

Project Title: A Decision-Support Framework for Sequencing Industry 4.0 Investments in Lean Surface-Mount (SMT) Electronics Production

Module code:	EG7302
Module name:	MEM Individual Project

Date of submission: 24-08-2026

Author(s):	Bindushree Sanjeev Murthy
Student ID(s):	259045211
Degree:	Master of Engineering Management
Tutor/Project supervisor:	Dr Imran Patel/ Dr Saravana Kumar K. Jaganathan

SUPERVISOR'S COPY/EXAMINER'S COPY [For 3rd/4th year projects only - delete as appropriate]

By submitting this report for assessment, I confirm that this assignment is my own work, is not copied from any other person's work (published or unpublished), and that it has not previously been submitted for assessment on any other module or course. By submitting this report for assessment, I confirm that this assignment is my own work, is not copied from any other person's work (published or unpublished), and that it has not previously been submitted for assessment on any other module or course.

I am aware of the University of Leicester's policy on plagiarism and have taken the online tutorial on avoiding plagiarism. I understand that plagiarism in this project report may result in severe penalties, up to and including expulsion from the University without a degree.

Table of Contents

Abstract.....	4
Chapter 1 Introduction	5
1.1 Background and industrial context	5
1.2 Problem statement: the I4.0 sequencing gap.....	8
1.3 Research question and sub-questions	8
1.4 Aim and objectives (O1-O5)	8
1.5 Scope and boundaries	8
1.6 Significance and contribution	9
1.7 Structure of the dissertation	9
Chapter 2 Literature Review	9
2.1 Review method and search protocol (PRISMA).....	9
2.2 Lean and Six Sigma in electronics manufacturing	10
2.3 Industry 4.0 technologies in SMT (IoT-SPC, MES, ML-AOI, digital twin).....	10
2.4 Lean-Industry 4.0 integration literature.....	11
2.5 Maturity and readiness models	12
2.5.1 Maturity-model development as a research object	12
2.5.2 Operationalisation and measurement theory	13
2.5.3 Readiness assessment and design-science evaluation.....	13
2.6 Investment appraisal and technology sequencing	13
2.7 Literature synthesis matrix and thematic findings	14
2.8 Research gap statement	15
Chapter 3 Methodology.....	15
3.1 Research design and philosophy (design-science, desk-based)	15
3.2 Overview of the four-phase approach	16
3.3 Systematic literature review protocol.....	16
3.4 Framework construction approach	16
3.4.1 Mapping from source instruments to framework dimensions	17
3.4.2 Measurement position and the aggregation rule	17
3.4.3 Instrument reliability protocol	18
3.5 Decision-support tool development method	18
3.6 Validation protocol (≥3 case studies)	20
3.6.1 Reliability - do different assessors obtain the same scores?	20
3.6.2 Repeatability - does the same assessor obtain the same result twice?.....	21
3.6.3 Reproducibility - can an independent researcher obtain the same result?	21
3.6.4 Robustness - do the conclusions survive plausible variation?	21
3.6.5 Predictive agreement - does the framework match documented outcomes?	22
3.6.6 Simulation - verification and validation	22
3.7 Assumptions and parameter ranges	23
3.8 Ethics and risk considerations	23
Chapter 4 Results: Framework, Tool and Validation.....	24
4.1 The Digital - Lean Synergy Model	24
4.1.1 Four-level Lean-maturity model	24
4.1.2 Industry 4.0 readiness assessment instrument.....	27
4.1.3 Technology-to-KPI mapping.....	29
4.1.4 Assessment procedure	29
4.2 Sequencing decision logic (rule set and decision tree)	30
4.3 The decision-support tool	31
4.3.1 Architecture (Python engine · Streamlit dashboard · MATLAB visual analytics)	31
4.3.2 ROI / NPV / payback model	32
4.3.3 Worked example and verification	33
4.4 Framework validation against published cases	34
4.4.1 Case selection and data extraction	34
4.4.2 Tool recommendation versus reported outcome.....	35
4.5 Sensitivity analysis	36
4.6 Cross-validation by MCDA, simulation and decision analysis	37
4.6.1 Multi-criteria decision analysis (MCDA)	38
4.6.2 Decision-analytic influence diagram.....	39
4.6.3 Discrete-event simulation of the SMT front-end line	40
4.6.4 Convergence and synthesis	41
4.7 Chapter summary	42
Chapter 5 Discussion.....	43
5.1 Interpretation of validation results	43
5.2 Comparison with existing literature	44
5.3 Managerial and practical implications	45
5.4 Contribution to theory and practice	45

5.5 Limitations and threats to validity	46
6.1 Answer to the research question	47
6.2 Achievement of objectives (O1-O5)	48
6.3 Key contributions.....	49
6.4 Future work	50
Chapter 7 Self-Learning Plan Execution	50
7.1 Skills audit	51
7.2 Learning objectives and resources.....	51
7.3 Evidence of applied learning (LM-01 to LM-05)	52
7.4 Reflection on effectiveness and next steps	52
References	54
Appendices.....	57
Appendix A Literature synthesis matrix	57
Appendix B Assessment instruments and user guide	57
Appendix C Case-extraction template	57
Appendix D Ethics position and project risk assessment.....	58
Appendix E Computational artefacts and reproduction record.....	58
Appendix F Using the decision-support application	59
Appendix G Abbreviations, symbols and notation.....	60

Note to the reader: *software used and how to access it*

The results in this report were produced with six tools. The decision-support engine and the cross-validation models are written in Python 3.9 (tool.py, mcda_ahp_topsis.py with pymcdm, and smt_line_des.py with SimPy). The independent ROI ledger is a Microsoft Excel workbook. The decision-analytic cross-check is a GeNIe Academic influence diagram. MATLAB scripts generate the published charts under the University's campus licence.

All the flow charts of this project were obtained from draw.io, and Graphics Images Obtained from Figma.com

The interaction layer is a Streamlit web application.

The framework can be exercised directly, without installing any of the above, at the deployed application: <https://dlsm-dashboard-cgwxs9tn32gzktrlky3khw.streamlit.app>. Appendix F provides a guide to every input and output.

All source code, the Excel ledger, the MATLAB scripts, the GeNIe model, the seven assessment forms (Forms B1–B7), the scorer's handbook and every software output reported here are deposited in the project repository at:

GitHub repository: <https://github.com/sbindushree2502-ux/EG7302-DLSM-Framework>

Abstract

iPhone MLB surface-mount assembly represents one of the most challenging high-volume manufacturing processes in consumer electronics, being concentrated in the EMS fleets of China and India. EMS firms are capital-intensive, but capital is allocated line by line: IoT-statistical process control (SPC), a manufacturing execution system (MES), ML-automated optical inspection (AOI), and digital twin technologies cannot be rolled out across several hundred lines at once. The literature establishes that Lean maturity conditions successful I4.0 adoption but offers no quantified approach to line-by-line prioritisation. This research addresses that gap.

Informed by the PRISMA framework, a systematic literature review of 19 sources (16 peer-reviewed) proves the sequencing issue. It identifies baseline SMT key performance indicators (first pass yield 92-98%, defects per million opportunities 50-500, overall equipment effectiveness (OEE) 60-85%). Thereafter, a desk-based design-science methodology develops the Digital–Lean Synergy Model, which includes: a four-level Lean maturity rubric which filters technology eligibility (Level 2 technologies include IoT-SPC and MES while ML-AOI and digital twin belong to Level 3); a readiness rubric (three dimensions, organisational, technical and financial scored 1-5) that determines the timing of deployment; mapping technology-to-KPI relationships together with the parametric payback/NPV model to rank eligible alternatives. The framework is implemented as a decision-support tool in Python and Excel and tested on five synthetic profiles, with a line calculated manually for a 100% match. Using base priors, IoT-SPC returns its costs within 26.7 months, with a five-year net present value of £70,585, and the recommended technology sequence is insensitive to $\pm 30\%$ cost and $\pm 20\%$ benefit variation in 1,000 simulation runs. Screening four published cases against pre-registered criteria admitted one line-resolved case; a blind run confirmed the maturity gate and KPI direction while revealing a refinement to the instrumentation rule. The recommended order was reproduced independently by entropy-weighted TOPSIS, VIKOR and PROMETHEE II, by an influence diagram under maturity uncertainty, and by a thirty-replication discrete-event simulation. Instrument reliability is specified but unexecuted and is carried as future work.

(IPC, 2022) This flow goes far beyond traditional SMT operations. Boards undergo pre-assembly functional testing (Pre AP in DFU, FCT and SOC areas); underfill is applied at component edges and drawn under the packages by capillary action to reinforce placements; MTS then joins CG and BG sub-boards via the interposer to form the final stacked assembly; and finally the product undergoes post-assembly testing (Post AP in DFU, LVCC-DC, FCT and WCS), RF testing at 13 stations, and graphite label marking with QR codes. In all these stages, solder-related issues such as defects, bridges, insufficient paste, cold joints and voids dominate SMT defect Pareto charts; this is why the inspection-intensive process steps (SPI, AOI and X-ray at test) carry disproportionate weight in terms of quality performance, and why inspection technology is central to this project. Figure 3 shows the same line as an equipment layout with a candidate sensor retrofit, which defines the physical scope of the IoT-enabled process control assessed in Chapters 2 and 4. Figure 4 plots India's share of that output and Tata Electronics' share within it.

