

A FOUR-LAYER MODULAR FRAMEWORK FOR KPI MEASUREMENT AND OPTIMIZATION IN SMART MANUFACTURING: A CONCEPTUAL MODEL FOR TRANSITIONAL ECONOMIES

Tanja Trifunović¹
Aleksandar Đorđević
Miladin Stefanović

Received 10.02.2026.
Revised 20.05.2026.
Accepted 10.06.2026.

Keywords:

Industry 4.0, Smart manufacturing,
KPI optimization, Modular
framework, AI/ML analytics,
Transitional economies

ABSTRACT

Industry 4.0 has opened the possibility of moving from reactive KPI monitoring toward predictive and even prescriptive performance management in manufacturing. Most existing optimization models, however, presuppose a high degree of digital maturity, tight OT/IT integration and ready access to specialized staff. These conditions are rarely met in economies still undergoing structural transition, where many factories continue to operate with fragmented systems and limited analytical capacity.

This paper develops a four-layer framework intended specifically for such settings. The layers address successive stages of capability: sensor-based data acquisition supported by 5G, hybrid Edge/Cloud processing, AI/ML models that generate both forecasts and actionable recommendations, and a closed feedback mechanism that continuously refines the system. Because the architecture is modular, firms can begin with basic real-time data collection and add further layers only when resources and skills allow.

The principal contribution lies in offering a staged pathway that lets smaller manufacturers start digital transformation without committing to large-scale, high-risk investments from the outset.



© 2026 Published by Faculty of Engineering

1. INTRODUCTION

1.1 Contextual problem

Contemporary manufacturing is undergoing a fundamental technological transformation driven by Industry 4.0. This paradigm, based on the integration of digital technologies, physical systems and human

resources into a single intelligent network, enables a shift from centralized and sequential production processes toward self-organizing, decentralized and highly connected systems (Drenik, 2025; Joković, 2020; Nikolić et al., 2023). Key Industry 4.0 technologies such as the Internet of Things, artificial intelligence, machine learning, Edge Computing, digital twins and 5G networks make real-time data collection and analysis

¹ Corresponding author: Tanja Trifunović
Email: tanjamtrifunovic94@gmail.com

possible, support predictive maintenance, allow supply-chain optimization and contribute to measurable gains in operational efficiency (Nikiforidis et al., 2025; Rajković et al., 2023; Kang & Kim, 2024; Yuldashev et al., 2023).

In such environment traditional approaches to performance measurement quickly prove insufficient. Key Performance Indicators, which remain the primary instrument for evaluating the efficiency of production processes, now require dynamic and integrated models of measurement and optimization (Međugorac et al., 2011; Velimirović et al., 2011). Instead of static, periodically compiled reports, modern manufacturing needs continuous monitoring, prediction of future KPI values and, where feasible, automatic generation of recommendations for optimal actions (Lepenioti et al., 2020; Zhang, 2025). These capabilities rest on the integration of operational and information systems, advanced data analytics and artificial intelligence (Kumar, 2022; Đorđević et al., 2025; Pratna et al., 2025).

Despite the clear potential, practical application of these concepts in transitional settings such as Serbia and the wider Western Balkans continues to lag behind theoretical possibilities (Joković, 2020; Rajković et al., 2023). Existing studies show that most manufacturing enterprises, particularly smaller ones, still face limited system integration, insufficient digital maturity, shortages of qualified personnel and high implementation costs (Akhmedova et al., 2023). Consequently KPIs are most often monitored periodically and reactively, which prevents timely optimization of production processes and the reduction of downtime, defects and costs.

1.2 Research Problem

The main research problem addressed in this paper is the lack of a contextually adapted, modular framework that would enable manufacturing enterprises operating under constrained conditions to progress gradually from basic real-time data collection to predictive and eventually prescriptive KPI optimization.

Existing KPI optimization models developed within the Industry 4.0 literature have largely been designed for highly digitalized economies that already possess dense IoT infrastructures, full integration of operational and information technologies, and ready availability of specialized personnel. Such models tend to be complex, resource-intensive and insufficiently flexible for firms whose financial and technical capacities remain limited, the very firms that constitute the backbone of Serbian industry (Joković, 2020).

A further difficulty lies in the limited modularity of many published solutions. Authors frequently offer either overly general theoretical frameworks or highly specific, costly systems that leave little room for stepwise adoption (Rajković et al., 2023). Full integration of IoT and 5G connectivity, Edge and Cloud infrastructure,

AI/ML analytics and closed feedback loops into a single, practically usable architecture remains rare, especially under the real operating conditions found in transitional economies.

1.3 Aim and Research Questions

The aim of the paper is therefore to develop a four-layer modular framework for KPI measurement and optimization in smart manufacturing that takes explicit account of the constraints typical of transitional settings. The framework is intended to allow manufacturing enterprises to advance digital capabilities in stages, beginning with reliable data acquisition, moving through processing and analysis, and eventually reaching predictive and prescriptive performance optimization.

Three research questions structure the inquiry:

1. How can a modular framework for KPI measurement and optimization in smart manufacturing be structured?
2. What are the key components of each of the four layers, and how are they integrated?
3. In what way does the modular approach enable gradual implementation in enterprises with limited resources?

Based on these aims and research questions we proposed hypothesis:

H1 states that a modular framework makes it possible to begin implementation without large initial investments.

H2 proposes that the integration of an AI/ML module produces improvements in core indicators such as Overall Equipment Effectiveness, downtime and energy efficiency.

H3 expects the feedback loop to support continuous performance improvement through dynamic adaptation of models and maintenance strategies.

It is important to underline that the paper presents a conceptual framework. The architecture is built on a systematic reading of the existing literature and is offered as a foundation for subsequent empirical validation. No primary data collection was undertaken; the text synthesizes published findings on digital maturity in transitional contexts in order to justify the need for a staged approach. Empirical testing of the framework is left as a direction for future research.

1.4 Contribution of the Paper

The paper offers theoretical, practical and methodological contributions.

Theoretically, it develops a modular architecture that brings four technological layers into a single coherent structure and thereby provides an alternative to

comprehensive, costly and high-risk digital-transformation projects. Such an integrated yet staged approach has not been systematically elaborated in the domestic and regional literature to date.

