Benchmarking: An International Journal

27

Received 27 November 2025
Revised 26 January 2026
8 April 2026
Accepted 14 April 2026

Benchmarking human, digital and green performance: a novel adaptive Total Productive Maintenance 5.0 maturity framework

Laura Lucantoni
Universita Politecnica delle Marche, Ancona, Italy
Sara Antomarioni
Universita degli Studi eCampus, Novedrate, Italy, and
Filippo Emanuele Ciarapica and Maurizio Bevilacqua
Universita Politecnica delle Marche, Ancona, Italy

Abstract

Purpose – This article proposes a benchmarking-oriented Total Productive Maintenance 5.0 (TPM5.0) maturity framework that jointly assesses human, digital and green (HDG) performance across all eight TPM pillars in the Industry 5.0 (I5.0) era, enabling peer-referenced maturity assessment across heterogeneous firms and industrial contexts.

Design/methodology/approach – Semi-structured interviews with companies, triangulated with a systematic literature review, informed an HDG–TPM questionnaire. The tool is operationalised into a scoring model that treats maturity as a multidimensional, peer-benchmarking problem, computed via a Multi-Attributive Border Approximation Area Comparison-based scheme advanced with sector-size normalisation and a context-aware weighting architecture. Explainable AI and natural language processing of qualitative comments enhance transparency and support periodic updating.

Findings – The framework produces HDG- and pillar-level maturity profiles, assigns firms to discrete maturity classes and quantifies deviations from sector–size benchmarks. A Food-Small illustrative case demonstrates how misalignments are detected and translated into prioritised improvement actions on specific TPM pillars and HDG practices.

Research limitations/implications – This pilot study covers a limited set of firms within a single sector-size cluster. Future work will expand the sample to test reliability and validity, enable longitudinal analysis and support framework updates.

Originality/value – The study introduces the first TPM5.0 maturity model that operationalises HDG benchmarking through an adaptive, context-aware weighting logic, extending TPM, Lean and operations benchmarking literature. The framework provides a practically deployable diagnostic tool that connects peer-referenced maturity gaps to coordinated HDG improvement priorities under an I5.0 lens.

Keywords Total Productive Maintenance, Industry 5.0, Lean production, Maturity model, Benchmarking

Paper type Research article

1. Introduction

Lean Manufacturing enhances quality and productivity by reducing waste and operating costs through continuous improvement (Ulhe *et al.*, 2024). Within this paradigm, Total Productive Maintenance (TPM) is pivotal for limiting downtime and defects while improving equipment utilisation. As manufacturing transitions towards Industry 5.0 (I5.0), priorities extend beyond Industry 4.0 (I4.0) digitalisation to include sustainable and human-centric trajectories (Ganesan *et al.*, 2025), consistent with the European Commission (2021). For maintenance,

© Laura Lucantoni, Sara Antomarioni, Filippo Emanuele Ciarapica and Maurizio Bevilacqua. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at http://creativecommons.org/licences/by/4.0/.

Benchmarking: An International Journal
Vol. 33 No. 13, 2026
pp. 27-55
Emerald Publishing Limited
e-ISSN: 1758-4094
p-ISSN: 1463-5771
DOI 10.1108/BIJ-11-2025-1075

Downloaded from http://www.emerald.com/bij/article-pdf/33/13/27/11604746/bij-11-2025-1075en.pdf by Universita Politecnica Delle Marche u

this shift requires rethinking TPM as human-oriented, digitally enabled and environmentally responsible.

The evolution from I4.0 to I5.0 is operational in nature. I4.0 is a technological paradigm focused on productivity, efficiency and automation through its enabling technologies (Cannavacciuolo *et al.*, 2023): it acknowledges human involvement, but workers are mainly considered users or beneficiaries of automation (Breque *et al.*, 2021). Sustainability is often treated as a by-product of efficiency, not as a primary goal. In contrast, I5.0 balances three operational pillars: human-centricity (workforce wellbeing, ergonomics, upskilling and inclusive participation), measured through Key Performance Indicators absent from I4.0 frameworks (Maddikunta *et al.*, 2022; Xu *et al.*, 2021); sustainability (resource reuse and emissions management integrated into maintenance and production decisions (Jin *et al.*, 2024)); and resilience through adaptive systems capable of absorbing disruptions without compromising the other two (Breque *et al.*, 2021).

Specifically regarding TPM, I4.0 integration is largely limited to the digital enablement of maintenance (e.g. predictive analytics, IoT-based condition monitoring, computerised maintenance management system (CMMS)). The TPM5.0 vision requires that each pillar be assessed simultaneously along Human (H), Digital (G) and Green (G) dimensions, so that maintenance decisions extend beyond maximising overall equipment effectiveness (OEE) to include ergonomic scheduling, eco-efficient interventions and operator empowerment. Each TPM pillar operationalises HDG dimensions in distinct ways. Autonomous maintenance (AM) combines operator empowerment (H), sensor-assisted inspections and real-time monitoring (D), and eco-autonomous routines (G). Planned maintenance (PM) integrates ergonomic scheduling (H), predictive analytics (H) and energy/waste-optimised maintenance planning (G). Quality maintenance (QM) links quality circles and operator-led anomaly detection (H), AI-based vision systems (D) and defect reduction as a proxy for resource efficiency (G). Training and Education (TandE) supports competency development (H), virtual and augmented reality-learning tools (D) and sustainability awareness (G). Early equipment management (EEM) embodies participatory procurement design (H), digital twins simulation (D) and eco-design criteria for equipment selection (G), consistently with circular economy principles. The safety, health and environment (SHE) pillar is the main vehicle for H, while also addressing D through wearable IoT-based safety monitoring and G through zero-emission and environmental compliance targets. Office TPM (O-TM) operationalises via administrative ergonomics (H), workflow automation (D) and paperless operations (G). Targeted improvement embeds H through bottom-up Kaizen involvement, D via data-driven improvement and G through resource efficiency and waste reduction initiatives. This pillar-level HDG framework makes TPM a suitable operational backbone for I5.0, as it provides a structured architecture for systematically accommodating multidimensional HDG indicators (Bashar *et al.*, 2020).

Although TPM assessment frameworks are widespread (Antomarioni *et al.*, 2023), maturity assessment remains challenging due to the multidimensional nature of TPM pillars and their organisational interdependencies, increasingly requiring multi-criteria approaches (Vikas and Mishra, 2024). Many studies still rely on post-hoc operational or financial indicators (e.g. OEE, mean time between failure and mean time to repair) (Paiva *et al.*, 2024), without jointly assessing HDG dimensions of maintenance practice. Three converging gaps motivate this study.

First, to the best of authors' knowledge, literature lacks an I5.0-aligned TPM maturity model that jointly assesses HDG performance across all pillars. While recent Lean maturity models integrate digital and environmental aspects (Benmamoun *et al.*, 2025; Gatell and Avella, 2024), the human dimension remains underdeveloped (Antomarioni *et al.*, 2025) and "digital sustainable maturity" is still scarce and methodologically fragmented (Silva *et al.*, 2025a). In TPM-specific research, only two structured maturity frameworks exist (Alidrisi, 2020; Shaaban and Awni, 2014), neither of which incorporates digital or environmental dimensions.

Downloaded from http://www.emerald.com/bij/article-pdf/33/13/27/11604746/bij-11-2025-1075en.pdf by Universita Politecnica Delle Marche u

Secondly, existing models rely on static expert-derived weights and internal or sample-limited comparisons, lacking adaptive benchmarking across contexts that vary by sector, size, and technological intensity. These constraints cross-firm learning and reduce practical relevance, especially for SMEs, which cannot meaningfully benchmark against context-agnostic standards (Ferraro *et al.*, 2023; Nascimento *et al.*, 2025).

Third, most models provide static assessments with limited ability to translate results into actionable improvement priorities. Mechanisms for data-driven updating, qualitative feedback integration and explainability are largely absent, increasing obsolescence risk and reducing managerial confidence.

To address these gaps, this article proposes an empirical HDG-TPM5.0 framework that formulates TPM maturity as a multidimensional, context-sensitive benchmarking problem by assessing HDG performance across all eight TPM pillars. Building on a Multi-Attributive Border Approximation Area Comparison (MABAC) computation scheme, it introduces a new contextual weighting architecture that preserves comparability across peer-clusters, while explainable artificial intelligence (XAI) techniques and natural language processing (NLP) based analysis of qualitative comments enhance interpretability and support periodic updating for long-term relevance. The remainder of the article is organised as follows: Section 2 reviews the relevant literature; Section 3 details the methodology; Section 4 presents an illustrative application; Sections 5 and 6 discuss implications and conclusions.

2. Literature review

The literature review was conducted following a structured and replicable procedure to ensure coverage, consistency and transparency in identifying and analysing maturity model contributions within operations management (OM). It follows a progressive logic: first examining general OM maturity models, then Lean-oriented models and finally TPM-specific frameworks. The process includes explicit keyword queries, database selection and screening criteria, as detailed below.

Within OM, maturity models are widely used as diagnostic, benchmarking, and decision-support tools to monitor capabilities, identify improvement areas and guide optimisation (Rendon, 2015; Shukla and Sushil, 2022). However, recurring limitations persist: (1) fragmented and sector-bound constructs with uneven environmental-social-economic coverage; (2) predominantly qualitative assessment and (3) maturity scores rarely anchored to cross-firm benchmarks or translated into actionable guidance (Ferraro *et al.*, 2023; Pavan *et al.*, 2023). Although recent contributions partially mitigate these issues (e.g. sustainability self-assessment (Correia *et al.*, 2023), and I4.0 indicator-based models (Hajoary *et al.*, 2024)), scoring remains subjective and weighting structures largely static. Digital sustainable maturity models remain scarce and heterogeneous (Silva *et al.*, 2025b), and even structured approaches rely on fixed weighting and context-generic benchmarks (Benmamoun *et al.*, 2025; Nascimento *et al.*, 2025). Overall, OM maturity research is evolving towards multidimensional, indicator-based approaches, but adaptive benchmarking architecture remains underdeveloped.

2.1 Lean-oriented maturity models

A Scopus search retrieved maturity contributions in manufacturing using the query: *TITLE-ABS-KEY* ((“maturity” AND (“model” OR “assessment” OR “framework” OR “readiness” OR “level” OR “dimension”)) AND (“lean”) AND (“manufactur*” OR “production”), reaching 361 records The keyword string was designed to capture the main terminological variations used in maturity model research.

Filtering for English-language documents in Engineering, Business, Management and Accounting, and Decision Sciences (261 records) ensured consistency with the OM domain. Limiting the timeframe to 2019 onwards resulted in 171 records, reflecting the convergence of

Downloaded from http://www.emerald.com/bij/article-pdf/33/13/27/11604746/bij-11-2025-1075en.pdf by Universita Politecnica Delle Marche u

Lean maturity research with I4.0 and I5.0 perspectives (Kolla *et al.*, 2019), aligning with the research objectives of this study.

After excluding reviews and proceedings, 152 abstracts were screened. Studies were retained only when including an explicit maturity construct (e.g. levels, dimensions, or scoring logic), resulting in 66 contributions. Among these, 14 integrated Lean with at least one HDG dimension and were analysed in detail (Table 1).

Table 1 shows uneven multidimensionality: the Digital dimension dominates, reflecting data-enabled automation for waste/variability reduction and real-time control (Lucantoni *et al.*, 2022). Green is emerging, and Human remains partial. This imbalance steams from the difficulty of formalising and quantifying human factors (skills, motivation, safety culture, human–machine interaction), which are qualitative and context-dependent, unlike immeasurable indicators such as OEE. This challenge is also evident in I5.0 and human–machine literature (Lucantoni *et al.*, 2026). Consequently, maturity models prioritise technical indicators, while the H dimension remains underdeveloped.

Benchmarking is mainly internal or sample-bound, and adaptive mechanisms are rarely formalised.

A relevant reference is the Green–Lean model by Benmamoun *et al.* (2025), which highlights Environmental Management Systems, Waste Minimisation, and Just-in-Time with TPM as key drivers. Despite these advances, Lean maturity models still rely on static weights and limited benchmarking approaches. These limitations are particularly critical in maintenance context, where multidimensional trade-offs are more pronounced. TPM therefore represents a natural extension of Lean for further analysis.