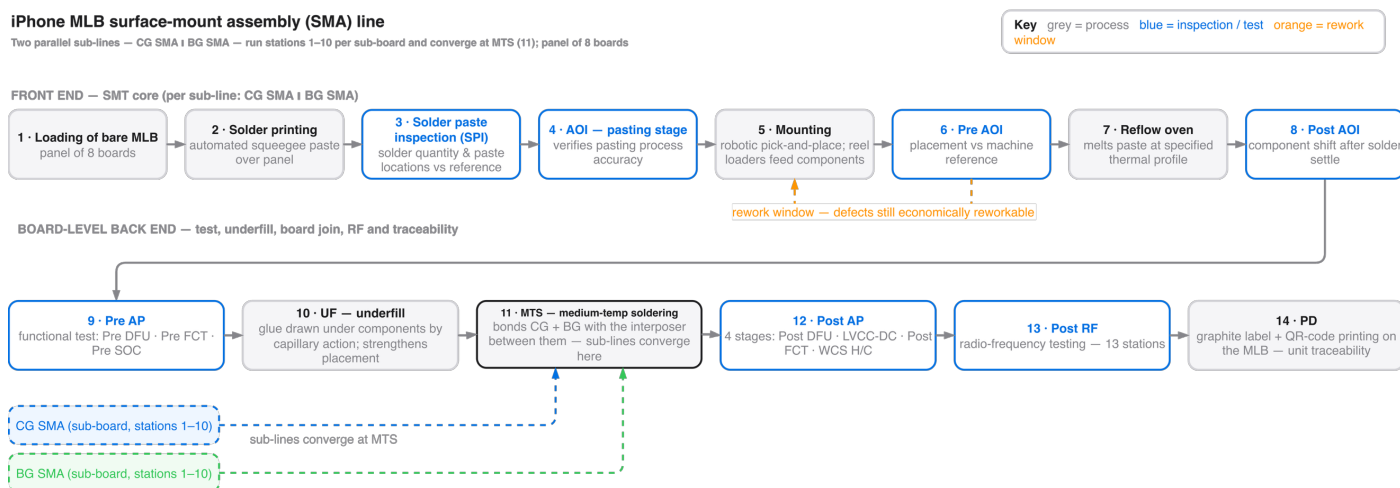


Figure 2 The iPhone MLB surface-mount assembly (SMA) line: fourteen stations spanning the front-end SMT core (1–8) and the board-level back-end (9–14), with the CG and BG sub-lines converging at MTS and the rework window at Pre AOI. Source: compiled from the Phase-1 research dossier (WBS 1.9).

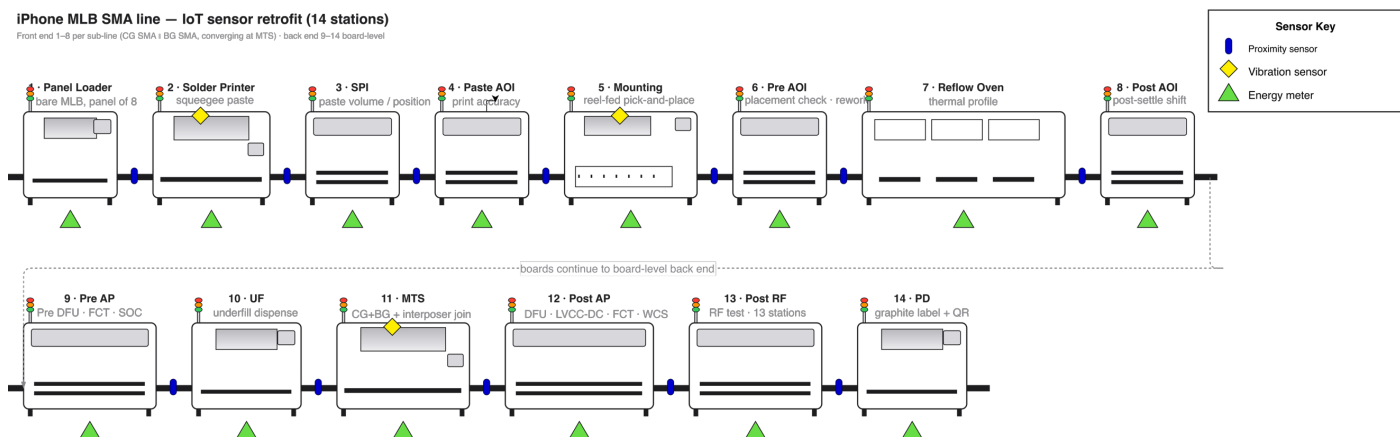


Figure 3: The same fourteen-station line viewed as an equipment layout with a candidate IoT sensor retrofit overlaid: proximity sensors at station boundaries for flow and cycle-time capture, vibration sensors on the printer, placement and joining stations, and energy meters per front-end station. The figure establishes the physical scope of the IoT-enabled statistical process control evaluated throughout this study and shows that the instrumentation is a retrofit to existing equipment rather than a line replacement, which is what places its capital requirement in the band reported in Table 1.

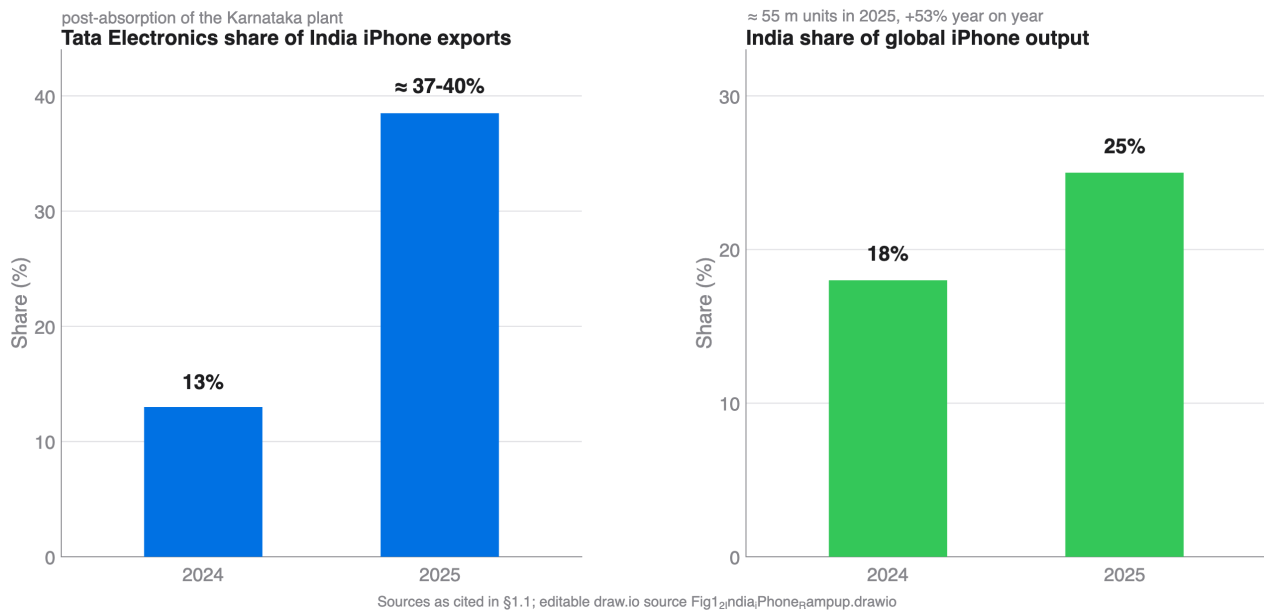


Figure 4 India's iPhone manufacturing ramp-up: Tata Electronics' share of India iPhone exports (left) and India's share of global iPhone output (right) (MeitY, 2025; Bloomberg, 2026; Free Press Journal, 2026). Values are from the cited in §1.1; the chart is a rendering of that source, not a new calculation.