Practically, the framework is designed to accommodate the conditions under which smaller manufacturing enterprises typically operate. By allowing firms to add capabilities incrementally, it reduces both the initial financial outlay and the risk of implementation failure while still opening a realistic path toward more advanced performance management.

Methodologically, the work demonstrates how findings drawn from multiple sources can be reorganized into a coherent conceptual model that addresses a clearly identified gap. The synthesis itself constitutes a contribution by showing how scattered insights can be assembled into a usable architecture tailored to a specific regional context.

1.5 Structure of the Paper

The paper is organized into five sections. After the introduction, Section 2 presents the theoretical framework and literature review, covering Industry 4.0, KPIs in smart manufacturing, integrated information systems, data analytics and the research gap. Section 3 sets out the proposed four-layer modular framework in detail: IoT and 5G data acquisition, Edge and Cloud infrastructure, AI/ML analytics, and the feedback loop. Section 4 discusses theoretical and practical implications, limitations and recommendations for future research. Section 5 offers concluding remarks.

2. THEORETICAL FRAMEWORK AND LITERATURE REVIEW

2.1 Industry 4.0 and Smart Enterprises

Industry 4.0 marks a far-reaching reorganisation of production and business processes. Digital technologies, physical systems and human expertise are brought together into interconnected networks that can operate with a degree of self-organisation (Drenik, 2025). First articulated in Germany in 2011, the concept moves away from strictly centralised, sequential workflows toward decentralised and highly linked arrangements (Joković, 2020; Brajković, 2019).

Contemporary smart enterprises rely on a cluster of enabling technologies. The Internet of Things, artificial intelligence, machine learning, Edge Computing, digital twins and Blockchain appear most frequently in recent accounts (Rajković et al., 2023; Othman et al., 2026). Together they support continuous data capture, predictive maintenance routines, more responsive supply chains and measurable gains in operational efficiency (Nikiforidis et al., 2025; Nikulin, 2025).

Several authors stress that the transformation is not purely technological. New business models, higher digital competencies among employees and sustained management commitment are equally necessary (Saini et al., 2025; Barbarić, 2025). When these organisational conditions are met, firms can move toward self-optimising operations, more flexible customer-specific production and noticeable reductions in resource use and waste (Woźniak et al., 2022; Woźniak et al., 2024).

Digital twin technology receives particular attention. By creating virtual counterparts of physical assets and processes, it improves monitoring, simulation and real-time decision-making (Rajković et al., 2020; Rajković et al., 2023). Parallel work on network infrastructure seeks to resolve latency, energy-efficiency and security problems that arise in industrial IoT environments (Othman et al., 2026).

Consensus on the potential of these developments is widespread. At the same time the literature repeatedly records obstacles: high capital requirements, shortages of suitably trained staff, cybersecurity risks and slower uptake among smaller manufacturers. These difficulties are especially pronounced in transition regions such as Serbia and the Western Balkans (Joković, 2020; Saini et al., 2025). The question that follows is how performance can be measured and continuously improved under precisely these constrained conditions.

2.2 KPIs and Performance Measurement

Key Performance Indicators remain the principal means by which organisations quantify progress toward strategic and operational goals (Medugorac et al., 2011; Velimirović et al., 2011; Balon et al., 2024; Knapcikova et al., 2024). In the literature they are commonly described as measurable, relevant, reliable and actionable measures that link day-to-day processes with higher-level objectives.

Earlier measurement systems relied heavily on financial metrics. More recent work insists on multidimensional frameworks that also capture quality, time, flexibility, sustainability, safety and the satisfaction of both employees and customers (Domanović, 2013; Balon et al., 2024). The Balanced Scorecard continues to serve as one of the most widely referenced models in this regard, combining financial, customer, internal-process and learning-and-growth perspectives (Čavalić et al., 2017; Damjanović & Domanović, 2023).

Within production and logistics settings, attention has shifted toward indicators tailored to specific industrial contexts. Automotive manufacturers, for instance, track Scheduled Sequence Achievement Ratio, Scheduled Time Achievement Ratio and Standard Production Lead Times Ratio, often in real time through Manufacturing Execution Systems (Đorđević et al., 2025). In the energy sector, indicators focused on equipment efficiency,

energy consumption and maintenance help isolate process signals most strongly correlated with desired outcomes (Lindberg et al., 2015; Rahimi et al., 2020).

Smart manufacturing places particular weight on indicators that can be monitored continuously. Overall Equipment Effectiveness, cycle time, machine utilisation, downtime frequency, predictive-maintenance metrics, throughput, first-pass yield and energy efficiency appear most frequently (Lindberg et al., 2015). Unlike traditional reporting cycles, the Industry 4.0 environment allows these values to be generated automatically from IoT and MES sources, analysed by AI algorithms and adjusted through predictive and prescriptive techniques (Lepenioti et al., 2020).

Methodological contributions have also expanded. Atanasov (2016) proposed a dynamic mathematical model for selecting performance indicators that can adapt to changing business conditions. Simeunović (2015) introduced the GPI model, which links goals, processes and indicators directly and thereby increases the flexibility of the measurement system; Šarović and Perović (2019) later demonstrated its practical utility. Corsini et al. (2025) explore how artificial intelligence, combined with internal and external collaboration, can support the definition of KPIs for circular-economy applications. Zahar Đorđević et al. (2020) developed a fuzzy model suited to the uncertainties typical of small and medium-sized firms, while Kojić et al. (2017) underline that indicators used in turbulent environments must remain measurable, reliable, realistic and actionable if timely corrective action is to be possible.

Despite the variety of available models, recurring difficulties persist: selecting an appropriate set of indicators, integrating them with real-time data streams, overcoming employee resistance, and adapting static frameworks to the rapidly changing conditions of smart production (Atanasov, 2016; Corsini et al., 2025). These persistent gaps explain the growing interest in hybrid approaches that combine established models such as the Balanced Scorecard or GPI with data analytics, artificial intelligence and integrated information systems.

2.3 Integrated Information Systems

Integrated information systems form the backbone of digital transformation in smart manufacturing. They move beyond the closed architectures of earlier generations and enable continuous data exchange across production, logistics, finance, sales and customer-facing functions. Recent literature traces a clear evolution from classical ERP, SCM and CRM platforms toward hybrid, cloud-based solutions that are tightly coupled with IoT technologies.