2.1.1 TPM maturity models. Compared with Lean and I4.0 maturity research, TPM-specific studies remain scarce. As a core Lean component, TPM links equipment reliability, resource efficiency and workforce involvement, enabling multidimensional performance trade-offs (Lucantoni *et al.*, 2023).

A preliminary Scopus search for ("maturity model" AND "TPM") returned only three studies, mostly addressing related concepts (e.g. 5S maturity or Total Project Management (Kumar *et al.* (2021), Žurga (2018)). Although Sanchez *et al.* (2019) propose a transition towards agile TPM, no structured quantitative framework is provided.

Expanding the query to ("maturity" AND "TPM") yielded 18 papers. Most adopt a maturity perspective to assess TPM outcomes (e.g. OEE, reliability) (Badiger *et al.*, 2008; Kwon and Lee, 2004) or Kaizen-based improvements (Vilarinho *et al.*, 2018), rather than providing structured frameworks. TPM is also embedded in broader digital maturity models, but without explicit integration of human and environmental dimensions (Terelak-Tymczyna and Niesterowicz, 2024).

Only two studies propose explicit TPM maturity frameworks. Shaaban and Awni (2014) offer a qualitative, context-specific model based on critical success factors and three stages, without digital or environmental dimensions. Alidrisi (2020) applies MABAC to compute quantitative maturity scores using TPM pillars but relies on static weights, firm-specific normalisation and a non-manufacturing context. Overall, existing frameworks provide useful elements (pillar structuring, staged logic, quantitative scoring) but remain insufficient for I5.0 decision-making.

Across both Lean and TPM models, a common limitation is their cross-sectional nature, with limited evidence on longitudinal robustness, raising concerns about stability and relevance over time.

2.2 Research gap and contribution

OM maturity models increasingly support benchmarking, but are still and mostly pre-I5.0. Specifically, three gaps emerge in Lean maturity literature from both theoretical and industrial perspective:

Downloaded from http://www.emerald.com/bij/article-pdf/33/13/27/11604746/bij-11-2025-1075en.pdf by Universita Politecnica Delle Marche u

Table 1. Lean maturity models overview

Ref	Method	H	D	G	Benchmarking	Dynamic /adaptive	Actionable insights	Evidence
Pulkkinen *et al.* (2019)	DEXTER digital maturity model	–	Y	–	Internal benchmarking across KPAs / units	Static	P – Generic guidance	Case study
Shi *et al.* (2019)	SMKL Kaizen-based smart manufacturing maturity	–	Y	–	Comparison across levels and sites	Static	P – Generic guidance	Case study
Santos *et al.* (2020)	LI4MTA Lean–I4.0 roadmap (BPM-based)	–	Y	–	Current vs target levels (internal benchmarking)	Partial	Y – Explicit roadmap	Case study
Sassanelli *et al.* (2020)	Digital–Lean maturity in product development	–	Y	–	Benchmarking among 3 ETO case companies	Static	Y – Explicit roadmap	Case studies
Zoubek *et al.* (2021)	SMKL 4.0 maturity model (Lean and I4.0)	P	Y	–	Internal benchmarking across plants/lines	Static	P – Generic guidance	Case study
Kandler *et al.* (2022)	Human-centred digital Shopfloor Management: implementation + maturity assessment	Y	Y	–	Current vs target state (internal)	Static	Y – Implementation process + element selection logic	Case study
Martins *et al.* (2023)	Hybrid MCDM for Lean and I4.0 prioritisation	–	Y	–	Ranking/prioritisation of technologies	Static	Y – Explicit technologies prioritisation	Case study
Skalli *et al.* (2023)	ADIME 4.0 digital maturity limited to Six Sigma	P	Y	–	Internal diagnostic across units	Static	P – Generic guidance	Theoretical
e Silva *et al.* (2024)	Lean–Circular Economy self-assessment (MCDA)	–	–	Y	Descriptive cross-company comparison	Static	Y–Explicit recommendations	Case studies
Sajjad *et al.* (2024)	Lean and I4.0 maturity (IMPULS, SIRI; survey)	P	Y	–	Firm segmentation/ clustering	Static	P – Generic suggestions	Survey
Gatell and Avella (2024)	Conceptual DGL maturity model (5 levels)	–	Y	Y	Conceptual comparison of DGL profiles	Static	P - High-level guidance	Theoretical
Benmamoun *et al.* (2025)	GL maturity index (BWM) for pharma SMEs	P	–	Y	Intra-sample GL index	Static	P - High-level guidance	Case study

(*continued*)

Downloaded from http://www.emerald.com/bij/article-pdf/33/13/27/11604746/bij-11-2025-1075en.pdf by Universita Politecnica Delle Marche user on 18 May 2026

Table 1. Continued

Ref	Method	H	D	G	Benchmarking	Dynamic /adaptive	Actionable insights	Evidence
Silva, Nakandala *et al.* (2025)	Lean-centric I4.0 readiness model (6 dimensions, 39 sub-dimensions)	Y	Y	–	Readiness profiling/ maturity matrix (internal/ self-assessment)	Static	P – Structured readiness profile (no stepwise roadmap)	Theoretical
Tadja *et al.* (2025)	Inclusive lean manufacturing maturity tool (size/location analysis)	P	–	–	Cross-company comparison within sample (size/location lens)	Static	P – Tool-based diagnosis	Survey
This study	HDG–TPM 5.0 maturity framework	Y	Y	Y	External benchmarking (contextual, cluster-based)	Dynamic/ adaptive	Y – Data-driven, explainable recommendations	Theoretical + pilot study

Note(s): H/D/G/L = Human/Digital/Green/Lean; Y = explicit, P = partial, – = not addressed
MCDA = Multi-criteria Decision Analysis; MCDM = Multi-criteria Decision Model; BPM = Business Process Management; BWM = Best Worst Method; SMKL = Samart Manufacturing Kaizen Level)
Source(s): Authors' own creation/work

Downloaded from http://www.emerald.com/bij/article-pdf/33/13/27/11604746/bij-11-2025-1075en.pdf by Universita Politecnica Delle Marche user on 18 May 2026

(1) *Multidimensional TPM maturity.* No model integrates HDG performance across all TPM pillars. This is critical, as improvement requires balancing human (skills and involvement, digital maintenance capabilities and green resource efficiency) dimensions across pillars. As shown in Table 1, among 14 Lean contributions integrating at least one HDG dimension, none provides full H×D×G integration; digital dominates, green is emerging and human remains partial, often limited to manufacturing skills or externalised to broader social effects. TPM-specific frameworks (Alidrisi, 2020; Shaaban and Awni, 2014) do not include environmental or digital dimensions.

(2) *Benchmarking architecture for heterogeneous contexts.* Adaptive benchmarking accounting for sector, size and technology intensity is absent. Existing models rely on static weights and internal comparisons, limiting cross-firm learning and practical relevance, because firms, especially SMEs, operate under fundamentally different baselines and cannot meaningfully compare themselves against absolute standards. Table 1 confirms the absence of adaptive weighting mechanisms, consistent with recent OM literature (Ferraro *et al.*, 2023; Nascimento *et al.*, 2025).

(3) *Practical usability.* Most models provide static assessments with limited translation into actionable priorities. Data-driven updating, feedback loops and explainability are rarely embedded, increasing obsolescence and reducing managerial trust. None of the 14 contributions reviewed in Table 1 (see "Dynamic/adaptive" column) includes structured adaptive mechanisms to keeping maturity tools current, despite growing interest in XAI and NLP for model interpretability and longevity documented in recent manufacturing intelligence literature (Lucantoni *et al.*, 2025).

Based on these gaps, this study advances a context-sensitive and multidimensional TPM framework aligned with I5.0, enabling cross-firm benchmarking and translating diagnostic outputs into actionable HDG improvement priorities.

3. Methodology

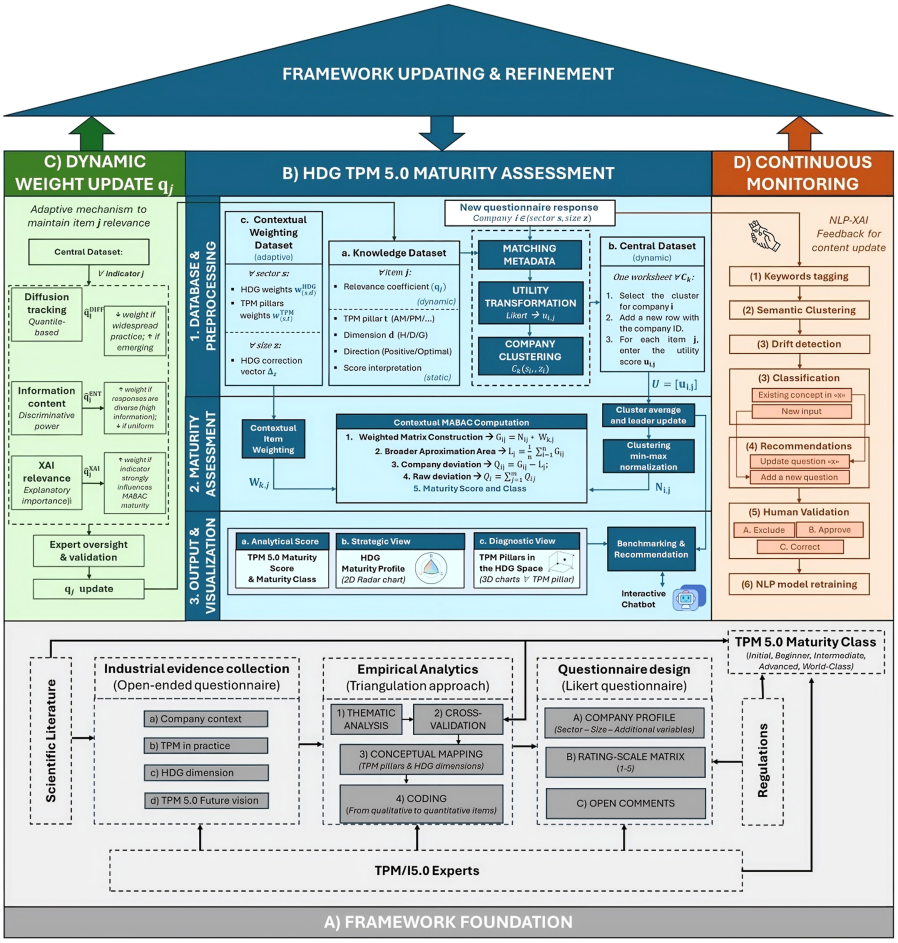
The HDG-TPM5.0 maturity framework (Figure 1) consists of a structured four-phase procedure designed to ensure that the tool is simultaneously grounded in empirical industrial evidence and scientific literature, and applicable to heterogeneous contexts. Phase A represents the foundation layer, where an initial industrial survey is developed, conducted and with evidence translated into measurable items and specific maturity taxonomy. Phase B operationalises the questionnaire into a scoring model through a cluster-based MABAC scheme, while XAI and NLP are used within support layers in Phases C and D. This sequence constitutes the methodological core of the framework and is detailed below.

3.1 Foundation layer (A)

The framework was developed through a multi-phase procedure combining industrial and scientific literature, resulting in a structured questionnaire designed in three steps: (1) evidence collection via a survey of TPM-oriented companies (Section 3.1.1), (2) translation of qualitative evidence into measurable items (Section 3.1.2) and (3) questionnaire design (Section 3.1.3). TPM5.0 maturity classes were also defined at this stage (Section 3.1.4).