(MeitY (Ministry of Electronics and Information Technology), 2025)(MeitY (Ministry of Electronics and Information Technology), 2025; Bloomberg, 2026; Tech, 2026)(MeitY (Ministry of Electronics and Information Technology), 2025; Bloomberg, 2026; Tech, 2026; Free, 2026)Geographically, this process has taken place in two countries. In China, mega-campus EMS companies Foxconn (Zhengzhou, Shenzhen, Chengdu), Pegatron, Luxshare, and BYD Electronics have assembled and performed final assembly for iPhones for more than a decade. In India, production capacity has been ramped up at an unprecedented pace under the Production-Linked Incentive (PLI) programme for electronics manufacturing on a large scale: by 2025, 25% of all iPhone output about 55 million units with year-to-year growth of 53%, was produced in India, smartphone exports reached an all-time high of more than ₹2 lakh crore in FY2024-25, with iPhones making up some 70% of exports (India TV News, 2025), and the share of Tata Electronics in the exports of iPhones from India increased from 13% in 2024 to 37-40% in 2025 after the company acquired the former Wistron plant in Karnataka and developed its capacity in Hosur. Now, India-based suppliers produce circuit boards and sub-assemblies for the Apple system (Figure 4). Strategically, this results in a manufacturing system on the scale of fleets: tens to hundreds of SMT lines per campus in two countries, renewed annually by new-product introductions.

(World Economic Forum, 2021) (World Economic Forum, 2023) In this way, two improvement paradigms contend for capital investment. In Lean Six Sigma (LSS), the process is stabilised and improved using standard work, value stream mapping, statistical process control (SPC) and DMAIC problem solving. Industry 4.0 enhances these with connected sensing, real-time execution control, machine-learning inspection, and simulation. Both of which have been proven effective at the smartphone-EMS scale: the Chengdu campus of Foxconn, named a WEF Global Lighthouse in 2021, employed mixed reality, AI and IoT to increase labour productivity by 200%, and OEE by 17%; the Shenzhen campus of Foxconn Industrial Internet, named in 2023, implemented 37 Fourth-Industrial-Revolution use cases at scale for agile smartphone new-product introduction. Table 1 lists the four Industry 4.0 technologies considered for this project, their indicative capital bands per line, and the Lean maturity level assumption for each.

Table 1 Candidate Industry 4.0 technologies for SMT lines: capital bands and Lean-maturity prerequisites (approved Project Proposal, Table 1).

Technology	Per-line capex band	Maturity gate	Primary KPI moved
IoT-enabled SPC	£80k - £120k	L2+	First-pass yield · defect PPM · Cpk
MES	£150k - £250k	L2+	Throughput · WIP · schedule adherence · OEE
ML-based AOI	£100k - £180k	L3+	Solder-defect PPM · false-call rate · escapes
Digital twin	£80k - £200k	L3+	OEE · predictive maintenance · energy

1.2 Problem statement: the I4.0 sequencing gap

(MeitY (Ministry of Electronics and Information Technology), 2025; Free, 2026) Proving feasibility through lighthouses doesn't provide replication strategy; even a financially well-endowed EMS company cannot digitise all lines at once because of the high capital intensity of digitisation shown in capital bands in Table 1, application of an advanced technology on a line with the process maturity level not sufficient enough to leverage it means wasting capital at fleet-level, a process without stability means a lack of consistent data for ML-AOI to learn from, and modelling a digital twin for a poorly understood line means learning from a noisy one. Moreover, heterogeneity compounds the challenge: mature Chinese brownfield lines are combined with recently built Indian greenfield lines that have less data infrastructure and younger staff, and the PLI incentive scheme links subsidies to incremental production. Each line must be prioritised by maturity, readiness, and the anticipated lift in KPIs. A systematic literature review in Chapter 2 shows that the literature consensus is that Lean maturity enables I4.0 success and presents integration concepts, but it does not provide a quantifiable, line-specific methodology for which technology to apply on which line and how to sequence it. This I4.0 knowledge gap is the challenge this dissertation addresses.

1.3 Research question and sub-questions

RQ. In what order should an Industry 4.0 implementation (IoT-powered SPC, MES, ML-driven AOI, Digital Twin) be introduced at a Chinese/Indian high-volume smartphone SMT contract manufacturing company, in terms of its lean capability and readiness, to get the most bang for the buck?

These sub-questions frame the research question:

SQ1 What is the minimum Lean maturity stage that each technology requires before it can pay off?

SQ2 What readiness factors (organisational, technical, financial) are the major bottlenecks to adoption of technologies in China/India smartphone-EMS?

SQ3 How sensitive is the suggested order to funding and the assumed improvement in KPIs of each technology?

1.4 Aim and objectives (O1-O5)

The objective of the study is to devise and validate a decision-support framework for sequencing investment in Industry 4.0 in lean SMT production lines to achieve maximum benefit per capital investment for KPIs. The five locked objectives (see Table 2) articulate the stated objective; each is specific and measurable.

Table 2 Locked objectives O1-O5 with verification evidence (Phase-1 dossier, WBS 1.11).

ID	Objective	Verification evidence
O1	Synthesise the literature on LSS–I4.0 integration in electronics manufacturing to define and evidence the sequencing gap, producing a matrix of at least fifteen peer-reviewed sources	PRISMA record; 16 peer-reviewed sources retained; synthesis matrix (§2.7, Appendix A)
O2	Develop a four-level Lean-maturity model and a three-dimensional I4.0-readiness rubric, with measurable indicators for each	Framework v1.1 signed off at supervisor review (WBS 2.8)
O3	Map the candidate technologies to SMT KPIs and build a parametric payback and net-present-value model quantifying expected return	Technology-to-KPI matrix; parametric ROI workbook
O4	Integrate maturity, readiness, KPI mapping and return logic into a single decision-support tool implemented in Python and Excel	tool.py + Excel dashboard verified against five hand-calculated synthetic line profiles
O5	Validate the framework against at least three published SMT case studies and a sensitivity analysis, and derive managerial recommendations	Validation results report; sensitivity tornado; limitations analysis

1.5 Scope and boundaries

This project is a desk-based design-science study comprising a systematic review, framework construction, tool development, and secondary validation using published cases. The analysis focuses on individual SMT lines within high-volume smartphone-EMS fleets in China or India. The technology set is limited to the four options listed in Table 1. The KPI set includes first-pass yield, defect PPM/DPMO, OEE, changeover time, and process capability (Cpk). No factory pilots are conducted, and primary data are restricted to optional, ethics-screened expert interviews, subject to approval. The findings apply analytically, not statistically, to other high-volume discrete assembly environments. All capital bands, effect sizes, and payback figures are

treated as editable design parameters, not guarantees, and are validated or adjusted against published case data during validation.

1.6 Significance and contribution

The contribution is threefold. Practically, the framework converts a recurring, high-stakes managerial judgement about where to allocate the next unit of I4.0 capital across an SMT fleet into a transparent, auditable procedure with an explicit maturity gate, a named binding readiness constraint and a quantified expected return. Theoretically, it answers calls for prescriptive adoption guidance in the Lean-I4.0 review literature by providing the ordering logic that integration models leave unquantified. Societally, sequencing capital towards the lines that can exploit it reduces wasted investment, energy and electronic scrap in the world's two largest smartphone-manufacturing economies, supporting sustainable industrialisation objectives; a full risk and sustainability appraisal accompanies the methodology chapter.

1.7 Structure of the dissertation

Chapter 2 reports the systematic literature review and the evidence gap. Chapter 3 sets out the design-science methodology, review protocol, framework-construction approach, and validation protocol. Chapter 4 presents the results: the Digital-Lean Synergy Model, the sequencing decision logic, the decision-support tool, and its validation against published cases, including a sensitivity analysis. Chapter 5 discusses the findings in relation to the literature and the managerial implications; Chapter 6 concludes and outlines future work; Chapter 7 presents evidence of the self-learning plan's execution. The appendices include the synthesis matrix, the tool user guide, the case-extraction templates, and the ethics record.

Chapter 2 Literature Review

2.1 Review method and search protocol (PRISMA)

(Page *et al.*, 2021) The review follows a PRISMA-informed systematic procedure, adapted to a thirteen-week MSc project, with the protocol fixed before screening to ensure the eventual gap claim is replicable and defensible. Three review questions bound the search: RQ-A, the documented relationship between Lean/Six-Sigma maturity and successful I4.0 adoption in discrete manufacturing; RQ-B, which I4.0 technologies map to which SMT KPIs and with what effect-size evidence; and RQ-C, what guidance exists for sequencing I4.0 investments line-by-line across large SMT fleets in the China/India EMS context. Inclusion criteria included English-language publications from 2015-2026 (with seminal earlier works identified via backward snowballing), peer-reviewed journals and conferences, and industry-grade grey literature (IPC, the WEF Global Lighthouse Network, government and analyst reporting); vendor marketing and opinion pieces without methods were excluded. A three-block Boolean string (lean, six sigma, DMAIC) AND digital technology (Industry 4.0, IoT, MES, digital twin, machine learning, AOI) AND application domain (SMT, electronics assembly, PCB, smartphone, EMS) was defined in Scopus syntax, with Web of Science and IEEE Xplore equivalents. Screening proceeded in two passes, with reasons for exclusion logged: 26 records were assessed at full text, seven were excluded (four wrong domain, two lacking empirical content, one superseded), and 19 sources were retained, 16 peer-reviewed and three industry-grade, exceeding the fifteen-source objective O1 (Figure 2). Database identification counts are recorded from the authenticated institutional re-run of the documented query strings, as noted in the methodology chapter.