The decisive step lies in linking IoT data streams with existing enterprise systems. Traditional ERP installations often struggle with the volume, velocity and

heterogeneity of sensor data, which leads to delayed decisions and reduced operational responsiveness. In response, architectures that insert intermediary layers, Edge Computing nodes and machine-learning routines have been developed. These arrangements allow sensor readings to flow directly into the ERP environment, where they can be processed in near real time, generate predictive insights and appear on interactive dashboards (Kumar, 2022).

Comparable integration efforts appear in the CRM domain. Distributed web systems make it possible to exploit rapid data collection from smart devices, personalise customer interactions, react more quickly to shifts in behaviour and strengthen loyalty (Ileana et al., 2025). At supply-chain level, microservice architectures that process IoT data deliver end-to-end visibility, ranging from goods tracking and delay prediction to risk and sustainability optimisation (Kousiouris et al., 2019).

Empirical accounts repeatedly show that organisations which succeed in connecting ERP and CRM systems, supported by Big Data analytics and AI, record gains in productivity and tighter synchronisation of the value chain. Success, however, is never solely technological. It depends on the digital maturity of the firm, management support, employee training and the ability to manage interoperability, data-security and cost challenges, especially among smaller manufacturers (Kowsar & Rahman, 2022; Lo & Pachamanova, 2023).

Cloud computing plays a pivotal enabling role. It allows IoT-generated data to be stored, processed and fed into ERP systems with minimal human intervention, opening the way to higher levels of automation and predictive management. Deng (2022) argues that precisely this interconnection among ERP, SCM and CRM platforms has become a prerequisite for successful digital transformation under contemporary conditions.

Once such integration is achieved, a continuous data path from sensors to managerial decision levels becomes available. That path, in turn, creates the necessary precondition for applying artificial intelligence to real-time data and for constructing models that can monitor and optimise KPIs in volatile production environments (Yormatov et al., 2024).

2.4 Data Analytics in Smart Manufacturing

In smart manufacturing environments vast quantities of data are generated continuously from IoT sensors, integrated information systems and production processes themselves. Analytics converts these raw streams into forms that decision-makers can use and, ultimately, into competitive advantage. The task is no longer limited to describing what has already occurred; it extends to forecasting future states and recommending actions aligned with strategic objectives.

Contemporary business analytics is usually organised into four successive layers. Descriptive and diagnostic techniques clarify past and present performance. Predictive methods, drawing on machine learning and statistical models, estimate what is likely to happen next. Prescriptive analytics occupies the highest level: it recommends concrete actions while taking account of constraints, risks and organisational goals (Lepenioti et al., 2020).

A noticeable shift in the literature moves away from purely data-driven exploration toward goal-oriented analytical decision-making. Instead of beginning with whatever patterns the data happen to reveal, the newer approach starts from a defined business objective and then combines predictive models with causal inference and operations-research techniques to generate optimal recommendations (Lo & Pachamanova, 2023). Hybrid methods that marry machine learning with classical operations research have shown particular strength in complex domains such as inventory management, where they outperform conventional heuristics (Bertsimas & Kallus, 2019).

Sustainability applications illustrate the same progression. Regression models and neural networks are used to assess how design choices affect resource use and environmental impact, while constraint-based prescriptive techniques identify specific opportunities for reducing waste and defective output already at the design stage (Relich, 2023). Tyagi (2021) emphasises the value of parallel processing architectures that can integrate heterogeneous sources, including IoT data, and thereby support the full spectrum of analytical techniques in logistics and transport.

One persistent difficulty remains the selection of relevant performance indicators and the modelling of their interrelationships. Frameworks that integrate multiple data-collection models support the construction of strategic performance maps and the forecasting of future KPI values, while also clarifying how leading indicators influence lagging outcomes (Lepenioti et al., 2020). Such maps enable firms to adjust strategies dynamically across different business models.

Under Industry 4.0 conditions analytics ceases to function merely as decision support and becomes an integral component of both operational and strategic management. By combining predictive and prescriptive capabilities it allows continuous measurement, monitoring and optimisation of KPIs during the actual execution of production processes rather than in periodic reports (Lo & Pachamanova, 2023). When these techniques are applied together, automatic detection of deviations, anticipation of downtime and dynamic adjustment of production parameters become feasible (Bertsimas & Kallus, 2019).

2.5 Research Gap

Although the literature on Industry 4.0 and KPI optimization is extensive, several important limitations persist. A large share of international studies (Lepenioti et al., 2020) develop advanced optimization models that presuppose a highly developed digital infrastructure, full OT/IT integration and ready access to specialized personnel. These assumptions make the models difficult to transfer to transitional settings such as Serbia, where many manufacturing firms still operate with fragmented systems and constrained analytical capacity.

Domestic and regional contributions (Joković, 2020; Rajković et al., 2023) correctly identify the principal barriers to shortage of qualified staff, high implementation costs and low digital maturity, yet they largely remain at a descriptive level. Problems are stated, but concrete, operational solutions that smaller manufacturers could actually adopt are rarely offered. The segment of small and medium-sized manufacturing enterprises, which account for the overwhelming majority of industrial activity in Serbia and lack the financial and technical means for comprehensive system overhauls, remains particularly under-examined.

A further shortcoming is the limited modularity of most proposed architectures. Authors tend either to present broad theoretical frameworks that leave implementation details unspecified or to describe highly specific, capital-intensive solutions that cannot be introduced in stages. Full integration of IoT and 5G connectivity, Edge and Cloud infrastructure, AI/ML analytics and closed feedback loops into a single, practically usable structure is still uncommon.

This paper seeks to address these gaps by constructing a four-layer modular framework oriented toward the conditions under which smaller manufacturing firms in Serbia typically operate. Unlike many existing models, the proposed architecture permits stepwise adoption: an enterprise can begin with basic sensor-based data collection and only later introduce Edge/Cloud processing, predictive and prescriptive analytics, and a feedback mechanism. In this way the work contributes both to theory by linking predictive and prescriptive analytics with an explicitly modular logic and to practice, by offering a pathway that takes account of constrained resources, incomplete system integration and limited specialized expertise.