3.1.1 Industrial evidence collection. Eight companies, differing in size and sector, were selected for semi-structured interviews to capture TPM implementation and its integration with HDG dimensions (Table 2). Interviews (60–90 minutes) involved personnel responsible for TPM coordination and covered: (1) company context, (2) TPM in practice, (3) human, (4) digital, (5) green and (6) future TPM5.0 perspective. The protocol was informed by TPM implementation models (Ahuja and Khamba, 2008; Ananth and Vinayagam, 2015), human-oriented TPM frameworks (Bashar *et al.*, 2020), digital TPM approaches (Lucantoni *et al.*,

Downloaded from http://www.emerald.com/bij/article-pdf/33/13/27/11604746/bij-11-2025-1075en.pdf by Universita Politecnica Delle Marche u



Source(s): Authors' own creation

Figure 1. HDG-TPM5.0 maturity framework

2025; Seedao *et al.*, 2025; Sivakumar *et al.*, 2022) and sustainability-oriented maintenance models (Chen *et al.*, 2019; Crosby and Badurdeen, 2022).

The cross-sector sample strengthened the identification of recurring practices and variability in HDG integration. The selection of eight companies ensures analytical generalisability, consistent with qualitative research standards: Eisenhardt (1989) recommends 4–10 cases as sufficient for constructing robust conceptual constructs in theory-building research, whilst (Yin, 2018) distinguishes between statistical and analytical generalisability; the latter involves the logical extension of theoretical constructs derived from a targeted set of cases and constitutes the appropriate criterion for studies developing a framework of this type. Maximum variation purposive sampling (Patton, 2002) was used, encompassing five NACE Rev.2 macro-sectors and three company size classes (small, medium and large according to the EU definition of SMEs), to capture differences in sector-specific priorities, digital maturity and resource constraints.

Downloaded from http://www.emerald.com/bij/article-pdf/33/13/27/11604746/bij-11-2025-1075en.pdf by Universita Politecnica Delle Marche u

Table 2. Companies surveyed

ID	Sector (NACE Rev.2)	Role of respondent	Employees	TPM level self-assessment	Key TPM pillars	Distinctive features
A	C10–C12 Food, Beverage, Tobacco	Maintenance Manager	341	Intermediate	Autonomous, Planned, Quality	Digital CMMS, strong operator involvement, energy monitoring
B	C19–C21 Pharmaceutical	Maintenance Director	491	Advanced	Planned, Quality, Education	Predictive maintenance, IoT integration, GMP-oriented TPM
C	C26–C28 Automotive	Maintenance Manager	184	Intermediate	Autonomous, Planned	Sensor-based maintenance, OEE-driven planning
D	C13–C15 Textile	Production Engineer	402	Early–Intermediate	Autonomous, Planned	Training initiatives, moderate digital maturity
E	C19–C21 Chemical	Technical Director	277	Intermediate	Planned, Quality	Simulation-based planning, environmental KPI tracking
F	Transportation and Storage (Logistics)	Maintenance Supervisor	110	Early	Autonomous, Planned	CMMS adoption, culture of continuous improvement
G	C22–C25 Metalworking	Maintenance Manager	65	Early	Autonomous	Manual TPM logs, low digitalisation
H	Electricity, Gas, Steam, Renewable Energy	Plant Manager	366	Intermediate–Advanced	Planned, Quality, Safety	Predictive analytics, energy efficiency projects

Source(s): Authors' own creation/work

The framework's validity is supported by content validity (grounded in the literature review) and face validity, confirmed through expert review and feedback from participating companies regarding the clarity and usefulness of the questionnaire and the results.

3.1.2 Empirical analysis. A triangulation approach (Denzin, 2012; Jick, 1979), combined practitioner interviews, the scientific literature and the HDG–TPM conceptual model to translate qualitative evidence into quantitative questionnaire items (Table 3). The procedure followed four steps: (1) thematic analysis of interview transcripts to extract TPM themes (Braun and Clarke, 2006; Fereday and Muir-Cochrane, 2006; Sovacool *et al.*, 2023); (2) cross-validation with literature; (3) mapping of themes to TPM pillars and HDG dimensions and (4) coding into measurable items. For example, "Operators perform daily checks" was operationalised as "Frequency of operator involvement in daily maintenance and cleaning", mapped to autonomous maintenance within the human dimension (AM–H), with positive scoring direction and ID1. These specifications populate the tool's Knowledge dataset (Section 3.2.1).

3.1.3 Questionnaire design. The three HDG dimensions are operationalised as latent constructs, each measured through behaviourally anchored items mapped to TPM pillars. This multi-level structure links observable item-level behaviours to pillar-dimension scores to construct-level HDG indices, ensuring that constructs are grounded in assessable practices rather than self-reported perceptions.

Downloaded from http://www.emerald.com/bij/article-pdf/33/13/27/11604746/bij-11-2025-1075en.pdf by Universita Politecnica Delle Marche u

Table 3. From qualitative evidence to quantitative items

Thematic analysis Empirical evidence from interviews	Cross validation Literature grounding	Conceptual mapping TPM pillar	HDG dimension	Coding Measurable items	Direction	ID
"Operators perform daily checks and minor cleaning"	Ahmed *et al.* (2023)	Autonomous Maintenance (AM)	Human	Frequency of operator involvement in Daily checks and cleaning	Optimal	1
"Competence matrix updated every six months"	Oeij *et al.* (2024)	Training and Education (T&E)	Human	Frequency of operator training in preventive maintenance tasks	Positive	2
"IoT sensors monitor vibration and temperature"	Mouhib *et al.* (2025), Zero *et al.* (2024)	Planned Maintenance (PM)	Digital	IoT sensors for health monitoring equipment (vibration, temperature)	Positive	3
"CMMS schedules maintenance based on cycles"	Hetmanczyk (2024)	Planned Maintenance (PM)	Digital	CMMS integration for maintenance planning, tracking, and reporting	Positive	4
"Replacing old compressors reduced energy use by 12%"	Jin *et al.* (2024)	Preventive Maintenance (PM)	Green	Energy monitoring and retrofitting in maintenance operations	Positive	5
"We use biodegradable cleaners and track waste"	Farajzadeh *et al.* (2024)	Preventive Maintenance (PM)	Green	Green maintenance materials and waste tracking	Positive	6

Source(s): Authors' own creation/work

(1) The H dimension includes items on operator training (E and T, AM), ergonomics and safety practices (SHE), operator autonomy and involvement (AM, FI), and inclusiveness of training and improvement programmes (E and T).

(2) The D dimension includes use of IoT and real-time monitoring (PM), CMMS and digital planning tools (PM, O-TM), predictive analytics (PM) and cross-functional digital tool use (O-TM).

(3) The G dimension includes energy monitoring in maintenance (PM), eco-efficiency in inspection and quality routines (FI), sustainable materials and waste minimisation (PM) and eco-design in equipment procurement (EEM).

All items are scored on a five-point scale with behaviourally defined levels (presence/absence of practices, tools or procedures), reducing interpretation variability and improving comparability across firms (Hajoary *et al.*, 2024).

The resulting questionnaire includes:

(1) Company profile, capturing contextual variables for clustering and benchmarking (NACE Rev.2 sector, EU SME size classification, TPM duration and involved personnel);.

(2) Rating-scale matrix, using a 5-point Likert scale (1 = Not implemented, 5 = Fully).

(3) Open comments, capturing qualitative feedback for refinement and updating.

Downloaded from http://www.emerald.com/bij/article-pdf/33/13/27/11604746/bij-11-2025-1075en.pdf by Universita Politecnica Delle Marche u

Items are designed as granular, diagnostically focused indicators, each capturing a single observable practice with a clear interpretation. Each item is linked to one HDG dimension and TPM pillar, while pillars include multiple dimensions and dimensions span multiple pillars.

This design ensures multidimensional assessment emerges at the aggregate level, without requiring additional weighting across dimensions or pillars, thus reducing subjectivity and improving transparency and comparability.

3.1.4 TPM5.0 maturity classes. The final maturity level is assigned using six classes (Sajjad *et al.*, 2024): Outsider [0, 10]%, Beginner]10, 20]%, Intermediate]20, 40]%, Experienced]40, 60]%, Expert]60, 80]% and Top performer]80, 100]%.

3.2 Maturity assessment (B)

The assessment includes three modules: *(1) database and preprocessing* (Section 3.2.1), *(2) maturity assessment* (Section 3.2.2) and (3) *output visualisation* for benchmarking (Section 3.2.3).

Conceptually, the pipeline operates as follows: (1) questionnaire responses are converted into utilities $u_{ij} \in [0, 1]$; (2) utilities are stored in a central dataset organised by sector–size clusters; (3) each item receives a contextual weight combining HDG, TPM and diagnostic components; (4) MABAC computes maturity as a weighted deviation from cluster benchmark and (5) good practices are identified from peers.

3.2.1 Module 1: databases and questionnaire preprocessing. Module 1 standardises questionnaire inputs for contextual benchmarking. Specifically, three datasets support the process: a *Knowledge Dataset* (versioned), a *Central Dataset* (dynamic) and a *Contextual Dataset* (adaptive).

(1) *Knowledge Dataset* (Table 4) *and utility conversion.* For each item j, the dataset includes: item ID, TPM pillar t_j, HDG dimension d_j, diagnostic leverage coefficient q_j, scoring direction and Likert interpretation.

Item tags (t_j, d_j) derived from the triangulation process (Section 3.1), while q_j and scoring rules are defined by TPM and I5.0 experts. The coefficient q_j captures the discriminative power of item j and is bound to preserve interpretability:

$$q_j \in \{q^B, q^S, q^A\}, 0 < q^B < q^S < q^A \tag{1}$$

Table 4. Knowledge dataset extraction

Item ID	TPM pillar	HDG dimension	Indicator q_j	Direction	Likert	Interpretation
1	AM	H	q^S	Optimal	1	Low autonomy
					2	Reactive maintenance remains reactive
					3	Balanced and efficient routine (*optimal*)
					4	Excessive frequency
					5	Inefficient or stressful
2	T&E	H	q^S	Positive	1	Competencies not tracked
					...	...
					5	Demonstrable skill transfer
3	PM	D	q^A	Positive	1	Manual inspection only
					...	...
					5	Continuous condition monitoring

Source(s): Authors' own creation/work

Downloaded from http://www.emerald.com/bij/article-pdf/33/13/27/11604746/bij-11-2025-1075en.pdf by Universita Politecnica Delle Marche u

The three tiers (basic/standard/advanced) are ordinal and do not imply equidistant maturity intervals. For instance, within the digital dimension and the PM pillar, Excel-based maintenance tracking represents a basic stage, whereas AI/ML-based predictive tools represent an advanced stage and are therefore assigned with higher leverage. that the coefficient q_j is not the final item weight, but is later combined with contextual parameters (Eqs. 4–8) to obtain effective weights.

To ensure comparability, each Likert response $x \in \{1, 2, 3, 4, 5\}$ is mapped to a utility $u_{ij} \in [0, 1]$. For positive-direction items, higher scores indicate higher maturity; for optimal-direction items, maturity peaks at a target $x^{\star}$:

$$\boldsymbol{u}(\boldsymbol{x}) = \begin{cases} (\boldsymbol{x} - \mathbf{1})/\mathbf{4}, & \text{positive direction} \\ \mathbf{1} - |\,\boldsymbol{x} - \boldsymbol{x}^{\star}|\,/\mathbf{2}, \;, & \text{optimal direction} \end{cases} \tag{2}$$

(2) *Central dataset and clustering*. For each company i, the utility vector $\mathrm{u}_i = [u_{ij}]$ is stored in the central dataset. Records are grouped into sector-size clusters C_k to ensure benchmarking against comparable peers:

$$\boldsymbol{C_k} = \{\, \boldsymbol{i} |\ (\boldsymbol{s_i}, \boldsymbol{z_i}) \in \Omega_{\boldsymbol{k}} \,\}, k = (sector\, s, size\, z) \tag{3}$$

Within each cluster, data are stored as a company-by-items matrix.