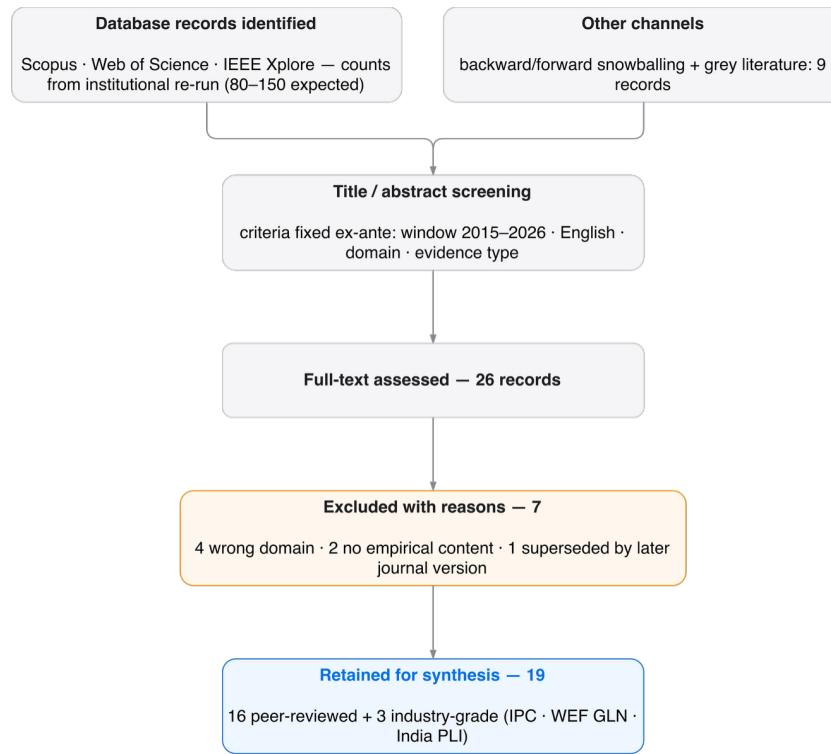


Figure 5 PRISMA-style flow of the systematic review: 26 records assessed at full text, 7 excluded with reasons, 19 retained for synthesis (16 peer-reviewed) (Page *et al.*, 2021).

2.2 Lean and Six Sigma in electronics manufacturing

(Guo *et al.*, 2019) (IPC, 2022) (IPC, 2022) Lean practice entered electronics assembly through the same routines it standardised elsewhere: standard work, 5S, value-stream mapping (VSM), SMED changeover reduction, and pull flow. Six Sigma contributed the DMAIC cycle and the process-capability discipline that high-density soldering demands. Their integration is well documented at the line level: VSM operates within a DMAIC cycle to drive concurrent Lean-Kaizen improvement on a high-volume assembly line, using the same indicator logic later applied to operationalise maturity levels in this project. In SMT specifically, the IPC benchmark programme quantifies the quality baseline through a biennial survey of more than fifty EMS and OEM assemblers, tracking process yields, DPMO at test, and non-conformance costs. Table 3 consolidates the KPI definitions and published baseline ranges that anchor this project's framework: they define what a 'good' line looks like before digitisation and provide the baseline vector against which technology benefits are later quantified. DPMO is preferred over percentage yield for cross-product comparability, because per-million-opportunity counts do not vary with board complexity or batch size.

Table 3 SMT KPI definitions and published baseline ranges (compiled from the IPC benchmark programme (IPC, 2022) and the Phase-1 dossier, WBS 1.9).

KPI	Definition	Typical range (mixed EMS)	Best-in-class
First-pass yield (FPY)	Boards passing all gates first time ÷ boards started	92–98% (complexity-dependent)	> 98%
Defect rate (DPMO/PPM)	Defects per million opportunities (component, termination, placement)	50–500 DPMO	< 10 PPM
OEE	Availability × performance × quality of the SMT line	60–85%	≥ 85%
Changeover time	Last good board (product A) to first good board (product B)	8–15 min (high-mix, SMED-dependent)	< 8 min
Process capability (Cpk)	Capability of critical paste-deposit and placement parameters	1.0–1.33	≥ 1.67

2.3 Industry 4.0 technologies in SMT (IoT-SPC, MES, ML-AOI, digital twin)

(Xu, Xu and Li, 2018)(Zhang *et al.*, 2024)(Schmidt *et al.*, 2020)(Karanjkar *et al.*, 2018)(World Economic Forum, 2021)(World Economic Forum, 2023) The state-of-the-art review defines the I4.0 technology stack: cyber-physical systems, IoT, cloud and analytics, from which the four SMT-relevant candidates in this project are drawn. IoT-enabled SPC extends classical control charts with real-time sensing of solder-paste volume,

placement offset and reflow profile, detecting drift before it becomes defects and tightening capability towards the Cpk gates in Table 3. MES digitises work orders, dispatch, traceability and downtime capture, acting on flow KPIs WIP, schedule adherence and the availability component of OEE. ML-based AOI inserts a trained classifier (typically convolutional) between the optical-inspection machine and manual verification: classify false calls from solder-paste-inspection and pre-reflow AOI data on a production SMT line, reaching 89.3% accuracy across 30,000 production data points; demonstrate CNN-enhanced X-ray inspection of solder joints; and industry implementations report false-call reductions of the order of 30-40% with detection rates of 98-99% on critical joints, figures this project treats as adjustable design priors rather than guarantees. For digital twins, provide the canonical SMT evidence: an IoT-driven discrete-event twin of a working SMT-PCB line whose buffer-insertion experiments yielded a 2.7× reduction in energy consumption. At fleet scale, the WEF Lighthouse designations of Foxconn Chengdu (+200% labour efficiency, +17% OEE) and Foxconn Industrial Internet Shenzhen (37 use cases for agile smartphone NPI), quantify what integrated deployment achieves, but only at single flagship sites, which is precisely what a sequencing method must generalise from. Figure 6 shows the sensor deployment such instrumentation requires on a working line, and Figure 6 shows the convolutional architecture on which machine-learning inspection depends.

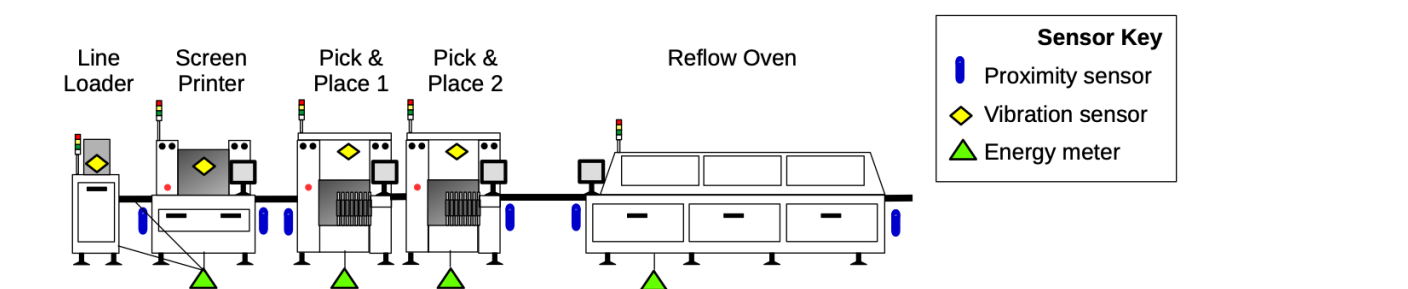


Figure 6 Sensor deployment on an instrumented SMT line for IoT-driven digital-twin energy optimisation (adapted from Karanjkar et al., 2018).