3. PROPOSED FOUR-LAYER FRAMEWORK

3.1 Overview of the Framework

Drawing on the literature review and the identified shortcomings of existing approaches, this section develops a conceptual framework for KPI measurement and optimization in smart manufacturing. The

architecture integrates selected Industry 4.0 technologies while remaining attentive to the constraints that characterize many manufacturing environments still far from full digital maturity.

The framework is intended to support four interrelated aims:

- 1) collection of data in real time from sensors and production processes;
- 2) processing and analysis of those data with appropriate tools;
- 3) generation of predictive and prescriptive outputs that can guide KPI improvement;
- 4) continuous feedback that allows the system itself to adapt over time.

Although the structure is general enough to be applied across different production settings, particular attention is given to the automotive, metallurgical and processing industries, which remain prominent in Serbia and the surrounding region.

Conceptual model with four layers form the architecture (Figure 1):

The four layers are:

- Layer 1: IoT and 5G data acquisition
- Layer 2: Edge/Cloud infrastructure
- Layer 3: AI/ML predictive and prescriptive analytics
- Layer 4: Feedback loop for dynamic KPI optimization

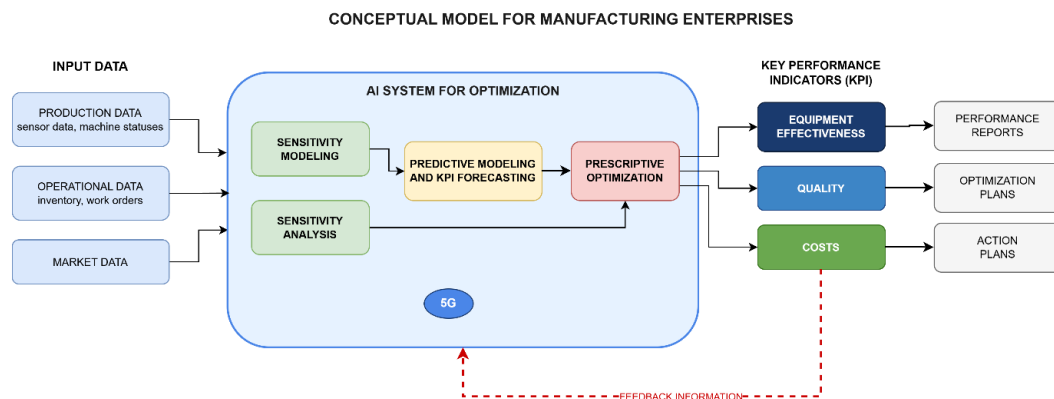


Figure 1. Conceptual model framework

Modularity is central. Each layer can be introduced independently, so that a firm may begin with basic data collection and only later add more advanced functions. The data flow is sequential yet iterative: raw signals originate in Layer 1, are processed and stored in Layer 2, analysed in Layer 3, and returned as recommendations through Layer 4. The cycle then repeats, allowing gradual refinement of performance.

3.2 Layer 1: IoT and 5G Data Acquisition

Reliable, timely data form the indispensable foundation of any subsequent analysis or optimization. Without them the remaining layers cannot function. Layer 1 therefore concentrates on establishing the physical and network infrastructure required for continuous data capture.

Sensors mounted on production equipment supply the primary signals. Temperature, vibration, current, pressure and rotational speed are among the most common parameters monitored. In parallel, devices

already present on the shop floor (PLC controllers, CNC machines equipped with IoT modules, industrial robots) can be networked and treated as additional data sources. Network connectivity is provided by 5G, chosen for its low latency, high throughput and capacity to support large numbers of simultaneous connections; private hybrid wireless access arrangements further address latency and security requirements typical of industrial settings. Edge nodes placed close to the equipment perform initial filtering and aggregation before data are forwarded.

The data themselves fall into four broad categories. Equipment-status readings (temperature, vibration, motor current, pressure) originate from IoT sensors. Production metrics such as Overall Equipment Effectiveness, cycle time, utilization rate and throughput are obtained from PLC and MES systems. Quality indicators (defect rate, first-pass yield) come from control devices. Downtime-related measures (frequency, Mean Time Between Failures, Mean Time To Repair) are drawn from monitoring systems (Table 1).

Table 1. Types of data collected in Layer 1

Data Category	Examples	Source
Equipment status	Temperature, vibration, motor current, pressure	IoT sensors
Production	Overall Equipment Effectiveness, cycle time, utilization rate, throughput	PLC, MES
Quality	Defect rate, first-pass yield	Control devices
Downtime	Downtime frequency, Mean Time Between Failures, Mean Time To Repair	Monitoring system

Because this layer demands primarily hardware investment and basic connectivity rather than advanced analytical skills, it constitutes a realistic starting point even for firms whose digital capabilities remain modest.

3.3 Layer 2: Edge and Cloud Infrastructure

Once data are being collected, they must be stored, processed and distributed. Layer 2 supplies a hybrid infrastructure that combines local Edge processing with centralized Cloud resources.

At the Edge, data are handled close to their source. Local servers or even the IoT devices themselves filter noise, aggregate readings and trigger immediate responses (for example, stopping a machine when a critical threshold is crossed). Pre-processing performed at this stage reduces the volume of data that must travel further and prepares cleaner inputs for later analytics.

The Cloud component complements these local functions by offering scalable storage and computational capacity. Historical records from multiple plants or locations can be retained, complex models trained, and interactive dashboards maintained for managerial oversight. Long-term archives also supply the longitudinal data needed for trend analysis and model retraining.

The division of labour is pragmatic: urgent events are resolved locally to minimize latency, while strategic and historical data move to the Cloud for deeper examination. Firms that cannot yet invest in full Cloud integration may begin with Edge capabilities alone and add the centralized layer later.

3.4 Layer 3: AI and ML Analytics

Layer 3 constitutes the analytical core. Artificial intelligence and machine-learning algorithms convert the data prepared in the preceding layers into forecasts and concrete recommendations.

Predictive applications focus on three domains. First, predictive maintenance: algorithms learn the patterns of equipment behaviour that typically precede failure and estimate the probability of breakdown within a defined horizon (the next shift, the next day, the next week). Maintenance can then be scheduled only when it is genuinely required. Second, quality prediction: process parameters are used to flag products likely to fall outside specification, allowing correction before scrap is generated. Third, cycle-time estimation: the system forecasts the duration needed to complete a given order, improving planning accuracy.