(3) *Contextual weighting dataset*. This dataset calibrates item weights for heterogeneous contexts. First, sector-based TPM pillar weights $\mathrm{w}^{TPM}_{(\mathrm{s,t})}$ define pillar relevance:

$$\sum\nolimits_{t \in \{AM, PM, QM, EEM, T\&E, SHE, FI, O-TM\}} \mathrm{w}^{TPM}_{(\mathrm{s,t})} = 1 \tag{4}$$

Second, sector-based HDG weights $w^{HDG}_{(s,d)}$ define dimension relevance:

$$\sum\nolimits_{d \in \{H, D, G\}} \mathrm{w}^{HDG}_{(\mathrm{s,d})} = 1 \tag{5}$$

Third, size effects are captured by a correction vector:

$$\Delta_{\mathrm{z}} = \left(\Delta_{(\mathrm{z},H)}, \Delta_{(\mathrm{z},D)}, \Delta_{(z,G)}\right) \tag{6}$$

Size-adjusted HDG weights are:

$$w'^{\mathrm{HDG}}_{(k,d)} = w^{HDG}_{(\mathrm{s},d)} * \Delta_{(\mathrm{z},d)}, \forall\, d \in \{H, D, G\} \tag{7}$$

and normalised:

$$\widetilde{w}^{HDG}_{(k,d)} = \frac{w'^{\mathrm{HDG}}_{(k,d)}}{\sum\limits_{d' \in \{H,D,G\}} w'^{\mathrm{HDG}}_{(k,\mathrm{d}')}}, \sum_{d} \widetilde{w}^{HDG}_{(k,d)} = 1 \tag{8}$$

3.2.2 Module 2: maturity assessment. Module 2 quantifies TPM5.0 maturity via peer benchmarking within clusters $k = (s, z)$. A multiple-criteria decision-making approach is adopted; MABAC is selected for its transparency and ability to express maturity as deviation from a reference level (Torkayesh *et al.*, 2023).

Step 1: Cluster normalisation. Utilities are rescaled within cluster C_k using min–max normalisation:

Downloaded from http://www.emerald.com/bij/article-pdf/33/13/27/11604746/bij-11-2025-1075en.pdf by Universita Politecnica Delle Marche u

$$N_{ij} = \frac{u_{ij} - \min(u_j)}{\max(u_j) - \min(u_j)} \tag{9}$$

where $min\,(u_j)$ and $max(u_j)$ are computed across firms in C_k. Values above (below) 0.5 indicate above- (below-)average performance.

Step 2: Contextual item weighting. Each item j is assigned a contextual weight that combines strategic relevance (HDG), operational relevance (TPM) and diagnostic leverage:

$$W_{k,j} = \widetilde{w}^{HDG}_{k,d_j} * w^{TPM}_{s,t_j} * q_j \tag{10}$$

Step 3: Weighted performance. Normalised values are weighted to obtain:

$$G_{ij} = N_{ij} * W_{k,j} \tag{11}$$

Step 4: Cluster benchmark. For each item j, the reference level is the cluster mean (with n firms in C_k):

$$L_j = \frac{1}{n} \sum_{i=1}^{n} G_{ij} \tag{12}$$

Step 5: Deviation from benchmark. Firm $i's$ deviation on item j is:

$$Q_{ij} = G_{ij} - L_j \tag{13}$$

with $Q_{ij} > 0$ indicating above-benchmark performance.

Step 6: Raw maturity deviation. Deviations are aggregated across the m questionnaire items:

$$Q_i = \sum_{j=1}^{m} Q_{ij} \tag{14}$$

Step 7: Maturity index and class. To preserve longitudinal comparability, each cluster uses fixed calibration bounds (Q_{min}, Q_{max}) estimated in the pilot phase. The maturity index is computed as:

$$Index_i = \frac{\mathrm{Q_i} - \mathrm{Q_{min}}}{\mathrm{Q_{max}} - \mathrm{Q_{min}}} * 100 \tag{15}$$

and then mapped to the six maturity classes defined in Section 3.1.4.

3.2.3 Module 3: benchmarking and recommendations. Module 3 translates results into an interpretable benchmarking narrative, combining maturity index, class and HDG/pillar-level diagnostics. Benchmarking is performed within sector-size clusters using aggregated data, ensuring comparability and confidentiality.

(1) *Strategic view (HDG radar)*. Dimension-level maturity is obtained by aggregating item-level deviations within each HDG dimension d and comparing them to the cluster average:

Downloaded from http://www.emerald.com/bij/article-pdf/33/13/27/11604746/bij-11-2025-1075en.pdf by Universita Politecnica Delle Marche u

$$Q_{(i,d)} = \sum_{j \in d} Q_{ij} \tag{16}$$

This view highlights whether maturity constraints originate primarily from Human routines, digital maintenance capabilities or green/resource-efficiency practices.

(2) *Diagnostic view (pillar in the HDG deviation space).* For each TPM pillar t, deviations are aggregated along the three HDG dimensions d (Eq. 17) and expressed relative to the corresponding cluster centroid (Eq. 18):

$$Q_{(i,t,d)} = \sum_{j \in (t,d)} Q_{ij} \tag{17}$$

$$\Delta_{(i,t,d)} = Q_{(i,t,d)} - \overline{Q}_{(k,t,d)} \tag{18}$$

Pillars near the origin align with peers; negative deviations indicate weaknesses. A colour scale conveys deviation intensity.

Once weak pillars and dimensions are identified, benchmarking against top performers highlights gap drivers and best practices within cluster C_k. Recommendations follow the same decomposition logic, linking HDG signals to pillar-level actions aligning with the TPM5.0 emphasis on explainability and adaptability. To support interactive exploration, a chatbot can provide clarifications on scores, benchmarks and recommended actions.

3.3 Framework updating and refinement

To ensure long-term relevance, the framework includes a structured updating mechanism combining dynamic weight update module (C) and continuous monitoring module (D), allowing indicators to evolve in response to technological and organisational changes. Data-driven process based on quantitative signals (e.g. weight dynamics, dispersion) and qualitative signals (NLP-detected trends) inform annual updates of q_j and periodic (3–5 years) revision of questionnaire and contextual weights. Any expert-approved changes are introduced through controlled roll-out and documented versioning.

3.3.1 Dynamic weight update mechanism (C). Item leverage q_j is updated to reflect periodic diffusion and relevance through an expert-mediated procedure. The rationale is that q_j should be downgraded when the practice captured by item j becomes mainstream and ceases to differentiate firms and upgraded when it captures emerging practices that still provide strong discriminatory power. Importantly, the mechanism supports experts by producing data-driven proposals for tier transitions over time (e.g. standard → basic). Each year, three signals are computed from item's response:

(1) *Diffusion tracking* (quantile-based): This signal down-weights items that become widely adopted and up-weights emerging differentiators. Let $F_j(r^*)$ be the share of firms that reach an "advanced adoption" threshold r^*. The diffusion adjustment is:

$$\widehat{\boldsymbol{q}}_j^{DIFF} = \boldsymbol{q}_j^t \left[\mathbf{1} + \boldsymbol{\alpha} \left(\mathbf{1} - \boldsymbol{F}_j(\boldsymbol{r}^*) \right) \right] \tag{19}$$

where $\boldsymbol{q}_j^t$ is the previous-year leverage weight and $\boldsymbol{\alpha}$ controls sensitivity (0.2–0.3). If most firms already implement the practice captured by item $\boldsymbol{j}$ (e.g. a CMMS), its relevance tends to decrease because it no longer separates firms; conversely, less-diffused practices (e.g. AI-based predictive maintenance) may gain relevance.

Downloaded from http://www.emerald.com/bij/article-pdf/33/13/27/11604746/bij-11-2025-1075en.pdf by Universita Politecnica Delle Marche u

(2) *Information content* (discriminative power): Diffusion alone may not capture whether an item still separates firms; therefore, a second signal measures response variability. Items with homogeneous responses carry limited information and are down-weighted, whereas items with heterogeneous adoption remain informative:

$$\hat{q}_j^{ENT} = q_j^t\left(1 + \beta \cdot EntropyGain_j\right) \tag{20}$$

where $EntropyGain_j$ quantifies response variability for item j, and β is typically 0.1–0.3.

(3) *XAI relevance*: A third signal uses explainability to quantify how much each item contributes to maturity classification in the previous-year models. Items with higher contribution are candidates for higher weights:

$$\hat{q}_j^{XAI} = q_j^t\left[1 + \gamma \parallel SHAP_{kj}\parallel_{norm}\right] \tag{21}$$

where $\parallel SHAP_j\parallel_{norm}$ is the normalised SHAP importance and $t\,\gamma \leq 0.2$ bounds the impact of XAI signal.

The three signals are combined into a single multiplicative proposal:

$$q_j^{prop} = q_j^t \cdot \left(1 + \alpha\left(1 - F_j(r^*)\right) * \left(1 + \beta\, EntropyGain_j\right) * \left(1 + \gamma \parallel SHAP_j\parallel_{norm}\right)\right. \tag{22}$$

Proposed updates are validated by experts and smoothed to prevent abrupt year-to-year changes:

$$q_j^{(t+1)} = (1 - \lambda)\, q_j^t + \lambda\, q_j^{approved}, \lambda \leq 0.3 \tag{23}$$

whit λ enforcing gradual evolution. After approval, the updated value is expressed as an ordinal tier (basic/standard/advanced).

3.3.2 Continuous monitoring (D). Open-ended responses are analysed to support structured feedback and guide expert-led questionnaire refinement without affecting scores. These questions are intentionally included to capture insights, emerging practices and contextual information not fully represented by structured items.

The NLP layer acts as a support tool for qualitative interpretation and continuous improvement. It does not directly influence maturity scoring; instead, it produces reviewable evidence to enhance understanding, detect emerging needs and support periodic questionnaire updates. The analysis is structured as follows:

(1) *Keyword tagging:* key terms are extracted and displayed as tags to speed up screening (e.g. “Power BI”, “REST API”, “MES”).

(2) *Semantic clustering*: responses are grouped into standardised categories to reduce linguistic variability and highlight recurring themes.

(3) *Drift detection*: changes in terminology and topic prevalence are monitored over time to identify emerging practices (e.g. “5G-enabled edge computing”, “generative-AI”).

(4) *Classification:* responses are classified as existing or emerging practices to structure the review process.

(5) *Recommendations panel:* the system suggests targeted updates (e.g. refining item wording or identifying candidate topics for future questionnaire inclusion).

Downloaded from http://www.emerald.com/bij/article-pdf/33/13/27/11604746/bij-11-2025-1075en.pdf by Universita Politecnica Delle Marche u

(6) *Expert validation and retraining:* all NLP outputs are reviewed by domain experts, who approve, refine or reject suggestions; validated annotations are used to periodically retrain the model, improving consistency.

The insights generated provide qualitative evidence complementing the quantitative signals used in layer C. They support expert interpretation of emerging practices and potential item obsolescence without directly modifying model parameters.

4. Illustrative application of the framework

This section provides an end-to-end illustration of the HDG–TPM5.0 maturity framework using Company-C07, a small firm in the Food, Beverage and Tobacco sector (NACE C10–C12). C07 is assigned to the Food-Small peer cluster $k = (s = \text{Food}, z = \text{Small})$. The purpose is to demonstrate the workflow from preprocessing to maturity index computation, class assignment and benchmark-driven recommendations.

4.1 Data preprocessing

Company C07 completed the full questionnaire. For each item j: (1) the responses were linked to the Knowledge dataset (pillar t_j, dimension d_j, direction/target, Likert interpretation and leverage weight q_j); and (2) the Likert score $x \in \{1, \ldots, 5\}$ was converted into a utility value $u_{ij} \in [0, 1]$ using Eq. (2). Most items are positively oriented; for optimal-direction items, the target is $x^\star = 3$. In this pilot case, leverage weights are set as $(q^B, q^S, q^A) = (0.8, 1.0, 1.2)$. Table 5 reports an extract of this step. The resulting utility vector for C07 is $u_{C07} = (0.50, 1.00, 0.75, \ldots)$.