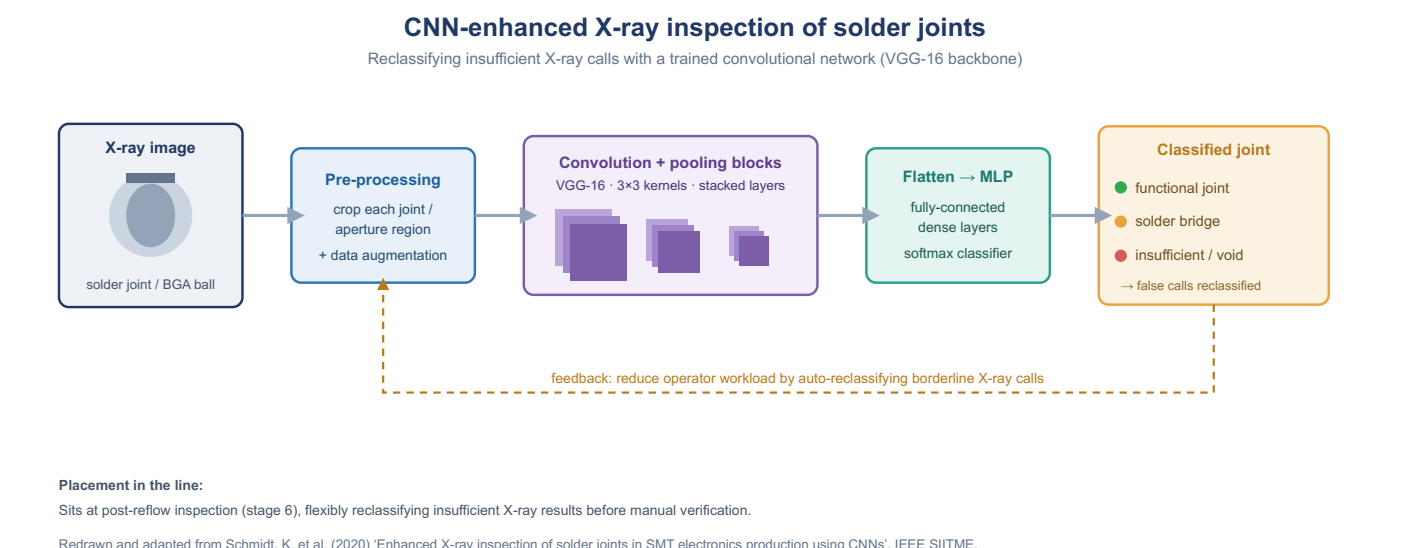


Figure 7 ML-based AOI: convolutional neural network architecture for false-call reduction in post-reflow X-ray inspection (adapted from Schmidt et al., 2020).

2.4 Lean-Industry 4.0 integration literature

The empirical stream consistently shows that Lean and I4.0 reinforce each other. (Tortorella, Giglio and van Dun, 2019) survey of 110 Brazilian manufacturers associates Lean practice with I4.0 adoption and finds that concurrent implementation yields larger performance gains; (Rossini et al., 2019) replicate this pattern across 108 European manufacturers, finding higher I4.0 adoption where Lean is extensively implemented; and (Tortorella, Giglio and van Dun, 2019) show that I4.0 adoption moderates the Lean-to-performance relationship. In India, directly relevant to this project's context, (Kamble, Gunasekaran and Dhone, 2020) find that Lean manufacturing mediates the path from I4.0 to sustainable organisational performance. The

conceptual stream supplies integration models: (Sony, 2018) proposes organisational integration propositions; the 'Lean 4.0' concept of (Sony, 2018) maps digital technologies onto classical Lean tools such as e-Kanban and digital VSM; (Ciano *et al.*, 2021) derive one-to-one relationships between specific I4.0 technologies and Lean techniques from multiple cases; and (Sanders, Elangeswaran and Wulfsberg, 2016) argue the reverse direction, I4.0 as an enabler of Lean, establishing that the causal ordering is contested and must be justified, not assumed. The review stream frames the agenda: the mapping by (Buer, Strandhagen and Chan, 2018) identifies four research streams and thin integration evidence; the systematic review by (Buer, Strandhagen and Chan, 2018; Pagliosa, Tortorella and Ferreira, 2021) explicitly calls for prescriptive adoption guidance; and the Lean-Six-Sigma-specific reviews, (Pongboonchai-Empl *et al.*, 2023) on DMAIC 4.0 and (Pongboonchai-Empl *et al.*, 2023; Skalli *et al.*, 2023) on the integrated LSS4.0 framework, position I4.0 as a driver or prerequisite of LSS while leaving capital constraints and sector-specific KPIs unaddressed. Across all three streams, the synergy is settled; the order of adoption is not.

2.5 Maturity and readiness models

(Schumacher, Erol and Sihm, 2016) (Tonelli *et al.*, 2022) Maturity and readiness models provide the assessment machinery with a sequencing method. The nine-dimension model is the reference instrument for assessing I4.0 readiness and maturity of manufacturing enterprises; the 3D-CUBE, reproduced as Figure 9, condenses readiness into technological, organisational and process enablers, evidence that a three-dimension structure retains discriminating power while remaining tractable for practitioners.

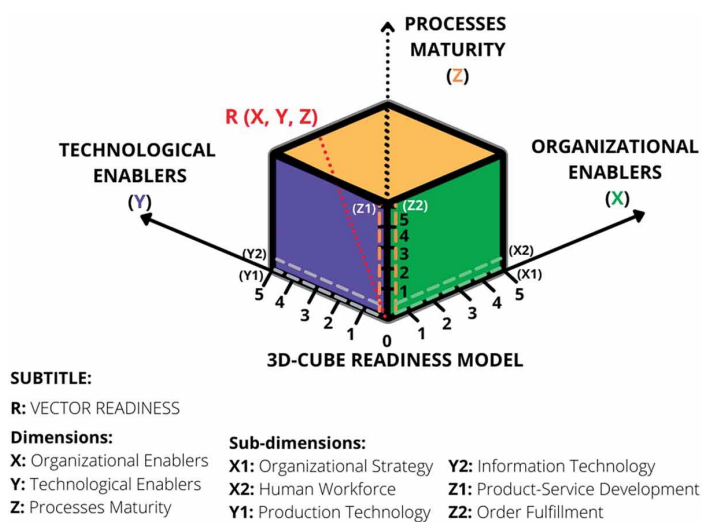


Figure 8: The 3D-CUBE readiness model of Tonelli *et al.* (2022): readiness expressed as a vector $R(X, Y, Z)$ over organisational enablers (X), technological enablers (Y) and process maturity (Z), each scored 1–5 with two sub-dimensions. The three-axis structure is the published precedent for the organisational, technical, and financial decomposition adopted in §4.1.2; Table 6 documents the condensation of this instrument and of Schumacher, Erol and Sihm (2016).

On the Lean side, long-established capability models describe progressions from reactive firefighting through formalised and deployed practice to continuous improvement as a 'way of life'. Two limitations matter for this project. First, these instruments are diagnostic: they locate a firm on a scale but do not prescribe which technology to buy next, for which line, at what expected return. Second, they assess the enterprise level, whereas capital in an SMT fleet is committed line by line. The framework developed in Chapter 4 therefore compresses the Lean ladders into a four-level line-level rubric used as an eligibility gate, and condenses the Schumacher and 3D-CUBE structures into a three-dimensional readiness rubric (organisational, technical, financial) scored 1-5, retaining the validated dimensional logic of (Schumacher, Erol and Sihm, 2016) and (Tonelli *et al.*, 2022) while converting diagnosis into decision.

2.5.1 Maturity-model development as a research object

Maturity models are designed artefacts with an established development methodology, not merely descriptive scales. The credibility of any new model rests on whether it followed that methodology. The capability-maturity tradition, originating with (Paulk *et al.*, 1993), establishes the defining structure: ordered levels, each a precondition for the next, with level membership determined by evidenced practice rather than by self-assessment. (de Bruin *et al.*, 2005) Decompose model development into scope, design, populate, test, deploy and maintain, and (Becker, Knackstedt and Pöppelbuß, 2009) add the requirement that every design decision be documented and justified against a comparison of existing models. (Mettler, 2011) distinguishes descriptive models, which report where an organisation stands, from prescriptive models, which state what should be done next; the artefact developed in Chapter 4 is deliberately prescriptive, which raises the evidential bar because a prescription that cannot be reproduced by a second assessor is not a decision aid

but an opinion. These four sources supply the procedure followed in §3.4 and the rationale for stating the maturity levels in §4.1.1 as mandatory quantitative criteria rather than descriptions.

2.5.2 Operationalisation and measurement theory

A construct becomes usable in research only when it is operationally defined and the operations used to measure it are specified closely enough that two investigators obtain the same value. (Churchill, 1979) sets out the sequence from construct definition through indicator generation to reliability assessment, and this remains the standard procedure for instrument construction. (Stevens, 1946) provides the measurement-theoretic constraint, whose classification of nominal, ordinal, interval and ratio scales determines which arithmetic operations are admissible: ordinal ratings support ordering, minima and medians, but do not licence unrestricted averaging. The applied literature is genuinely divided on how strictly to enforce this. (Jamieson, 2004) argues that computing means of Likert-type responses is a category error, while (Jamieson, 2004; Norman, 2010) shows empirically that parametric treatment of ordinal data is robust in practice. §3.4.2 takes a position on that debate rather than ignoring it. Reliability, finally, is not a property of an instrument in the abstract but of an instrument in use (Nunnally and Bernstein, 1994), which is why §3.4.3 specifies a reliability protocol rather than asserting that the rubrics are reliable.

2.5.3 Readiness assessment and design-science evaluation

Readiness instruments in the industry 4.0 literature are typically self-assessed and dimension-scored, which places the burden of objectivity on the wording of the anchors; when anchors describe attitudes rather than artefacts, the instrument measures optimism. The readiness rubric of §4.1.2 therefore requires documentary evidence for every score above the lowest anchor. The evaluation of the resulting artefact is situated within design-science research: (March and Smith, 1995) separate building an artefact from evaluating it; (Gregor and Hevner, 2013) require that a design contribution state its knowledge claim explicitly; and Venable, (Venable, Pries-Heje and Baskerville, 2016) provide the framework that distinguishes artificial from naturalistic evaluation and is applied in §3.6 and §5.1. (Straub, Boudreau and Gefen, 2004) provide the validation categories of content, construct, reliability and predictive validity against which §3.6 is organised, so the validation protocol addresses recognised categories rather than an ad hoc list.