Prescriptive applications move beyond forecasting to recommendation. Genetic algorithms and related techniques suggest machine settings that maximize Overall Equipment Effectiveness while limiting energy consumption. Scheduling algorithms propose operation sequences that reduce idle time and tool changes. When a KPI deviation is anticipated, the system can also recommend a specific corrective action (resetting a parameter, adjusting speed, initiating a targeted maintenance task). Table 2 summarizes how these two forms of analytics map onto selected KPIs.

Table 2. Application of predictive and prescriptive analytics to target KPIs

KPI	Predictive Application	Prescriptive Application
Overall Equipment Effectiveness	Prediction for the next two shifts	Recommendation for line balancing
Downtime frequency	Prediction of downtime probability	Recommendation to perform maintenance before downtime
Cycle time	Prediction of processing time	Recommendation for tool order optimization
Energy efficiency	Prediction of energy consumption	Recommendation for shutting down backup systems

Because the analytical demands of this layer are higher, firms may begin with relatively simple predictive models and only later progress to full prescriptive capability.

3.5 Layer 4: Feedback Loop

Without continuous adaptation the system would gradually lose accuracy as production conditions change. Layer 4 therefore closes the loop.

After each production cycle the predicted KPI values are compared with the values actually achieved. The resulting error is used to retrain the models, typically on a weekly or monthly schedule. Alarm thresholds are adjusted automatically on the basis of historical false-positive rates. Maintenance strategies themselves are revised when previously unpredicted failures occur.

The overall effect is an iterative optimization cycle: data continue to be collected in Layer 1, processed in Layer 2, analysed in Layer 3, and returned as refined recommendations through Layer 4. Over successive iterations the system improves its own performance.

3.6 Modular Implementation

The architecture is deliberately staged so that enterprises can match investment to current capacity. Four successive phases are envisaged.

Phase one establishes Layer 1: sensors are installed on critical equipment, network connectivity is secured, and real-time data streams begin to flow. Hardware and basic

connectivity constitute the main outlay; advanced analytical skills are not yet required.

Phase two adds Layer 2. Edge nodes handle local processing, Cloud storage is introduced for historical data, and simple dashboards make KPI values visible to managers.

Phase three activates Layer 3. Predictive models for maintenance, quality and cycle time are deployed, followed by prescriptive algorithms that suggest parameter adjustments and schedule improvements. Outputs are linked back to the production systems.

Phase four completes the loop with Layer 4. Mechanisms for comparing predicted and actual results are put in place, regular model retraining is scheduled, and dynamic adjustment of thresholds and maintenance rules becomes routine.

Each phase delivers usable value on its own. A firm need not complete the entire sequence in order to benefit; progress can stop or pause at any stage according to resources and priorities. This staged logic is what renders the framework accessible to manufacturers whose digital starting point remains modest.

4. DISCUSSION

4.1 Theoretical Implications

The four-layer framework contributes to the theory of performance measurement under Industry 4.0 conditions in three distinct ways. First, it underscores that KPI optimization models developed for highly digitalized environments cannot be transferred unchanged to settings where digital maturity remains uneven. Full OT/IT integration, dense sensor coverage and readily available specialist expertise (the preconditions assumed by many existing models) are still absent in a substantial share of manufacturing firms in Serbia and the Western Balkans (Joković, 2020). Studies consistently show that domestic enterprises remain at an early stage of digital transformation, with limited real-time data collection and fragmented information systems (Rajković et al., 2023). Any attempt to introduce advanced analytics must therefore begin by securing basic data-acquisition infrastructure. The proposed architecture acknowledges this sequence and is built accordingly.

Second, the paper formalizes the idea of a staged digital pathway. Rather than treating transformation as a single, comprehensive project, the framework treats capability as something that can be added layer by layer: data collection first, then Edge and Cloud processing, then AI/ML analytics, and finally a closed feedback loop. This logic reflects a pattern already observable among resource-constrained firms (Saini et al., 2025), yet it has rarely been articulated as an explicit design principle within the regional literature. Hypothesis H1, which

posits that modularity enables progress without large initial outlays, finds its theoretical grounding precisely in this staged conception.

Third, the framework brings predictive and prescriptive analytics together inside a single, modular structure and links both to a continuous feedback mechanism. Many existing models concentrate on either forecasting or recommendation, and still fewer embed both within an architecture that can itself evolve over time (Bertsimas & Kallus, 2019; Lepenioti et al., 2020). By locating predictive and prescriptive functions in one layer and coupling them to Layer 4, the framework offers a more complete account of how KPI optimization can be sustained rather than merely initiated. Hypotheses H2 and H3, concerning expected gains in core indicators and the adaptive role of the feedback loop, follow directly from this integrated design. These contributions remain theoretical. The framework has not been tested in operating plants, and its practical effectiveness is still to be demonstrated. Nevertheless, the conceptual structure supplies a coherent basis for subsequent empirical work.

4.2 Practical Implications

The framework carries several practical implications for manufacturing enterprises, especially smaller ones, and for the institutions that support industrial development.

For manufacturing enterprises the first implication is straightforward: digital transformation must begin with data. Before resources are committed to advanced analytics or artificial intelligence, reliable real-time collection from production equipment has to be established. Sensors, network connectivity and an unbroken data path from the shop floor to higher decision levels are prerequisites. Without them no amount of subsequent analysis can compensate for missing or unreliable inputs.

A second implication concerns the selection of indicators. At the outset firms are advised to concentrate on a limited set. Overall Equipment Effectiveness and downtime frequency are particularly suitable starting points: they are widely understood, relatively straightforward to measure, and directly linked to operational performance (Đorđević et al., 2025; Lindberg et al., 2015). Once these are monitored reliably, further indicators such as cycle time, defect rate and energy efficiency can be added.

A third practical point is that existing information systems should be integrated before new analytical modules are introduced. Many plants already operate ERP, MES or SCADA systems, yet these often function in isolation. Fragmented data cannot support advanced analytics. Prioritizing the connection of systems already in place therefore precedes investment in AI or machine-learning tools.

A fourth implication relates to human resources. The shortage of staff who combine production knowledge with data-science skills remains one of the most frequently cited obstacles (Joković, 2020). Because simultaneous recruitment of specialists in all required domains is rarely feasible, collaboration with universities or the formation of small interdisciplinary internal teams can provide a more realistic route to the necessary expertise.