Utilities are then appended to the cluster-specific matrix U, which stores responses from firms in the same sector–size cluster (C07 and six peers). Rows represent companies, columns items and each cell stores u_{ij}:

$$\mathbf{U_{i,j}} = \begin{pmatrix} \mathbf{0,5} & \mathbf{0,75} & \mathbf{0,5} & \ldots \\ \mathbf{0,75} & \mathbf{1,00} & \mathbf{0,75} & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ \mathbf{0,5} & \mathbf{1,00} & \mathbf{0,75} & \ldots \end{pmatrix}$$

4.2 Contextual weighting

Contextual parameters are retrieved from the contextual dataset, combining sector priorities (HDG and TPM pillar weights) and size correction factors. For the Food-Small cluster:

(1) Size correction vector (Small): $\Delta_{\text{small}} = (1{,}05;\ 1{,}00;\ 0{,}95)$

(2) Sector-based HDG weights (Food): $\mathrm{w}_{Food}^{HDG} = (0,45;\ 0,30;\ 0,25)$

Table 5. Preprocessing (extract)

New INPUT		Matching METADATA (from the knowledge dataset)						Utility TRANSFORMATION
Item	Likert	Interpretation	Pillar	Dim.	Direction	$x^\star$	q_j	Utility u_{ij} (0–1)
ID1	4	Excessive frequency	AM	H	optimal	3	1,0	$1 - \frac{\lvert 4-3 \rvert}{2} = 0.50$
ID2	5	Continuous training	T&E	H	positive	–	1,0	$\frac{5-1}{4} = 1.00$
ID3	4	Analytical capability	PM	D	positive	–	1,2	$\frac{4-1}{4} = 0.75$
...	...	...	...	...	...	...	...	...

Source(s): Authors' own creation/work

Downloaded from http://www.emerald.com/bij/article-pdf/33/13/27/11604746/bij-11-2025-1075en.pdf by Universita Politecnica Delle Marche u

(3) Sector-based TPM pillar weights (Food): $w_{Food}^{TPM} = (0,14;\ 0,12;\ 0,13;\ 0,10;\ 0,10;\ 0,19;\ 0,13;\ 0,09)$

Following Eq. (7), size-adjusted HDG weights are:

$w'_{(\text{Food}|\text{Small},\mathbf{H})} = 0{,}45*1{,}05 = 0{,}4725$; $w'_{(\text{Food}|\text{Small},\mathbf{D})} = 0{,}30*1{,}00 = 0{,}3000$; $w'_{(\text{Food}|\text{Small},\mathbf{G})} = 0{,}25*0{,}95 = 0{,}2375$

Their sum is 1.01; thus, normalised weights (Eq. 8) are:

$$\widetilde{w}^{HDG}_{(Food|Small,H)} = 0,4673;\ \widetilde{w}^{HDG}_{(Food|Small,D)} = 0,2970;\ \widetilde{w}^{HDG}_{(Food|Small,G)} = 0,2356$$

Hence, $\widetilde{w}^{HDG}_{Food|Small} \approx (0,467;\ 0,297;\ 0,236)$

4.3 Maturity assessment (worked example on item ID3)

To illustrate the intra-cluster MABAC logic, the full computation is shown for item ID3, classified as digital-planned maintenance (D-PM).

Step 1 Cluster normalisation (Eq. 9)*:* In the Food|Small cluster, for ID3*:* $\min(u) = 0,25$, $\max(u) = 1,00$ and $u_{C07,ID3} = 0,75$.

$N_{C07,ID3} = \frac{0{,}75-0{,}25}{1{,}00-0{,}25} = 0,667$

Step 2 Contextual item weight (Eq. 10)*:* For D–PM: $\widetilde{w}^{HDG}_{(Food|Small,D)} = 0,297$, $w^{TPM}_{(Food,PM)} = 0,12$ and $q_{ID3} = 1,2$.

$W_{ID3} = 0,297 * 0,12 * 1,2 = 0{,}0428$

Step 3 Weighted value (Eq. 11):

$G_{C07,ID3} = 0,667 * 0,0428 = 0,0285$

Step 4 Cluster reference level (Eq. 12)*:* With $n = 7$ firms in the cluster:

$\boldsymbol{L_{ID3}} = \frac{1}{7}(\mathbf{0,0178} + \mathbf{0,0234} + \ldots + \mathbf{0,0285}) = 0{,}0260$

Step 5 Deviation (Eq. 13):

$Q_{C07,ID3} = 0{,}0285 - 0{,}0260 = +\,0{,}0025$

Thus, C07 performs slightly above the Food|Small cluster benchmark on this specific D–PM item (ID3).

4.4 Maturity score and class assignment

Applying the procedure to all items yields the deviation profile $Q_{(C07,j)}$. Table 6 reports deviations aggregated by pillar and HDG dimension, supporting both analytical results and the visual diagnostics.

This matrix provides the basis for computing the maturity index and benchmarking visualisations. Specifically, it reports: (a) the overall deviation $Q_{C07} = 0,0078$ (Eq. 14); (b) HDG-level deviations (Eq.16), e.g. $Q_{(C07,H)} = 0,0034$ and (c) pillar-dimension deviations (Eq. 17), e.g. $Q_{(C07,AM,H)} = 0{,}0009$.

As this is a pilot application, calibration bounds $(Q_{\min}, Q_{\max})$ for the Food-Small cluster will be finalised with a larger sample. For illustrations, provisional bounds are used:

$$Q^{pilot}_{\min} = -0,0001, Q^{pilot}_{\max} = 0,0109.$$

The maturity index is computed via min–max rescaling (Eq. 15):

Downloaded from http://www.emerald.com/bij/article-pdf/33/13/27/11604746/bij-11-2025-1075en.pdf by Universita Politecnica Delle Marche u

Table 6. Deviation profile for C07

TPM pillar	Human	Digital	Green	Total per pillar
AM	0,0009	0,0006	0,0002	0,0017
PM	0,0003	0,0012	0,0002	0,0017
QM	0,0004	0,0007	0	0,0011
MP	0,0002	0,0004	−0,0009	0,0006
FI	0,0003	0,0004	−0,0002	0,0007
EEM	0,0006	0,0005	0,0001	0,0012
T&E	0,0005	0,0001	0	0,0006
O-TM	0,0002	0	−0,0001	0,0002
TOTAL per dimension	*0,0034*	*0,0039*	*0,0005*	*0.0078*

Source(s): Authors' own creation/work

$$Index_{C07} = \frac{0,0078 - (-0,0001)}{0,0109 - (-0,0001)} * 100 \approx 72\%$$

According to the taxonomy, $72\% \in]60, 80]$, thus C07 is classified as Expert.

4.5 Output visualisation, benchmarking and recommendations

The output layer reports the analytical score (index and class) and two views (Figure 2):

(1) *Strategic view (a)*: the HDG radar shows an asymmetric profile: Human maturity aligns with the Food-Small benchmark, while Digital and Green are below average.

(2) *Diagnostic view (b):* the 3D-HDG space highlights three critical pillars (SHE, PM, FI) in red. SHE and FI are mainly displaced along the negative G axis, while PM lies mainly in the negative D region. EEM and O-TM are moderately weak (yellow), with negative D and G partially offset by positive H. Similar in D and G dimensions, QM is moderate (orange), with an H near the cluster mean (coordinate close to zero). T&E is near the centroid, indicating a balanced performance, while AM emerges as a relative strength across all dimensions.

Benchmarking against cluster frontrunners clarifies gap drivers. In Food-Small, digital leaders combine CMMS integration, sensor-based condition monitoring and predictive analytics, while Green leaders, rely on energy dashboards, eco-efficiency routines and environmental Key Performance Indicators (KPIs). Consistent with C07's profiles, two priorities emerge: strengthening the instrumentation layer (sensing, monitoring, data collection) and the analytical layer (predictive tools and sustainability metrics).

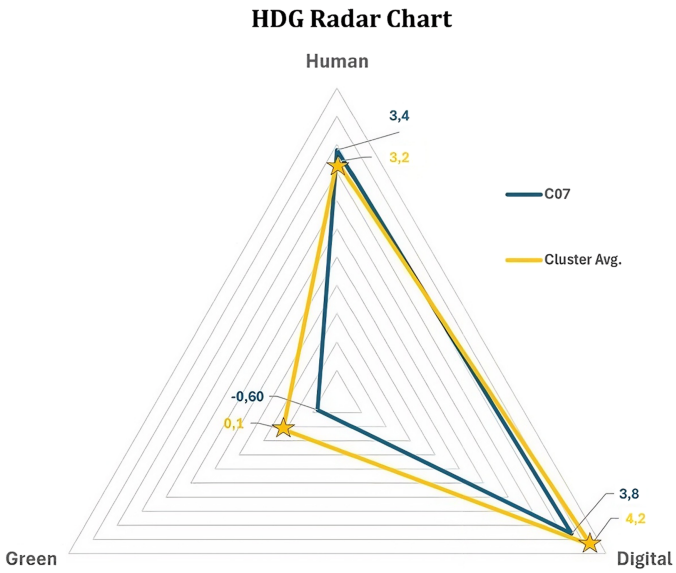
Accordingly, recommendations are expressed as follows: *Given the current position in the digital dimension, particularly within the PM pillar, priority actions include deploying a connected CMMS, implementing IoT sensing on critical assets and adopting basic predictive analytics, which cluster frontrunners use as primary drivers. In addition, given the negative green deviations in SHE and FI, energy and waste mapping tools, eco-efficiency projects and dashboards tracking key environmental KPIs represent the practices that distinguish environmental leaders in the Food|Small cluster.*

4.6 Inputs for updating and refinement

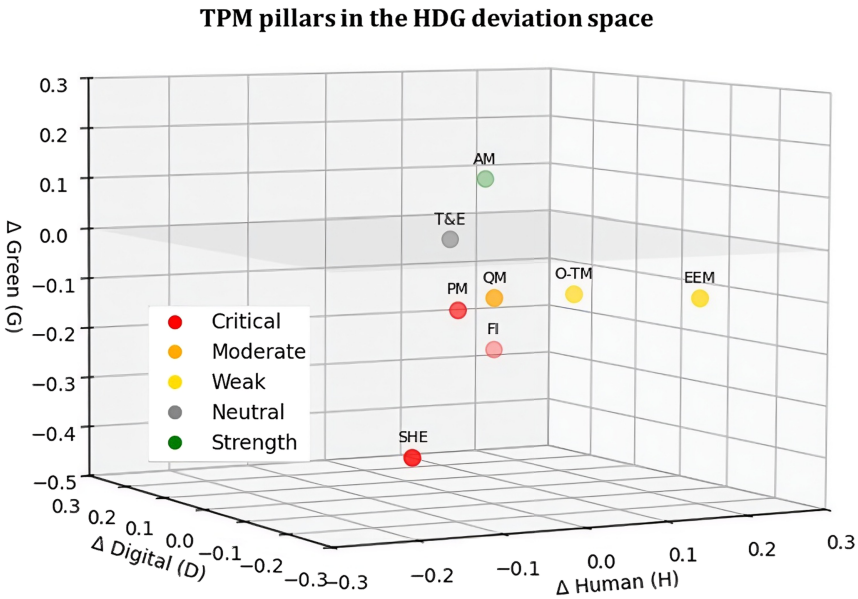
Applying the framework to C07 feeds both the quantitative adaptation loop and qualitative learning loop (Sections 3.3.1-3.3.2).

At the *quantitative level*, adding C07 updates cluster statistics and influences the annual q_j review (Table 7). In this case, digital–PM (ID3) and green–FI (ID18) remain differentiating

Downloaded from http://www.emerald.com/bij/article-pdf/33/13/27/11604746/bij-11-2025-1075en.pdf by Universita Politecnica Delle Marche u



(a)



(b)

Source(s): Authors' own creation

Figure 2. Output visualisation

Downloaded from http://www.emerald.com/bij/article-pdf/33/13/27/11604746/bij-11-2025-1075en.pdf by Universita Politecnica Delle Marche u

Table 7. Impact of C07 on weights

Item ID	Dimension–Pillar and item focus	Effect on cluster percentiles (α)	Effect on entropy (β)	Effect on SHAP relevance (γ)	Implication for q_j^{t+1}
ID3	D-PM: use of IoT devices and real-time sensors for equipment health monitoring	Widens the gap between upper and lower quantiles due to C07's low adoption	High dispersion preserved	High SHAP, as it strongly contributes to the Digital deficit	Potential weight increase (emerging differentiator)
ID5	G-MP: energy monitoring and improvement actions in maintenance operations	Reinforces a low lower-quartile, confirming systematic underperformance	Moderate entropy (stable weakness across firms)	Moderate SHAP as a persistent negative driver	Slight weight increase (still discriminative)
ID12	D-OTM: cross-functional use of digital tools in operations and teamwork	Convergence of scores around the upper quantiles	Entropy decreases (practice becoming standard)	Low SHAP, as it no longer explains maturity differences	Likely weight decrease (no longer informative)
ID18	G-FI: integration of eco-efficiency criteria in inspection and quality routines	Increases heterogeneity in the central and upper quantiles	Entropy increases (diverging adoption patterns)	High SHAP for explaining the Green deficit	Potential weight increase (key indicator of G maturity)
ID22	H-T&E: frequency and coverage of operator training and education in maintenance-related tasks	Upper quantiles further consolidated at high values	Low entropy (uniform good practice across the cluster)	Low SHAP, as it rarely drives maturity differences	Likely weight decrease (mature, non-discriminative item)

Source(s): Authors' own creation/work

practices, while widely adopted items (e.g. human–T&E, ID22 or common digital O-TM tools) lose discriminative power and are down-weighted candidates.