2.6 Investment appraisal and technology sequencing

The strongest direct evidence that sequencing matters comes from (Frank, Dalenogare and Ayala, 2019), whose survey of adoption patterns shows that I4.0 technologies are implemented in layers, with base connectivity and monitoring technologies preceding advanced analytics and simulation, rather than as a big bang. This layered pattern is an observed regularity, not a decision rule: it describes what adopters did, not what a specific line should do next. Standard investment-appraisal theory provides the quantification machinery, including payback period, net present value and internal rate of return under a corporate discount rate (Brealey, Myers and Allen, 2020), but appraises projects one at a time; it does not encode technological prerequisites (a digital twin consumes the data MES and IoT-SPC create) or process prerequisites (ML-AOI requires a capable, stable process to learn from). Between the descriptive adoption literature, the diagnostic maturity instruments and project-level appraisal theory, no published method combines eligibility gating, absorption readiness and per-line return ranking into a single sequencing procedure for SMT fleets. This project contributes that combination.

Beyond single-project appraisal, three established method families provide the machinery for this framework. Multi-criteria decision analysis ranks alternatives against several conflicting criteria at once; the analytic hierarchy process (Saaty, 2008), TOPSIS (Hwang and Yoon, 1981), VIKOR (Ogrizovic and Tzeng, 2004) and PROMETHEE (Brans and Vincke, 1985) are the dominant formulations and so address the multi-KPI prioritisation that payback alone cannot. Decision analysis under uncertainty, and specifically the influence-diagram representation (Brans and Vincke, 1985; Howard and Matheson, 2005), models a capital decision when a prerequisite such as process maturity is only partially observed, and quantifies the value of confirming it before committing. Discrete-event simulation, the canonical tool for estimating line-level throughput and yield (Banks *et al.*, 2010), generates performance evidence where published deltas are sparse. None of these, in isolation, yields a maturity-gated sequencing procedure; the present framework combines their logic and, in Chapter 4, uses all three as independent cross-checks on its recommendation.

When simulation is used as corroborating evidence rather than as an illustration, the literature is explicit that the model itself must be shown to be credible before its outputs may be interpreted. (Sargent, 2013) distinguishes verification, which establishes that the model implements its specification correctly, from validation, which establishes that the specification adequately represents the system, and recommends degenerate-case testing as a minimal verification instrument. (Law, 2015) adds the terminating-versus-steady-state distinction that governs warm-up exclusion and the replication logic by which stochastic outputs are reported as interval rather than point estimates. These conventions are adopted in the validation protocol of §3.6.6 and applied to the discrete-event results of §4.6.3, so that the simulation evidence is reported with the same auditability as the financial results.

2.7 Literature synthesis matrix and thematic findings

(Tortorella and Fetterman, 2018; Rossini et al., 2019; Tortorella, Giglio and van Dun, 2019; Kamble, Gunasekaran and Dhane, 2020)(Buer, Strandhagen and Chan, 2018; Sanders, Elangeswaran and Wulfsberg, 2016; Mayr et al., 2018; Sony, 2018; Ciano et al., 2021; Pagliosa, Tortorella and Ferreira, 2021; Pongboonchai-Empl et al., 2023; Skalli et al., 2023)(MeitY (Ministry of Electronics and Information Technology), 2025; Bloomberg, 2026; Free, 2026)(Karanjkar et al., 2018; Schmidt et al., 2020; Zhang et al., 2024)(World Economic Forum, 2023; World Economic Forum, 2021)Table 4 condenses the synthesis matrix developed in the Phase-1 dossier (WBS 1.8; the full version, including method and sector detail per source, is reproduced in Appendix A). Four thematic findings emerge. T1 Lean-first evidence: survey studies across Brazil, Europe and India converge on Lean maturity as a precondition or amplifier of I4.0 benefits. T2 Integration without ordering: conceptual and review work provide integration models and technology-practice mappings but leave the investment order unquantified. T3 SMT evidence is isolated: sector-specific studies quantify single technologies on single lines or aggregate campus-level transformations, neither of which yields a per-line sequencing rule. T4 the China/India context raises the stakes: fleet heterogeneity and PLI-driven greenfield expansion multiply low-maturity lines awaiting prioritisation, asking sequencing a live managerial question rather than an academic one.

Table 4 Condensed literature synthesis matrix: 19 retained sources by theme, key finding and gap relevance (full matrix with methods and sectors in Appendix A).

Source	Theme	Key finding	Relevance to the sequencing gap
Tortorella and Fettermann (2018)	Lean–I4.0 association	Lean practices positively associated with I4.0; concurrent implementation yields larger gains	Establishes lean-first logic; no investment ordering
Buer et al. (2018)	Lean–I4.0 mapping (SLR)	Four research streams; integration evidence thin	Names the integration question; sequencing unaddressed
Xu, Xu and Li (2018)	I4.0 state of the art	Defines the I4.0 technology stack and adoption trajectory	Technology taxonomy for the framework profiles
Tortorella et al. (2019)	Moderation effect	I4.0 adoption moderates the Lean-to-performance relationship	Supports the maturity-gating assumption
Rossini et al. (2019)	Lean–I4.0 in Europe	I4.0 adoption easier where Lean is extensively implemented	Direct empirical support: Lean maturity as precondition
Mayr et al. (2018)	Lean 4.0 concept	Maps digital technologies onto classical Lean tools	Tool-level mapping informs the tech-to-KPI matrix
Kamble et al. (2020)	Performance path (India)	Lean mediates the I4.0-to-sustainable-performance path	Reinforces the sequencing premise in the Indian context
Pagliosa et al. (2021)	Lean–I4.0 SLR	Calls for prescriptive adoption guidance	Explicitly flags the gap this project fills
Ciano et al. (2021)	Technology–practice pairs	One-to-one links between I4.0 technologies and Lean techniques	Granular pairings feed the decision-tree rule set
Sanders et al. (2016)	I4.0 enables Lean	I4.0 research activities function as enablers of Lean	Counter-direction view; ordering must be justified both ways
Pongboonchai-Empl et al. (2023)	DMAIC 4.0 (SLR)	Analyse phase best supported by I4.0 (ML, IoT, process mining)	Anchors the LSS-side integration of the framework
Skalli et al. (2023)	LSS4.0 framework (SLR)	I4.0 mostly seen as driver/prerequisite of LSS	Closest antecedent framework; lacks capital constraints and SMT KPIs

Source	Theme	Key finding	Relevance to the sequencing gap
Schumacher, Erol and Sihn (2016)	I4.0 readiness model	Nine-dimension readiness/maturity assessment instrument	Methodological basis condensed into the 3-dimension rubric
Sony (2018)	Integration model	Organisational integration propositions for Lean and I4.0	Integration antecedent; sequencing left unquantified
Frank, Dalenogare and Ayala (2019)	Adoption patterns	I4.0 adopted in layers: base technologies precede advanced ones	Strongest empirical support for staged sequencing
Karanjkar et al. (2018)	Digital twin in SMT	IoT-driven twin of an SMT-PCB line; 2.7× energy reduction	Validation-grade SMT case for the case pool
IPC (2022)	SMT quality baselines	DPMO, yields and non-conformance costs benchmarked across 50+ assemblers	Quantitative KPI anchors for the baseline vector
World Economic Forum (2021; 2023)	Lighthouse effect sizes	Foxconn Chengdu +200% labour efficiency, +17% OEE; FII Shenzhen 37 use cases	Proves feasibility at flagship sites; cannot be replicated fleet-wide at once
MeitY (2025); Bloomberg (2026); Free Press Journal (2026)	India ramp-up (PLI)	≈25% of global iPhone output in India by 2025; Tata share 13% → ≈37–40%	Greenfield expansion multiplies low-maturity lines awaiting prioritisation

2.8 Research gap statement

The literature consistently finds that Lean/LSS maturity conditions successful Industry 4.0 adoption (Tortorella and Fettermann, 2018; Rossini *et al.*, 2019; Tortorella, Giglio and van Dun, 2019; Kamble, Gunasekaran and Dhone, 2020), offers integration concepts and mappings (Mayr *et al.*, 2018; Sony, 2018; Ciano *et al.*, 2021; Pongboonchai-Empl *et al.*, 2023; Skalli *et al.*, 2023), and shows that adoption proceeds in layered patterns (Frank, Dalenogare and Ayala, 2019). Yet no study provides a quantified, per-line sequencing method that tells a high-volume smartphone SMT contract manufacturer in China or India which technology to deploy on which line next, in what order, and with what expected KPI and payback effect. Even capital-rich EMS firms must ration fleet-wide capital line-by-line: lighthouse-grade adoption is demonstrated at flagship sites (World Economic Forum, 2021; World Economic Forum, 2023) but cannot be replicated simultaneously across hundreds of lines, and rapid PLI-driven greenfield expansion in India multiplies low-maturity lines awaiting prioritisation (Free, 2026; MeitY (Ministry of Electronics and Information Technology), 2025). SMT-specific evidence remains isolated technology cases (Schmidt *et al.*, 2020; Karanjkar *et al.*, 2018; Zhang *et al.*, 2024) rather than a sequencing framework. The following chapters close this gap by combining a maturity gate, a readiness modulator, and a per-line return ranking into one auditable decision procedure.