For institutions and policymakers the findings suggest several concrete actions. Targeted support for basic IoT infrastructure, especially among smaller manufacturers, would lower the most immediate financial barrier. Educational programmes that combine production engineering, data analytics and artificial intelligence would help close the skills gap. Demonstration projects and publicly accessible case studies of modular implementation would reduce uncertainty among firms still hesitant to begin. Finally, public-private financing of carefully designed pilot applications would generate the empirical evidence needed both to refine the framework and to illustrate its value under local conditions.

4.3 Limitations

The most important is that the framework remains conceptual. No primary data were gathered and no pilot implementation was carried out. The architecture rests on a synthesis of published literature and on reasoned argument about how the four layers should interact. While this approach is legitimate for conceptual work, it leaves the practical effectiveness of the framework untested.

A second limitation concerns the uneven coverage of the underlying literature. Certain elements (predictive maintenance, OEE monitoring) are well documented; others (the combination of 5G with Edge computing under the specific constraints of transitional industrial settings) are less thoroughly explored. In those areas the framework necessarily relies on extrapolation.

A third limitation is the assumption of a minimum level of existing infrastructure. The model presupposes that production equipment can accept sensors, that some form of network connectivity is available, and that at least rudimentary information systems are already in place. In older facilities these conditions may not hold, and the framework does not address the particular difficulties posed by fully legacy equipment or by plants that currently possess no digital infrastructure at all.

A fourth limitation is the relatively light treatment of organizational and cultural factors. Successful digital transformation depends not only on technology but also on management commitment, employee engagement and broader organizational readiness. Although the framework recognizes these factors, it does not provide detailed guidance on how they should be managed.

Finally, the quality of the sources consulted varies. Some references are recent and rigorously documented; others are older or less thoroughly verified. Every effort was made to rely on reliable material, yet the limitations of the available literature remain.

4.4 Future Research

The limitations point to several clear directions for further work.

Empirical validation is the most urgent. Two complementary routes are possible. A large-scale survey of manufacturing firms across Serbia and the Western Balkans could map current levels of digital maturity, identify the most common implementation barriers, and gauge the perceived usefulness of a modular architecture. A second, more demanding route would be pilot implementation of one or more layers in an operating plant, with systematic measurement of effects on selected KPIs. Such pilots would supply stronger evidence than survey data alone, although they would also require substantial time and resources.

A second research avenue concerns algorithms capable of working with incomplete or noisy data. Many firms in transitional settings cannot supply the clean, high-volume datasets that advanced machine-learning models normally require. Methods that tolerate irregularity (fuzzy approaches, probabilistic models, or techniques designed for sparse data) would therefore be of particular practical value.

Comparative analysis across several transitional economies would test the extent to which the framework is regionally robust or, conversely, context-specific. Organizational and cultural dimensions also merit closer examination: technical architectures alone rarely determine success or failure. Finally, mapping the four layers onto established reference architectures such as RAMI 4.0 or IIRA would clarify the framework's position within the wider Industry 4.0 literature and facilitate dialogue with existing standardization efforts.

5. CONCLUSION

The transition from reactive to predictive and prescriptive KPI management cannot be treated as a single leap. For manufacturing enterprises that still operate with incomplete digital infrastructures, limited analytical capacity and constrained budgets, the process has to be incremental. Existing optimization models developed for highly digitalized settings rarely accommodate these starting conditions. They tend either to assume resources and skills that are not yet available or to offer solutions too rigid for gradual adoption.

The framework presented here responds to that gap by treating modularity as a core design principle. Each of the four layers (data acquisition, Edge/Cloud infrastructure,

AI/ML analytics, and the feedback loop) can be introduced independently. Firms may therefore begin with basic real-time data collection and add further capabilities only when resources, skills and operational needs permit. This staged logic lowers the initial investment, reduces the risk of failure, and renders digital transformation accessible to smaller manufacturers that cannot undertake comprehensive, high-cost projects.

Three contributions follow. First, the work assembles four technological layers into a single, coherent architecture that has not previously been articulated in this form within the regional literature. Second, the architecture is explicitly adapted to the constraints typical of transitional industrial environments. Third, a phased implementation path is outlined that allows enterprises to start with low-cost, low-risk steps and to build toward a more fully integrated system over time. The three working hypotheses advanced earlier (H1 on the feasibility of modular adoption without large up-front outlays, H2 on expected gains in core indicators once the AI/ML layer is active, and H3 on the adaptive role of the feedback loop) remain to be tested, yet they supply clear benchmarks for future empirical evaluation.

The limitations of the study must be kept in view. The framework is conceptual; no primary data were collected and no pilot implementation was conducted. The underlying literature is uneven in coverage, and the model assumes a minimum level of existing

infrastructure that may not be present in every plant. Organizational and cultural factors receive only limited attention. These constraints do not negate the contribution, but they do mean that the framework should be regarded as a starting point rather than a finished solution.

Future research should therefore concentrate on empirical validation. Large-scale mapping of digital maturity and implementation barriers across Serbia and the Western Balkans would clarify the practical relevance of the architecture. Pilot applications of one or more layers in operating plants would generate stronger evidence of effectiveness and allow the model to be refined on the basis of real production experience. Complementary work on algorithms tolerant of incomplete data, comparative studies across transitional economies, and closer examination of organizational factors would further strengthen the foundation laid here. In short, the shift toward predictive and prescriptive KPI management is a process, not an event. For enterprises whose digital starting point remains modest, that process must be gradual, modular and realistic about local constraints. The framework offered in this paper supplies one structured route for such a transition. Without pathways of this kind the competitive distance between domestic industry and more advanced manufacturing economies is likely to widen. With them, firms gain a practical means of beginning the transformation and of accumulating capability over time.