At the *qualitative level*, open-ended responses enter the NLP monitoring cycle. Signals such as "energy dashboard prototypes" and "sensor calibration issues" are clustered across Food-Small firms, revealing: (1) diffusion of early-stage environmental dashboards (supporting future refinement of G-related items) and increasing reliance on sensing infrastructures (reinforcing drift detection in digital terminology). After expert validation, these signals inform next questionnaire updates and the refinement of contextual weights.

4.7 Results

The pilot illustrates how the framework translates questionnaire responses into benchmarked and actionable maturity insights. C07 shows an asymmetric HDG profile: Human practices align with peer-cluster reference, while digital and green maturity lag. The HDG radar highlights these imbalances, and the 3D pillar representation further localises them across TPM pillars: PM drives the digital deficit, while SHE and FI account for most of the digital and green shortfall, indicating relatively strong people-related routines despite gaps in digitalisation and eco-efficiency.

Consistent with benchmarking, actions target gap-driving pillars, prioritising digital enablement in PM and eco-efficiency routines in SHE/FI. Priority actions include: (1) in digital (PM), deploying connected CMMS, IoT sensing and predictive analytics and (2) in

Downloaded from http://www.emerald.com/bij/article-pdf/33/13/27/11604746/bij-11-2025-1075en.pdf by Universita Politecnica Delle Marche u

green (SHE, FI), energy and waste mapping tools, eco-efficiency projects and environmental KPI dashboards.

From a managerial perspective, these results indicate that firms should prioritise strengthening the digital infrastructure of maintenance systems while integrating environmental practices into routine operations. The framework supports this process by linking HDG gaps to specific pillars, enabling managers to allocate resources more effectively and align improvement initiatives with peer-based benchmarks.

A sensitivity analysis was conducted on the contextual weighting structure. For HDG weights, one-at-a-time perturbations of ±20% show limited variation in the raw maturity deviation (within approximately ±4.3%), with the largest impact associated with the digital dimension. A similar analysis on TPM pillar weights yields even lower variation (within approximately ±2%), confirming that results are not driven by moderate changes in operational weight allocation. At the item level, weight dynamics are governed by the adaptive update mechanism (Module C), which adjusts leverage based on diffusion, information content and XAI relevance. This further reduces sensitivity to static weight assumptions over time. Overall, these results confirm that moderate variations in the weighting structure do not alter the interpretation of the maturity profile or the identification of priority areas, supporting the robustness of the framework.

The case also illustrates the adaptive layer: differentiating items (e.g. ID3 and ID18) retain relevance, while widely adopted practices progressively lose weight.

5. Discussions

This section discusses the study's contributions in relation to the three research gaps identified in the introduction. Gap (1), namely the absence of a multidimensional TPM maturity model integrating HDG performance across all pillars, is addressed by the framework's core architecture, operationalising TPM maturity as a three-dimensional H×D×G construct mapped onto all pillars and demonstrated in the illustrative application. Gap (2), i.e. the lack of adaptive benchmarking for heterogeneous industrial contexts, is addressed by contextual weighting architecture and cluster-based peer normalisation, enabling cross-firm comparison across sector, size and technological intensities without requiring a common absolute reference. Gap (3), i.e. the limited usability of existing models due to static logic and absence of updating mechanisms, is addressed by the XAI-based explainability layer, the NLP-based monitoring and recalibration cycle, thereby reducing obsolescence risk and improving managerial relevance. The remainder of this section elaborates on these contributions in terms of theoretical significance and practical implications.

5.1 Theoretical contributions

As highlighted in the literature review, TPM maturity research remains limited, with few structured frameworks and none fully integrating digital and environmental dimensions. This gap is not only theoretical but reflects a concrete managerial need, as confirmed by interview evidence, showing demand for integrated tools capturing HDG dimensions simultaneously.

The review of Lean assessment tools highlights substantial heterogeneity in scope and resource requirements, with assessments often relying on context-specific selection or ad hoc integration rather than robust benchmarking (Närvä and Bicheno, 2026). In this context, the proposed HDG–TPM5.0 framework advances TPM maturity research and the broader maturity-model literature by combining peer-cluster benchmarking with mechanisms for controlled model evolution, improving cross-firm comparability and addressing the obsolescence in static tools (Table 8).

The framework contributes to three main ways. First, it conceptualises TPM maturity as a multidimensional construct articulated along human-digital-green axes and mapped onto the TPM pillars, enabling a diagnostic interpretation of whether underperformance relates to

Downloaded from http://www.emerald.com/bij/article-pdf/33/13/27/11604746/bij-11-2025-1075en.pdf by Universita Politecnica Delle Marche u

Table 8. Key differences among TPM maturity models

Features	Shaaban and Awni (2014)	Alidrisi (2020)	Proposed HDG TPM 5.0 model
Maturity levels	3 stages (Pilot, Deployment, Continuous Improvement)	5 levels (Basic to World Class)	5 levels extended to HDG multidimensional structure
Focus areas	TPM pillars	TPM pillars	TPM pillars + HDG dimensions
Digital dimension	Not addressed	Not addressed	Explicitly modelled across all pillars
Human dimension	Limited	Limited (AM only)	Explicitly modelled across all pillars
Green dimension	Not considered	Not considered	Explicitly modelled across all pillars
Scoring logic	Qualitative categorisation	Quantitative per pillar	Three-layer scoring: Overall→ HDG (2D)→ TPM pillars on HDG space (3D)
Normalisation	Implicit; based on qualitative expert judgement	Firm-specific bounds (non-comparable)	Cluster-referenced, comparable across firms
Weighting mechanism	Static	Static fuzzy linguistic weights	Contextual weighting (sector*size*item relevance) + dynamic updates
Benchmarking logic	Not defined (case-based insights only)	Global ranking	Cluster-specific, context sensitive
Explainability	None	None	Integrated
Qualitative input	Not systematically included	None	NLP-driven analysis, drift detection
Operator interface	N/A	N/A	Interactive dashboard and chatbot
Improvement recommendations	General advice	Implicit logic	Data-driven (based on cluster frontrunners' practices), dimension- and pillar-specific

Source(s): Authors' own creation/work

people, technologies or environmental practices. Second, it replaces context-independent importance schemes with contextual weighting architecture $\mathbf{W_{k,j}}$, incorporating sector priorities, pillar relevance and size constraints. Methodologically, it adapts MABAC into a benchmarking-oriented maturity assessment by combining peer-cluster normalisation with contextual weights and item-level leverage, enabling comparison within homogeneous references rather than firm-specific rescaling. Third, it introduces an adaptive layer extending assessment beyond static measurement: diffusion dynamics, entropy-based signals and SHAP relevance support controlled weight updates, while NLP monitoring of open-ended inputs captures emerging practices and terminology shifts, strengthening continuous learning and refinement.

5.2 Practical implications

The HDG–TPM5.0 framework provides a decision-support tool applicable across heterogeneous firms by interpreting maturity relative to sector–size peers, rather than absolute scores. This enables fair benchmarking under different technological and organisational baselines and supports balanced HDG improvement pathways, while aggregated cluster benchmarking preserves data confidentiality, as firms are not required to disclose sensitive information while still benefiting from peer comparison.

A key innovation is the pillar-level HDG visualisation, which translates results into an actionable diagnostic by positioning each TPM pillar in the HDG space and highlighting dimension-specific gaps. This supports coordinated interventions (e.g. combining training, digital tools and sustainability practices) and facilitate prioritisation and resource allocation.

Downloaded from http://www.emerald.com/bij/article-pdf/33/13/27/11604746/bij-11-2025-1075en.pdf by Universita Politecnica Delle Marche u

Adoption is strengthened by explainability and learning layers. Explainability clarifies item-level drivers, improving interpretability and trust, while the monitoring layer processes open-ended inputs to detect emerging practices and support repeated use over time, reducing managerial effort and obsolescence. Weights recalibration is implemented as a centralised periodic process managed by the benchmark administrator and does not impose additional burden on firms that interact exclusively with the questionnaire. The framework remains therefore scalable for SMEs and large enterprises. AI, NLP and XAI operate only as supporting infrastructure at the monitoring and explainability levels (sections C and D of Figure 1) and do not influence HDG scores. The HDG balance is ensured through three independent weighting flows and maturity scores, and the contextual weighting architecture can amplify the H and G where digital maturity is already high, counteracting any tendency for the digital dimension to dominate. The 3D pillar-level visualisation makes dimension-specific gaps directly actionable, ensuring that high digital performance does not mask human or environmental weakness, allowing improvement priorities for H and G, not for digital. This puts into practice the fundamental principle of I5.0, according to which the digital and green transitions must reinforce rather than overshadow the human dimension (Xu *et al.*, 2021).

Beyond managerial use, the framework supports executive education and policy design by enabling evidence-based benchmarking and targeted incentives for digitalisation, sustainability and safety improvements. The three dimensions are consistent with EU regulatory and reporting instruments, extending the framework beyond operational benchmarking to corporate compliance. The Human dimension aligns with the EU Strategic Framework for Safety and Health at Work 2021–2027 and the European Sustainability Reporting Standards (ESRS) S1 "Own workforce" under the Corporate Sustainability Reporting Directive (CSRD, Directive, 2022/2464/EU): its indicators, covering training equity, ergonomic safety and workforce empowerment, are structurally consistent with the ESRS-S1 disclosure requirements. The digital dimension aligns with the EU AI Act requirements on transparency and human oversight in industrial AI systems; the XAI component directly addresses these explainability-by-design obligations. The green dimension aligns with the EU Taxonomy Regulation (Regulation, 2020/852), particularly its technical screening criteria for climate change mitigation and the circular economy, and with ESRS E1 (Climate Change) and E5 (Resource Use and Circular Economy) under the CSRD: indicators related to energy monitoring in maintenance operations and eco-efficiency in EEM are directly consistent with these reporting areas. The HDG-TPM5.0 framework can therefore support manufacturing companies in assessing operational readiness for CSRD social, digital and environmental disclosure requirements.

6. Conclusions, limitations and future research directions

Effective TPM remains central to equipment reliability, safety and resource efficiency, and must evolve towards human-centric, digital and sustainable practices under I5.0. This article introduced an HDG-oriented TPM5.0 maturity framework enabling heterogenous firms to assess maintenance maturity across Human, Digital, and Green dimensions and link results to TPM pillars.

The framework delivers: (1) a multidimensional benchmarking-oriented maturity index and (2) a contextual weighting architecture combining item-level leverage with sector–size and pillar-specific priorities, enabling peer-referenced assessment rather than absolute scoring. Benchmarking plays a dual role: it provides the maturity criterion while identifying best practices characterising higher performance across HDG dimensions and TPM pillars. Results are communicated through a 3D-HDG representation positioning each TPM pillar in the H–D–G space and comparing the firm profile with peer-clusters, reference making both the magnitude and direction of maturity gaps immediately interpretable for managers. XAI and NLP enhance transparency and support continuous learning and refinement.