Chapter 3 Methodology

3.1 Research design and philosophy (design-science, desk-based)

The project adopts a design-science research (DSR) strategy: knowledge is generated by constructing and evaluating a purposeful artefact, here the Digital–Lean Synergy Model and its decision-support tool, against a documented organisational problem (Hevner *et al.*, 2004). The six-activity DSR process of (Peppers *et al.*, 2007) maps directly onto the work: problem identification and motivation (Chapters 1–2), definition of solution objectives (O1–O5, §1.4), design and development (the framework and tool, §4.1–4.3), demonstration (worked line profiles, §4.3.3), evaluation (case-study validation and sensitivity analysis, §4.4–4.5), and communication (this dissertation and the colloquium presentation). The stance is pragmatist: constructs are judged by their usefulness to the deciding manager rather than by correspondence to a single observer-independent truth, which suits a study whose output is a working decision procedure. The study is entirely desk-based: all inputs are secondary peer-reviewed literature, industry-grade benchmark and policy reporting, and published case studies, and the unit of analysis is the individual SMT line within a high-volume smartphone-EMS fleet in China or India. Generalisation is analytic, to comparable high-volume discrete assembly settings, not statistical (Yin, 2018). Evaluation is planned against the FEDS framework of Venable, (Venable, Pries-Heje and Baskerville, 2016), which separates artificial evaluation, establishing that an artefact behaves as specified, from naturalistic evaluation, establishing that it performs in its intended setting; the two classes are reported separately in §5.1 rather than aggregated into a single validation claim.

3.2 Overview of the four-phase approach

The work was executed in four phases (Figure 9), each closing with a named deliverable and mapped to the locked objectives; the full 46-item work breakdown was tracked with supervisor gates at framework sign-off (WBS 2.8) and at full-draft review (WBS 4.11). Table 5 summarises the phase structure. This staging enforces methodological discipline: the framework could not be encoded until its design parameters were fixed and reviewed, and the tool could not be validated until its logic was verified against hand calculation.

Design-science research in four phases

Supervisor gates at WBS 2.8 (framework v1.1) and WBS 4.11 (full draft) · 46 work items tracked in Jira

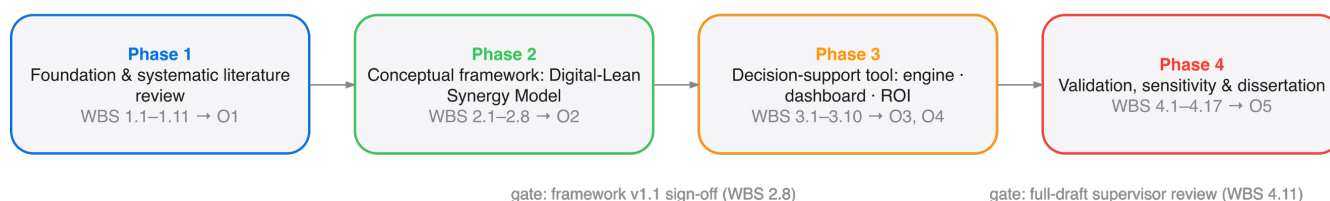


Figure 9 The four-phase design-science plan with supervisor gates and objective mapping (editable draw.io source provided).

Table 5 Phase structure, work packages and deliverables (Phase dossiers, WBS 1-4).

Phase	Purpose and core activities	Work packages	Deliverable → objective
1. Foundation	Systematic literature review; SMT process and KPI baselines; China/India EMS landscape; refined RQ and objectives	WBS 1.1–1.11	Research dossier; 19-source synthesis matrix → O1
2. Framework	Four-level Lean-maturity rubric; three-dimension readiness rubric; technology-to-KPI mapping; integration into one model	WBS 2.1–2.8	Digital–Lean Synergy Model v1.1 (supervisor-reviewed) → O2
3. Tool	Benchmark dataset; parametric ROI model; sequencing rule set; decision tree; implementation; verification	WBS 3.1–3.10	Decision-support tool v1.0 + user guide → O3, O4
4. Validation	Blind case runs against published SMT cases; sensitivity analysis; limitations; write-up	WBS 4.1–4.17	Validation report; sensitivity output; dissertation → O5

3.3 Systematic literature review protocol

The review protocol was finalised before screening and is reported alongside its results in §2.1: three review questions (RQ-A/B/C), explicit inclusion and exclusion criteria, a three-block Boolean string defined in Scopus syntax with Web of Science and IEEE Xplore equivalents, two-pass screening with logged exclusion reasons, and a PRISMA-style flow record (Page *et al.*, 2021). Twenty-six records were assessed at full text, and nineteen retained (sixteen peer-reviewed), exceeding the fifteen-source target of O1. Methodologically, two design choices matter. First, the protocol admits industry-grade grey literature (IPC benchmarks, WEF Global Lighthouse Network reports, government and analyst reporting) alongside peer-reviewed work, because the KPI baselines and the China/India industrial context that the framework needs are not carried by journals alone; grey sources are used for context and parameterisation, never as the sole support for a theoretical claim. Second, the documented query strings are re-executed through the university's institutional subscriptions so that identification counts derive from authenticated database exports; the full protocol and flow record are included as an appendix for audit.

3.4 Framework construction approach

The framework was constructed through deliberate condensation of validated instruments rather than de novo invention, so that every design choice traces to an evidence-based antecedent. The construction followed the six-phase development procedure for maturity models set out by de (de Bruin *et al.*, 2005)(Becker, Knackstedt and Pöppelbuß, 2009). At the scope stage, the unit of analysis was fixed as the individual SMT line and the purpose as prescriptive rather than descriptive (Mettler, 2011), which commits the artefact to stating what should be done next and therefore to being reproducible. At the design stage, the lean-maturity ladder compresses the five-stage capability progressions of the maturity-assessment literature into four operational levels, because the framework uses maturity as an eligibility gate rather than as a

cultural-transformation roadmap. This compression preserves the reactive-to-embedded logic while making each level verifiable on a line walk. At the populate stage, we operationalised every level and indicator with observable artefacts and quantitative thresholds, following the principle that a construct is usable only when the operations that measure it are specified (Churchill, 1979). At the test stage, the assembled framework was stress-tested against six critique questions at a structured supervisor review (WBS 2.8) covering gate placement, threshold calibration, weighting and ranking logic, and was frozen as v1.1 before validation.

3.4.1 Mapping from source instruments to framework dimensions

(Schumacher, Erol and Sih, 2016) (Tonelli et al., 2022) Reducing a nine-dimensional readiness instrument to three dimensions is a substantive design decision and is therefore documented, not asserted. Table 6 summarises the instrument's condensation and three-axis structure, indicating for each source construct whether it was merged, excluded, or relocated, and the rationale. Two decisions warrant emphasis: enterprise- and product-level constructs are excluded because they carry no line-level decision relevance and would introduce variance the framework cannot act on; and a financial dimension absent from both sources is added, because per-line capital allocation is the binding constraint the framework exists to address, and omitting it would make the instrument diagnostic rather than prescriptive.

Table 6 Documented mapping from the source readiness instruments to the three framework dimensions, with the rationale for every merge, exclusion, relocation and addition (Schumacher, Erol and Sih, 2016; Tonelli et al., 2022).

Source construct	Treatment	Framework dimension	Rationale
Strategy; Leadership; Culture; Employees	Merged	Organisational	All four concern sponsorship, ownership and workforce capability to absorb and sustain a deployment; separating them at line level would create four indicators resolved from the same evidence pack
Technology; Operations (data and IT-OT elements)	Merged	Technical	Both concern the data-and-systems substrate a technology requires; the line-level question is whether the data exist, are retained and are integrated
Governance; Products; Customers	Excluded	-	Enterprise- or product-level constructs with no line-level decision relevance; retaining them would add unactionable variance
Process-maturity axis (3D-CUBE)	Relocated	→ maturity gate	Process capability is measured by the lean-maturity rubric of §4.1.1; retaining it inside readiness would double-count the same evidence
(absent from both sources)	Added	Financial	Per-line capital allocation and payback appetite is the binding constraint the framework addresses; without it the instrument diagnoses but cannot prescribe

3.4.2 Measurement position and the aggregation rule

Readiness dimensions are rated on ordinal anchors, and ordinal ratings do not permit unrestricted arithmetic (Stevens, 1946). The applied literature is divided on how strictly to enforce this. (Jamieson, 2004) treats averaging ordinal responses as a category error, while (Norman, 2010) demonstrates that parametric treatment is robust in practice. This framework resolves the disagreement by not depending on it. For a line with dimension ratings O, T and F, two statistics are computed and always reported: the limiting dimension $R_{\min} = \min(O, T, F)$, and the composite $R_{\text{mean}} = (O + T + F) \div 3$. $R_{\min}$ is the primary decision variable and is a legitimate operation on an ordinal scale, since the minimum of a set of ordered categories is itself an ordered category and requires no assumption of equal spacing. R_{mean} is reported only as a secondary magnitude indicator, and the decision rule is constructed so that no action depends on R_{mean} alone.

