committee on Banking Supervision: Principles for sound stress testing practices and supervision, https://www.bis.org/publ/bcbs155.pdf, (2009).
9. Gross, M., Henry, J., Rancoita, E.: Macrofinancial stress test scenario design—for banks and beyond. In: Farmer, J.D., Kleinnijenhuis, A.M., Schuermann, T., and Wetzer, T. (eds.) Handbook of Financial Stress Testing. Cambridge University Press (2022).
10. Li, Z.: Analysis of the potential application of GenAI in supply chain optimization technology. In: Proceedings of the 2025 International Conference on Generative Artificial Intelligence for Business. pp. 299-304 (2025).
11. Tao, Z., Jin, Z., Zhang, Y., et al., Wong, K.-F.: A comprehensive evaluation on event reasoning of large language models. In: Proceedings of the AAAI Conference on Artificial Intelligence. pp. 25273–25281 (2025).
12. Shi, X., Xue, S., Wang, K., et al.: Language models can improve event prediction by few-shot abductive reasoning. In: Proceedings of the 37th International Conference on Neural Information Processing Systems. pp. 29532–29557 (2023).
13. Schoenegger, S., Tuminauskaite, T., Park, P., et al.: Wisdom of the silicon crowd: LLM ensemble prediction capabilities rival human crowd accuracy. Sci.Adv. 10, eadp1528 (2026).
14. Rauch, E., Ansari, F., Francalanza, E., et al.: AI as an enabler for long-term resilience in manufacturing. White paper. (2021).
15. Papacharalampopoulos, A., Michail, C.K., Stavropoulos, P.: Manufacturing resilience and agility through processes digital twin: design and testing applied in the LPBF case. Procedia CIRP. 103, 164–169 (2021).
16. Omar, M.M., Sormaz, D.N.: Analysis of manufacturing system resilience through dynamic resource allocation and scheduling. In: Srihari K., Khasawneh M.T., Yoon S., and Won D. (eds.) Lecture Notes in Mechanical Engineering. pp. 379–386. Springer Science and Business Media Deutschland GmbH (2026).
17. Tarekegn, N.A.: Large language model enhanced clustering for news event detection. arXiv.org. (2024).
18. Hong, D., Fu, Z., Zhang, X., Pan, Y.: Research on the development and application of the GDELT event database. Data (Basel). 10, 158 (2025).
19. European Commission's Directorate-General for Economic and Financial Affairs: European economic forecast. Autumn 2025. Institutional paper 327 (2025).
20. International Monetary Fund: World economic outlook update. Global economy: steady amid divergent forces (2026).