References:

- Atanasov, N. S. (2016). *Model za izbor adekvatnog skupa indikatora performansi u upravljanju proizvodnjom* (Doctoral dissertation). University of Belgrade, Faculty of Organizational Sciences.
- Akhmedova, M. Sh., Meliksetyan, K. A., Krutinin, A. A., & Boris, O. A. (2023). The role of quality management in the development of high-tech industrial enterprises in the context of industry 4.0. *Proceedings on Engineering Sciences*, 5(S2), 207–220. <https://doi.org/10.24874/PES.SI.02.003>
- Balon, U., Dziadkowiec, J. M., & Niewczas-Dobrowolska, M. (2024). Key performance indicators (KPIs) in the quality management system. *International Journal for Quality Research*, 18(2), 473–486. <https://doi.org/10.24874/IJQR18.02-10>
- Barbarić, E. (2025). *Korištenje tehnologija Industrije 4.0 u optimizaciji poslovnih procesa: studije slučaja* (Master's thesis). University of Zagreb, Faculty of Economics and Business. Retrieved from <https://urn.nsk.hr/urn:nbn:hr:148:783684>
- Bertsimas, D., & Kallus, N. (2019). From predictive to prescriptive analytics. *Management Science*, 66(3), 1025–1044.
- Brajković, L. (2019). *Industrija 4.0 i digitalna transformacija* (Undergraduate thesis). University of Zagreb, Faculty of Humanities and Social Sciences. Retrieved from <https://urn.nsk.hr/urn:nbn:hr:131:245279>
- Corsini, F., Iannuzzi, T., Fundoni, M., & Frey, M. (2025). Can (A)I help you measure your circularity? An AI-driven and collaborative framework for strategic KPI definition. *Journal of Cleaner Production*, 524, Article 146488. <https://doi.org/10.1016/j.jclepro.2025.146488>
- Čavalić, A., Terzić, N., & Čavalić, S. (2017). Mjerenje poslovnih performansi: Balanced Scorecard. U *Zbornik radova 12. međunarodnog simpozija „Cirkularna ekonomija i održivi razvoj“* (str. 704–716).
- Damjanović, R., & Domanović, V. (2023). Opportunities and limitations of the implementation of contemporary performance measurement models in the context of pre-university education. *International Journal for Quality Research*, 18(2), 383–404. <https://doi.org/10.24874/IJQR18.02-04>
- Deng, L. (2022). Demonstration comparison of the features for ERP, SCM and CRM. *BCP Business & Management*, 34, 68–76. <https://doi.org/10.54691/bcpbm.v34i.2866>

- Domanović, V. (2013). Efektivnost sistema merenja performansi u uslovima savremenog poslovnog okruženja. *Ekonomski horizonti*, 15(1), 31–44. <https://doi.org/10.5937/ekonhor1301031D>
- Drenik, G. (2025). *Industrija 4.0 i pametna proizvodnja* (Undergraduate thesis). University of Rijeka, - Faculty of Engineering. Retrieved from <https://urn.nsk.hr/urn:nbn:hr:190:442833>
- Dorđević, M., Puškarić, H., Vasiljević, S., Kostić, S., Dorđević, M., Trifunović, T., & Đorić, D. (2025). SSAR and STAR as logistics quality KPIs. *International Journal for Quality Research*, 19(4), 1219–1238. <https://doi.org/10.24874/IJQR19.04-11>
- Ileana, M., Petrov, P., & Milev, V. (2025). Optimizing customer experience by exploiting real-time data generated by IoT and leveraging distributed web systems in CRM systems. *IoT*, 6(2), Article 24. <https://doi.org/10.3390/iot6020024>
- Joković, B. (2020). Industry 4.0 implementation challenges in transitional economies. *Management: Journal of Sustainable Business and Management Solutions in Emerging Economies*, 25(2), 15–27.
- Kang, B. G., & Kim, B. S. (2024). Attachable IoT-based digital twin framework specialized for SME production lines. *International Journal of Simulation Modelling*, 23(3), 471–482. <https://doi.org/10.2507/IJSIMM23-3-694>
- Knapcikova, L., Sukic, E., Behunova, A., Tauberova, R., Ljajic, S., & Dedic, N. (2024). Digitalization and technology as main tools to assure the company's performance. *International Journal for Quality Research*, 18(4), 953–964. <https://doi.org/10.24874/IJQR18.04-02>
- Kojić, N., Dajić, M., & Vučković, N. (2017). Merenje i unapređenje indikatora ključnih performansi u savremenoj organizaciji. *Ekonomski signali*, 12(2), 1–9.
- Kousiouris, G., Tsarsitalidis, S., Psomakelis, E., Koloniaris, S., Bardaki, C., Tserpes, K., Nikolaidou, M., & Anagnostopoulos, D. (2019). A microservice-based framework for integrating IoT management platforms, semantic and AI services for supply chain management. *ICT Express*, 5(2), 141–145. <https://doi.org/10.1016/j.icte.2019.04.002>
- Kowsar, M. M., & Rahman, M. A. (2022). Enterprise resource planning and customer relationship management integration: A systematic review of adoption models and organizational impact. *Review of Applied Science and Technology*, 1(2), 26–52. <https://doi.org/10.63125/tn7g2v08>
- Kumar, N. (2022). IoT-enabled real-time data integration in ERP systems. *International Journal of Scientific Research in Science, Engineering and Technology*, 9(6), 393–410. <https://doi.org/10.32628/IJSRSET2215479>
- Lepenioti, K., Bousdekis, A., Apostolou, D., & Mentzas, G. (2020). Prescriptive analytics: Literature review and research challenges. *International Journal of Information Management*, 50, 57–70.
- Lindberg, C.-F., Tan, S. T., Yan, J., & Starfelt, F. (2015). Key performance indicators improve industrial performance. *Energy Procedia*, 75, 1785–1790. <https://doi.org/10.1016/j.egypro.2015.07.474>
- Lo, V. S. Y., & Pachamanova, D. A. (2023). From meaningful data science to impactful decisions: The importance of being causally prescriptive. *Data Science Journal*, 22(1), 1–18. <https://doi.org/10.5334/dsj-2023-008>
- Međugorac, M., Savić, M., Sikimić, U., & Šakota, M. (2011). Razvoj i implementacija ključnih indikatora performansi kao sredstvo modernog upravljanja kompanijom. In *XVII Skup Trendovi razvoja: Evropa 2020 – Društvo zasnovano na znanju* (str. 1–3). Kopaonik, Serbia.
- Nikiforidis, K., Kyrtoglou, A., Vafeiadis, T., Kotsiopoulos, T., Nizamis, A., Ioannidis, D., Votis, K., Tzouvaras, D., & Sarigiannidis, P. (2025). Enhancing transparency and trust in AI-powered manufacturing: A survey of explainable AI (XAI) applications in smart manufacturing in the era of Industry 4.0/5.0. *ICT Express*, 11, 135–148. <https://doi.org/10.1016/j.icte.2024.12.001>
- Nikulin, D. (2025). Utjecaj IOT-a na održivost u industriji 4.0 (Master thesis). Retrieved from <http://ebiblioteka.efsa.unsa.ba/xmlui/handle/EFSA/946>
- Nikolić, N., Djapan, M., Damjanović, A., Radenković, M., & Mačuzić, I. (2023). Implementation of robotics for lean manufacturing improvement. *International Journal for Quality Research*, 17(4), 1127–1140.
- Othman, N. I., Syed, H. J., Ramly, A. M., Suhaimi, N. H. S. B., Ali, A., Abdalnabi, M., & Ismail, A. F. (2026). Beyond 5G: PHWAN – A secure, low-latency, and cost-effective framework for Industry 4.0 smart manufacturing. *Egyptian Informatics Journal*, 33, Article 100859. <https://doi.org/10.1016/j.eij.2025.100859>
- Priatna, D. K., Roswinna, W., Limakrisna, N., Khalikov, A., Abdullaev, D., & Hussein, L. (2025). Optimizing smart manufacturing processes and human resource management through machine learning algorithms. *International Journal of Industrial Engineering and Management*, 16(2), 176–188. <https://doi.org/10.24867/IJIEEM-382>
- Rahimi, Y., Matysenko, I., Kapitan, R., & Pronchakov, Y. (2020). Organization the information support of full logistic supply chains within the Industry 4.0. *International Journal for Quality Research*, 14(4), 1279–1290. <https://doi.org/10.24874/IJQR14.04-19>
- Rajković, D., Petrović, M., & Lazarević, S. (2023). Digitalna transformacija i industrija 4.0 u Srbiji: Izazovi i perspektive. *Tehnika – Menadžment*, 73(1), 45–52.