Downloaded from http://www.emerald.com/bij/article-pdf/33/13/27/11604746/bij-11-2025-1075en.pdf by Universita Politecnica Delle Marche u

Key limitations expandable in future researches include: (1) the pilot nature of the empirical application, limited to a single sector–size cluster; results should therefore illustrative and analytically generalisable rather than statistically; (2) reliance on expert elicitation in questionnaire design and weighting, although grounded in literature, may introduce potential subjectivity; (3) the exploratory nature of the dynamic update mechanism and NLP/XAI monitoring layer.

In terms of data intensity, the questionnaire minimises the administrative burden through a structured Likert-scale format with open-ended comments, administered to TPM-related personnel. NLP automates signal extraction from qualitative responses, reducing analytical workload; data intensity depends on benchmark size rather than firm resources. Expert judgement follows standard MCDM practices when large datasets are unavailable (Alidrisi, 2020; Benmamoun *et al.*, 2025; Hajoary *et al.*, 2024; Vikas and Mishra, 2024), with bias mitigated through: (a) triangulation with literature and interviews, (b) dynamic updating progressively replacing expert-defined weights with data-driven estimates and (c) structured consistency review of the panel's calibration decisions.

Future research will expand the empirical base across sectors and firm sizes to build robust benchmark frontiers and enable formal validation (reliability, construct and criterion validity). Larger samples will also support longitudinal analyses linking HDG–TPM5.0 maturity to operational and sustainability outcomes and enable versioned updates. Planned validation includes: (1) at least three multi-cluster deployments for reliability testing, (2) confirmatory analysis of the HDG construct via structural equation modelling and (3) criterion validity testing against independent operational and sustainability indicators.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used generative AI to improve readability and language. After using this service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

Ahmed, M.S., Alhammadi, A.Y., Ali, A., Ahmed, S.M., Al Raeesi, H.J.H.A., Eid, A.D., Sumaida, N.S., Aslam, M., Al Bloushi, M.A., Al Mazrouei, M.K. and Al Awadhi, H.K. (2023), "Resource optimization at remote facilities by leveraging advance maintenance practices TPM (total productive maintenance), ODR (operator driven reliability) and industry 4.0 technologies", *ADIPEC*, SPE, doi: 10.2118/216229-MS.

Ahuja, I.P.S. and Khamba, J.S. (2008), "Total productive maintenance implementation in a manufacturing organisation", *International Journal of Productivity and Quality Management*, Vol. 3 No. 3, p. 360, doi: 10.1504/IJPQM.2008.017504.

Alidrisi, H. (2020), "Maintenance maturity level identification using MABAC method: an adaptation of TPM pillars in a public service sector", *Journal of Scientific and Industrial Research*, Vol. 79 No. 10, pp. 914-917, doi: 10.56042/jsir.v79i10.43545.

Ananth, G. and Vinayagam, B.K. (2015), "Effectiveness improvement through total productive maintenance using particle swarm optimisation model for small and micro manufacturing enterprises", *International Journal of Productivity and Quality Management*, Vol. 16 No. 4, p. 473, doi: 10.1504/IJPQM.2015.072425.

Antomarioni, S., Lucantoni, L., Ciarapica, F.E. and Bevilacqua, M. (2023), "A preliminary implementation of data-driven TPM: a real case study", in *Lecture Notes in Mechanical Engineering*, Springer, Cham, pp. 14-22, doi: 10.1007/978-3-031-25448-2_2.

Antomarioni, S., Fani, V., Bandinelli, R., Ciarapica, F.E. and Bevilacqua, M. (2025), "Toward quality 5.0: integrating industry 4.0, human-centricity, and quality management", *IFAC-Papers OnLine*, Vol. 59 No. 10, pp. 1414-1419, doi: 10.1016/j.ifacol.2025.09.238.

Downloaded from http://www.emerald.com/bij/article-pdf/33/13/27/11604746/bij-11-2025-1075en.pdf by Universita Politecnica Delle Marche u

Badiger, A.S., Gandhinathan, R. and Gaitonde, V.N. (2008), "A methodology to enhance equipment performance using the OEE measure", *European Journal of Industrial Engineering*, Vol. 2 No. 3, p. 356, doi: 10.1504/EJIE.2008.017690.

Bashar, A., Hasin, A.A. and Jahangir, N. (2020), "Linkage between TPM, people management and organizational performance", *Journal of Quality in Maintenance Engineering*, Vol. 28 No. 2, pp. 350-366, doi: 10.1108/JQME-11-2019-0105.

Benmamoun, Z., Fethallah, W. and Raouf, Y. (2025), "New green lean assessment framework for enhancing enterprise performance", *Benchmarking: An International Journal*, Vol. 33 No. 2, pp. 633-662, doi: 10.1108/BIJ-08-2023-0581.

Braun, V. and Clarke, V. (2006), "Using thematic analysis in psychology", *Qualitative Research in Psychology*, Vol. 3 No. 2, pp. 77-101, doi: 10.1191/1478088706qp063oa.

Breque, M., De Nul, L. and Petridis, A. (2021), "Industry 5.0: towards a sustainable, human-centric and resilient European industry", in *EU Research and Innovation Paper Series KI-BD-20-021-EN-N*, Directorate General for Research and Innovation (DG RTD) of the European Commission.

Cannavacciuolo, L., Ferraro, G., Ponsiglione, C., Primario, S. and Quinto, I. (2023), "Technological innovation-enabling industry 4.0 paradigm: a systematic literature review", *Technovation*, Vol. 124, 102733, doi: 10.1016/j.technovation.2023.102733.

Chen, P.-K., Fortuny-Santos, J., Lujan, I. and Ruiz-de-Arbulo-López, P. (2019), "Sustainable manufacturing: exploring antecedents and influence of total productive maintenance and lean manufacturing", *Advances in Mechanical Engineering*, Vol. 11 No. 11, doi: 10.1177/1687814019889736.

Correia, E., Garrido-Azevedo, S. and Carvalho, H. (2023), "Supply chain sustainability: a model to assess the maturity level", *Systems*, Vol. 11 No. 2, p. 98, doi: 10.3390/systems11020098.

Crosby, B. and Badurdeen, F. (2022), "Integrating lean and sustainable manufacturing principles for sustainable total productive maintenance (Sus-TPM)", *Smart and Sustainable Manufacturing Systems*, Vol. 6 No. 1, pp. 68-84, doi: 10.1520/SSMS20210025.

Denzin, N.K. (2012), "Triangulation 2.0", *Journal of Mixed Methods Research*, Vol. 6 No. 2, pp. 80-88, doi: 10.1177/1558689812437186.

Eisenhardt, K.M. (1989), "Building theories from case study research", *The Academy of Management Review*, Vol. 14 No. 4, pp. 532-550, doi: 10.2307/258557.

e Silva, M.H.P., Coser Mergulhão, R., Geraldo Vidal Vieira, J., Brasco Pampanelli, A., Salvador, R. and Aparecido Lopes Silva, D. (2024), "Lean-circular maturity model (LCMM) for companies' self-assessment in terms of process, product and life cycle thinking", *Waste Management*, Vol. 173, pp. 172-183, doi: 10.1016/j.wasman.2023.11.013.

European Commission (2021), *Industry 5.0 – Towards a Sustainable, Human-Centric and Resilient European Industry*, Publications Office of the European Union, Luxembourg. doi: 10.2777/308407.

Farajzadeh, L., Salem, A. and Salem, S. (2024), "Clean-up of wastewater discharged from zinc industry through application of zeolite 4A produced from a waste contaminated by lubricant oil components", *Journal of Hazardous Materials*, Vol. 480, 136204, doi: 10.1016/j.jhazmat.2024.136204.

Fereday, J. and Muir-Cochrane, E. (2006), "Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development", *International Journal of Qualitative Methods*, Vol. 5 No. 1, pp. 80-92, doi: 10.1177/160940690600500107.

Ferraro, S., Leoni, L., Cantini, A. and De Carlo, F. (2023), "Trends and recommendations for enhancing maturity models in supply chain management and logistics", *Applied Sciences*, Vol. 13 No. 17, 9724, doi: 10.3390/app13179724.

Ganesan, S., Lih, T.C. and Binti Mustafa, S.A. (2025), "Assessing lean transformation maturity: a real-case study on integrating technical and cultural practices for sustainable manufacturing in semiconductor memory assembly", *International Journal of Mechanical Engineering*, Vol. 12 No. 11, pp. 86-97, doi: 10.14445/23488360/IJME-V12I11P109.

Downloaded from http://www.emerald.com/bij/article-pdf/33/13/27/11604746/bij-11-2025-1075en.pdf by Universita Politecnica Delle Marche u

Gatell, I.S. and Avella, L. (2024), "A maturity model for assessing digital green lean leadership and culture implementation in manufacturing companies", *Total Quality Management and Business Excellence*, Vol. 35 Nos 7-8, pp. 860-897, doi: 10.1080/14783363.2024.2347373.

Hajoary, P.K., MA, A. and Garza-Reyes, J.A. (2024), "Industry 4.0 maturity assessment: a multi-dimensional indicator approach", *International Journal of Productivity and Performance Management*, Vol. 73 No. 4, pp. 981-1004, doi: 10.1108/IJPPM-07-2022-0325.

Hetmanczyk, M.P. (2024), "A method for evaluating the maturity level of production process automation in the context of digital transformation—polish case study", *Applied Sciences*, Vol. 14 No. 11, 4380, doi: 10.3390/app14114380.

Jick, T.D. (1979), "Mixing qualitative and quantitative methods: triangulation in action", *Administrative Science Quarterly*, Vol. 24 No. 4, p. 602, doi: 10.2307/2392366.

Jin, Q., Chen, H. and Hu, F. (2024), "Proposal of industry 5.0-Enabled sustainability of product–service systems and its quantitative multi-criteria decision-making method", *Processes*, Vol. 12 No. 3, p. 473, doi: 10.3390/pr12030473.

Kandler, M., Gabriel, P., Schröttle, V., May, M.C. and Lanza, G. (2022), "Modular, digital shopfloor management model – a maturity assessment for A human-oriented transformation process", *Proceedings of the Conference on Production Systems and Logistics*, pp. 642-651, doi: 10.15488/12153.

Kolla, S., Minufekr, M. and Plapper, P. (2019), "Deriving essential components of lean and industry 4.0 assessment model for manufacturing SMEs", *Procedia CIRP*, Vol. 81, pp. 753-758, doi: 10.1016/j.procir.2019.03.189.

Kumar, R.R., Ganesh, L.S. and Rajendran, C. (2021), "An entropy based approach to 5S maturity", *Materials Today: Proceedings*, Vol. 46, pp. 8103-8110, doi: 10.1016/j.matpr.2021.03.048.

Kwon, O. and Lee, H. (2004), "Calculation methodology for contributive managerial effect by OEE as a result of TPM activities", *Journal of Quality in Maintenance Engineering*, Vol. 10 No. 4, pp. 263-272, doi: 10.1108/13552510410564882.

Lucantoni, L., Antomarioni, S., Ciarapica, F.E. and Bevilacqua, M. (2022), "Implementation of industry 4.0 techniques in lean production technology: a literature review", *Management and Production Engineering Review*, Vol. 13 No. 3, pp. 83-93, doi: 10.24425/mper.2022.142385.

Lucantoni, L., Antomarioni, S., Ciarapica, F.E. and Bevilacqua, M. (2023), "A rule-based machine learning methodology for the proactive improvement of OEE: a real case study", *International Journal of Quality and Reliability Management*, Vol. 41 No. 5, pp. 1356-1376, doi: 10.1108/IJQRM-01-2023-0012.

Lucantoni, L., Antomarioni, S., Ciarapica, F.E. and Bevilacqua, M. (2025), "A data-driven framework for supporting the total productive maintenance strategy", *Expert Systems with Applications*, Vol. 268, 126283, doi: 10.1016/j.eswa.2024.126283.

Lucantoni, L., Ciarapica, F.E. and Bevilacqua, M. (2026), "From predictive maintenance 4.0 to 5.0: bringing humans back into the loop with a self-learning platform and implementation roadmap on automated production lines", *Journal of Quality in Maintenance Engineering*, Vol. 32 No. 5, pp. 58-75, doi: 10.1108/JQME-12-2025-0147.

Maddikunta, P.K.R., Pham, Q.-V., Prabadevi, B., Deepa, N., Dev, K., Gadekallu, T.R., Ruby, R. and Liyanage, M. (2022), "Industry 5.0: a survey on enabling technologies and potential applications", *Journal of Industrial Information Integration*, Vol. 26, 100257, doi: 10.1016/j.jii.2021.100257.

Martins, G.R.D.N., Ramos, L.F.P., Loures, E.F.R., dos Santos, V.A., do Amaral, L.M.B. and Deschamps, F. (2023), "Prioritization of industry 4.0 technologies to increase maturity in lean manufacturing from the perspective of enterprise interoperability", in *Lecture Notes in Mechanical Engineering*, Springer, Cham, pp. 379-386, doi: 10.1007/978-3-031-17629-6_39.

Mouhib, Z., Gallab, M., Merzouk, S., Soulhi, A. and Di Nardo, M. (2025), "Total productive maintenance and industry 4.0: a literature-based path toward a proposed standardized framework", *Applied System Innovation*, Vol. 8 No. 4, p. 98, doi: 10.3390/asi8040098.

Downloaded from http://www.emerald.com/bij/article-pdf/33/13/27/11604746/bij-11-2025-1075en.pdf by Universita Politecnica Delle Marche u

Närvä, M. and Bicheno, J. (2026), "Leanness – Comparison of different assessment methods", *Advances in Lean Manufacturing*, Vol. 2, pp. 136-145, doi: 10.1007/978-3-032-10111-2_11.

Nascimento, L.da.S., Trizotto, R.C.A., Sott, M.K., Lago, N.C., Pivoto, D., Costa Júnior, J.C.da and Cortimiglia, M.N. (2025), "Sustainable business model innovation: bibliometric review and maturity matrix for benchmarking social, environmental and economic dimensions", *Benchmarking: An International Journal*, Vol. 32, pp. 1-28, doi: 10.1108/BIJ-11-2024-0982.

Oeij, P., Lenaerts, K., Dhondt, S., Van Dijk, W., Schartinger, D., Sorko, S. and Warhurst, C. (2024), "A conceptual framework for workforce skills for industry 5.0: implications for research, policy and practice", *Journal of Innovation Management*, Vol. 12 No. 1, pp. 205-233, doi: 10.24840/2183-0606_012.001_0010.

Paiva, R.G.N., Melo, Y.R.A.V., Cavalcante, C., Tenório, V.A.S. and Do, P. (2024), "Developing data-driven O&M policy through sequential pattern mining: a case study", *Computers and Industrial Engineering*, Vol. 193, 110318, doi: 10.1016/j.cie.2024.110318.

Patton, M.Q. (2002), *Qualitative Research and Evaluation Methods*, 3rd ed., Sage Publications, Thousand Oaks, CA.

Pavan, R.O., Ferreira, M.A., Stefanelli, N.O. and Leal, G.C.L. (2023), "Maturity models in SSCM: a systematic review aimed at consolidating models and outlining possibilities for future research", *Benchmarking: An International Journal*, Vol. 30 No. 10, pp. 4076-4099, doi: 10.1108/BIJ-05-2022-0328.

Pulkkinen, A., Anttila, J.-P. and Leino, S.-P. (2019), "Assessing the maturity and benefits of digital extended enterprise", *Procedia Manufacturing*, Vol. 38, pp. 1417-1426, doi: 10.1016/j.promfg.2020.01.146.

Rendon, R.G. (2015), "Benchmarking contract management process maturity: a case study of the US Navy", *Benchmarking: An International Journal*, Vol. 22 No. 7, pp. 1481-1508, doi: 10.1108/BIJ-10-2014-0096.

Sajjad, A., Ahmad, W., Hussain, S., Chuddher, B.A., Sajid, M., Jahanjaib, M., Ali, M.K. and Jawad, M. (2024), "Assessment by lean modified manufacturing maturity model for industry 4.0: a case study of pakistan's manufacturing sector", *IEEE Transactions on Engineering Management*, Vol. 71, pp. 6420-6434, doi: 10.1109/TEM.2023.3259005.

Sanchez, F., Micaëlli, J.P., Bonjour, E. and Monticolo, D. (2019), "The transition between traditional project management to agile project management: using a project management maturity model", *Journal of Modern Project Management*, Vol. 7 No. 1, pp. 102-119, doi: 10.19255/JMPM01906.

Santos, V.A.D., Loures, E.D.F.R., Ramos, L.F.P. and Deschamps, F. (2020), "Lean manufacturing and industry 4.0: integrated roadmap proposal for process management maturity assessment", *IJCIEOM 2020 - International Joint Conference on Industrial Engineering and Operations Management*, doi: 10.14488/IJCIEOM2020_ABST_0001_37370.

Sassanelli, C., Rossi, M. and Terzi, S. (2020), "Evaluating the smart maturity of manufacturing companies along the product development process to set a PLM project roadmap", *International Journal of Product Lifecycle Management*, Vol. 12 No. 3, p. 185, doi: 10.1504/IJPLM.2020.10032146.

Seedao, S., Pannakkong, W., Huynh, V.-N., Sanglerdsinlapachai, N., Kojima, F., Fukuda, Y. and Tanaka, K. (2025), "Improvement steps towards digital transformation: enhancing productivity in developing manufacturing companies through total productive maintenance", *International Journal of Lean Six Sigma*, Vol. 16 No. 7, pp. 1674-1705, doi: 10.1108/IJLSS-07-2024-0149.

Shaaban, S.M. and Awni, H.A. (2014), "Critical success factors for total productive manufacturing (TPM) deployment at Egyptian FMCG companies", *Journal of Manufacturing Technology Management*, Vol. 25 No. 3, pp. 393-414, doi: 10.1108/JMTM-09-2012-0088.

Shi, X., Baba, T., Osagawa, D., Fujishima, M. and Ito, T. (2019), "A maturity model for sustainable system implementation in the era of smart manufacturing", *2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, IEEE, pp. 1649-1652, doi: 10.1109/ETFA.2019.8869446.

Downloaded from http://www.emerald.com/bij/article-pdf/33/13/27/11604746/bij-11-2025-1075en.pdf by Universita Politecnica Delle Marche u

Silva, N., Nakandala, D. and Yang, R.C. (2025), "A novel lean-centric readiness model for harnessing lean capabilities in industry 4.0 digital technology adoption", *Journal of Industrial Engineering and Management*, Vol. 18 No. 3, p. 459, doi: 10.3926/jiem.8651.

Shukla, S.K. and Sushil (2022), "Benchmarking the practices of flexibility with maturity models and frameworks of organizational capabilities", *Benchmarking: An International Journal*, Vol. 29 No. 2, pp. 664-682, doi: 10.1108/BIJ-08-2020-0459.

Silva, J., dos, S., Bouzon, M. and Chalmeta, R. (2025a), "Measuring digital sustainable maturity in supply chains: a maturity grid proposal", *International Journal of Productivity and Performance Management*, Vol. 75 No. 4, pp. 1-30, doi: 10.1108/IJPPM-01-2025-0012.

Silva, J., dos, S., Bouzon, M. and Chalmeta, R. (2025b), "Advancing sustainability in the digital age: defining and measuring maturity models for supply chains", *Benchmarking: An International Journal*, Vol. 33 No. 5, pp. 1376-1400, doi: 10.1108/BIJ-07-2024-0595.

Sivakumar, A., Saravanakumar, S., Sathiamurthi, P. and KarthiVinith, K.S. (2022), "Forecasting the equipment effectiveness in total productive maintenance using an intelligent hybrid conceptual model", *Transactions of FAMENA*, Vol. 46, pp. 29-40, doi: 10.21278/TOF.463042822.

Skalli, D., Charkaoui, A. and Cherrafi, A. (2023), "Design of industry 4.0 maturity model and readiness assessment under the digital transformation towards the implementation of LSS4.0", in *Artificial Intelligence and Industrial Applications (A2IA 2023)*, pp. 167-180, doi: 10.1007/978-3-031-43524-9_12.

Sovacool, B.K., Iskandarova, M. and Hall, J. (2023), "Industrializing theories: a thematic analysis of conceptual frameworks and typologies for industrial sociotechnical change in a low-carbon future", *Energy Research and Social Science*, Vol. 97, 102954, doi: 10.1016/j.erss.2023.102954.

Tadja, D.D., Lehyani, F., Bassetto, S.-J., Zouari, A., Tollenaere, M. and Wong, T. (2025), "Inclusive tool to assess lean manufacturing maturity and its relationship with the size or location of the company in greater Montreal, Canada", *Cogent Engineering*, Vol. 12 No. 1, 2535692, doi: 10.1080/23311916.2025.2535692.

Terelak-Tymczyna, A. and Niesterowicz, B. (2024), "Lean manufacturing in digital transformation of manufacture", in *Advances in Manufacturing IV. MANUFACTURING 2024. Lecture Notes in Mechanical Engineering*, Springer, Cham., pp. 74-83, doi: 10.1007/978-3-031-56444-4_6.

Torkayesh, A.E., Tirkolaee, E.B., Bahrini, A., Pamucar, D. and Khakbaz, A. (2023), "A systematic literature review of MABAC method and applications: an outlook for sustainability and circularity", *Informatica*, Vol. 34, pp. 415-448, doi: 10.15388/23-INFOR511.

Ulhe, P.P., Dhepe, A.D., Shevale, V.D., Warghane, Y.S., Jadhav, P.S. and Babhare, S.L. (2024), "Flexibility management and decision making in cyber-physical systems utilizing digital lean principles with Brain-inspired computing pattern recognition in Industry 4.0.", *International Journal of Computer Integrated Manufacturing*, Vol. 37 No. 6, pp. 708-725, doi: 10.1080/0951192X.2023.2257633.

Vikas and Mishra, A. (2024), "Evaluation of TPM adoption factors in manufacturing organizations using fuzzy PIPRECIA method", *Journal of Quality in Maintenance Engineering*, Vol. 30 No. 1, pp. 101-119, doi: 10.1108/JQME-11-2020-0115.

Vilarinho, S., Lopes, I. and Sousa, S. (2018), "Developing dashboards for SMEs to improve performance of productive equipment and processes", *Journal of Industrial Information Integration*, Vol. 12, pp. 13-22, doi: 10.1016/j.jii.2018.02.003.

Xu, X., Lu, Y., Vogel-Heuser, B. and Wang, L. (2021), "Industry 4.0 and industry 5.0—inception, conception and perception", *Journal of Manufacturing Systems*, Vol. 61, pp. 530-535, doi: 10.1016/j.jmsy.2021.10.006.

Yin, R.K. (2018), *Case Study Research and Applications: Design and Methods*, 6th ed., Sage Publications, Thousand Oaks, CA.

Zero, E., Sallak, M. and Sacile, R. (2024), "Predictive maintenance in IoT-Monitored systems for fault prevention", *Journal of Sensor and Actuator Networks*, Vol. 13 No. 5, p. 57, doi: 10.3390/jsan13050057.

Downloaded from http://www.emerald.com/bij/article-pdf/33/13/27/11604746/bij-11-2025-1075en.pdf by Universita Politecnica Delle Marche u

Zoubek, M., Poor, P., Broum, T., Basl, J. and Simon, M. (2021), "Industry 4.0 maturity model assessing environmental attributes of manufacturing company", *Applied Sciences*, Vol. 11 No. 11, 5151, doi: 10.3390/app11115151.

Žurga, G. (2018), "Project management in public administration. TPM – total project management maturity model. The case of Slovenian public administration", *Transylvanian Review of Administrative Sciences*, Vol. 14 No. 53 E, pp. 144-159, doi: 10.24193/tras.53E.9.

Corresponding author
Laura Lucantoni can be contacted at: l.lucantoni@staff.univpm.it

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: **permissions@emeraldinsight.com**

Downloaded from http://www.emerald.com/bij/article-pdf/33/13/27/11604746/bij-11-2025-1075en.pdf by Universita Politecnica Delle Marche u