- Rajković, T., Lečić-Cvetković, D., & Kalem, M. (2020). Digitalni bliznac: Primena savremene tehnologije u proizvodnji. In *XIV Skup privrednika i naučnika – SPIN '20* (pp. 1–10). University of Belgrade, Faculty of Organizational Sciences.
- Relich, M. (2023). Predictive and prescriptive analytics in identifying opportunities for improving sustainable manufacturing. *Sustainability*, 15(9), Article 7667. <https://doi.org/10.3390/su15097667>
- Saini, R., Sharma, S., & Soni, V. (2025). Exploring themes and determinants of Industry 4.0 adoption in smart manufacturing sector: A mixed-methods approach. *Procedia Computer Science*, 277, 197–206. <https://doi.org/10.1016/j.procs.2026.02.061>
- Simeunović, B. P. (2015). *Razvoj modela za mjerenje performansi procesa* (Doctoral dissertation). University of Belgrade, Faculty of Organizational Sciences.
- Šarović, D., & Perović, M. J. (2019). Pristup mjerenju performansi procesa u funkciji ostvarivanja ciljeva organizacije. U *Zborniku radova Festivala kvaliteta 2019* (str. 25–30). Centar za kvalitet, Fakultet inženjerskih nauka u Kragujevcu.
- Tyagi, P. (2021). Diagnostic, descriptive, predictive and prescriptive analytics with geospatial data. *International Journal of Computer Trends and Technology*, 69(1), 18–22. <https://doi.org/10.14445/22312803/IJCTT-V69I1P104>
- Velimirović, D., Velimirović, M., & Stanković, R. (2011). Role and importance of key performance indicators measurement. *Serbian Journal of Management*, 6(1), 63–72.
- Woźniak, J., Budzik, G., Przeszłowski, Ł., Fudali, P., Dziubek, T., & Paszkiewicz, A. (2022). Analysis of the quality of products manufactured with the application of additive manufacturing technologies with the possibility of applying the Industry 4.0 conception. *International Journal for Quality Research*, 16(3), 831–850. <https://doi.org/10.24874/IJQR16.03-12>
- Woźniak, J., Paszkiewicz, A., Budzik, G., Przeszłowski, Ł., Kielbicki, M., & Zimon, D. (2024). Development of additive technologies in the light of the Industry 4.0 and Industry 5.0 conception in accordance with the IoE idea. *International Journal for Quality Research*, 18(2), 643–658. <https://doi.org/10.24874/IJQR18.02-21>
- Yormatov, I. T., Kenzhebekov, N. D., Kadirov, A. M., & Umnova, M. G. (2024). Digital institutes of knowledge and information systems management in support of high-tech production in Industry 4.0. *Proceedings on Engineering Sciences*, 6(3), 1187–1196. <https://doi.org/10.24874/PES06.03A.012>
- Yuldashev, N. M., Yugay, L. Yu., Akopova, E. S., & Fomenko, N. M. (2023). The place of knowledge management in the quality assurance system of industry 4. 0: A view from the point of view of sustainability and competitiveness. *Proceedings on Engineering Sciences*, 5(S2), 371–384. <https://doi.org/10.24874/PES.SI.02.014>
- Zahar Đorđević, M., Simeunović, B., Nestić, S., Aleksić, A., & Puškarić, H. (2020). Study on performance evaluation of the production process – fuzzy MCDM approach. *Journal of Intelligent & Fuzzy Systems*, 39(3), 4009–4026. <https://doi.org/10.3233/JIFS-200159>
- Zhang, Y. X. (2025). Simulation-driven explainable AI for quality prediction and control. *International Journal of Simulation Modelling*, 24(3), 533–544. <https://doi.org/10.2507/IJSIMM24-3-CO14>

Tanja Trifunović

Faculty of Engineering,
Kragujevac,
Serbia

tanjamtrifunovic94@gmail.com

ORCID 0009-0005-1454-5540

Aleksandar Đorđević

Faculty of Engineering,
Kragujevac,
Serbia

adjordjevic@kg.ac.rs

ORCID 0000-0003-2856-6578

Miladin Stefanović

Faculty of Engineering,
Kragujevac,
Serbia

miladin@kg.ac.rs

ORCID 0000-0002-2681-0875