(Norman, 2010; Belton and Stewart, 2002) The resulting rule is a conjunctive screen with a compensatory magnitude term, in the sense that the screen on $R_{\min}$ cannot be passed by strength elsewhere, and R_{mean} only separates actions among profiles that have already cleared the screen. Table 7 states this formally. Two consequences follow. First, the framework requires no dimension-weighting vector for the deploy, stage or defer decision, which removes the customary objection that equal weighting across readiness dimensions is asserted rather than justified. Under non-compensatory screening, the question does not arise, because no weight can offset an absent capability. Second, the rule is substantively identical to the narrative rule already applied in §4.1.2 and §4.3.3, so no previously published recommendation changes; the formalisation makes the existing logic auditable rather than altering it.

Table 7 The readiness decision rule as conjunctive screening with a compensatory magnitude term. $R_{\min}$ governs; R_{mean} separates only among profiles that have cleared the screen (Keeney and Raiffa, 1976; Belton and Stewart, 2002).

Action	Formal condition	Interpretation
DEPLOY	$R_{\min} \geq 3$ AND $R_{\text{mean}} \geq 3.5$	Every dimension is at least Defined, and the profile is above Defined overall; capital may be committed now
STAGE	$R_{\min} \geq 2$ AND $R_{\text{mean}} \geq 2.5$, and not DEPLOY	No dimension is Absent, but one binds; deployment proceeds only with a named remediation on the limiting dimension
DEFER	$R_{\min} = 1$ OR $R_{\text{mean}} < 2.5$	An Absent dimension, or an overall profile below Emerging, blocks deployment irrespective of strength elsewhere

3.4.3 Instrument reliability protocol

An instrument that yields different scores in different hands is not a measurement (Nunnally and Bernstein, 1994), and reliability is a property of an instrument in use rather than of its design. The protocol of Table 8 is therefore specified in full and in advance, with acceptance criteria fixed before any execution, so that it can be run without further design decisions and its outcome cannot be adjusted after the fact. It was not executed within this project. Executing it requires independent human raters, which engages research ethics approval; the activity was raised with the module leader and the project supervisor before any participant was approached, and, on the guidance received, it was deferred rather than compressed into the remaining timeframe (§3.8). The protocol is consequently carried as the first item of future work (§6.4, FW2), and the resulting limitation is stated at full strength in §5.5 rather than softened: the instrument's reliability is specified and testable but asserted rather than demonstrated. Where the protocol is subsequently run, the commitment recorded here holds. Coefficients are reported as obtained, and a failing indicator is flagged as a limitation rather than quietly revised, since revising an indicator after seeing its disagreement statistic would invalidate the test.

Table 8 pre-registered instrument reliability protocol. Acceptance criteria were fixed before execution. The protocol was specified but not executed within this project and is carried to §6.4 as FW2; the consequent limitation is recorded in §5.5.

Test	What it establishes	Statistic	Acceptance criterion
Inter-rater agreement on maturity level	Two assessors given the same evidence pack assign the same level	Cohen's weighted kappa, quadratic weights (Cohen, 1960)	$\kappa_w \geq 0.61$, the lower bound of substantial agreement (Landis and Koch, 1977)
Inter-rater agreement on readiness dimensions	Dimension scores do not depend on who scores them	Intraclass correlation ICC(2,k), two-way random, absolute agreement (Shrout and Fleiss, 1979)	$\text{ICC}(2,k) \geq 0.75$ on each of O, T and F
Internal consistency of each dimension	The five indicators of a dimension measure one construct	Cronbach's alpha (Cronbach, 1951)	$\alpha \geq 0.70$; a dimension below this is reported and its indicators reviewed
Test–retest stability	The same assessor reaches the same result on re-scoring	Exact-agreement rate on level; ICC on dimensions	Re-score after ≥ 14 days, blind to the original; $\geq 90\%$ exact agreement on level
Disagreement analysis	Which indicators generate disagreement, and why	Tabulation by indicator across the vignette set	Any indicator disagreeing in more than one vignette is rewritten and re-tested

3.5 Decision-support tool development method

The tool was developed verification-first and in three coordinated layers. The computation engine is a single Python module (**tool.py**) that implements the benchmark dataset, the ROI functions, the sequencing rule set, the decision-tree recommendation, and the Monte Carlo sensitivity. It is built on pandas for data handling (McKinney, 2010), NumPy for NPV, IRR, and sampling (Harris *et al.*, 2020), and scikit-learn for an interpretable surrogate decision tree fitted to the rule engine's outputs (Pedregosa *et al.*, 2011). The interaction layer upgrades the originally planned Excel-only dashboard to Streamlit, an open-source application framework that turns the existing Python engine into an interactive local web application on macOS at no cost (Streamlit, no date): the practitioner enters a line profile via form controls and receives the readiness action, limiting dimension, ranked sequence and ROI metrics live, with the Excel workbook retained as the transparent, formula-auditable ROI ledger. Two free visual-workflow alternatives were evaluated and rejected: KNIME (Berthold *et al.*, 2009) and Orange (Demšar *et al.*, 2013) offer strong no-code pipelines, but both would require re-implementing logic that already exists and is verified in Python, introducing translation

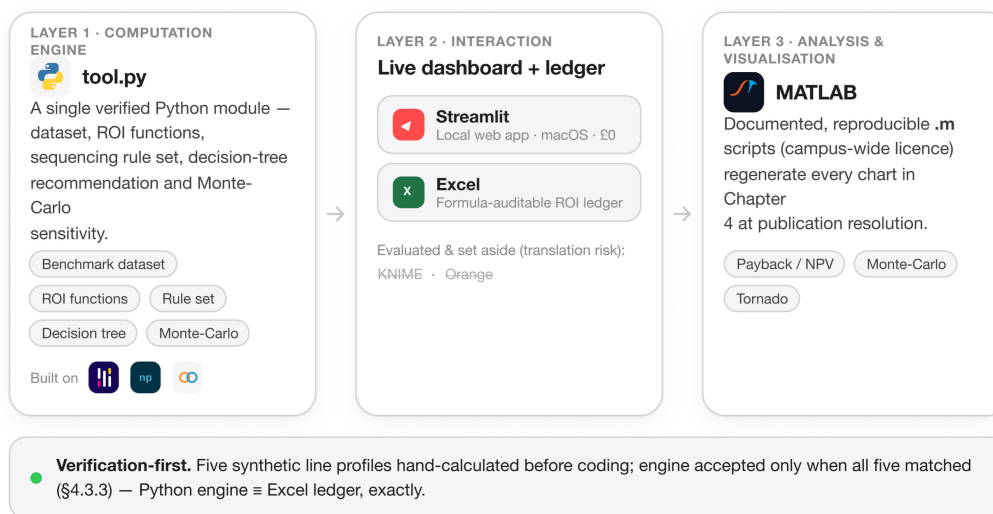
risk for no analytical gain; Streamlit reuses tool.py directly. The analysis and visualisation layer uses MATLAB, available at **no cost through the university's campus-wide licence** (MathWorks, no date): reproducible scripts read the model outputs (or the project workbook via readtable) and generate the payoff/NPV, Monte-Carlo and tornado charts reported in Chapter 4, each delivered as a documented .m file. Development was verification-first: five synthetic line profiles spanning the maturity-readiness space were hand-calculated from the rule set before coding began, and the engine was accepted only when its actions matched all five (§4.3.3); the Excel ledger was independently recalculated and agrees with the Python engine exactly.

Figure 10 maps the complete apparatus, including the cross-validation methods introduced below. The map is organised to make two distinctions auditable. The upper band is the baseline tool as originally specified, comprising the computation engine, the interaction and ledger layers, and the visualisation layer. These are not three tools but three views of one logic: the engine is the single place where a rule or a coefficient is defined; the dashboard and the ledger read from it rather than reimplementing it; and the MATLAB scripts regenerate every chart from the same dataset. That structure is what makes the exact Python-to-Excel agreement reported in §4.3.2 a meaningful check rather than a coincidence, because the two layers arrive at the same figures through independent arithmetic. The lower band is the set of methods added specifically to test the recommendation from outside the engine's own logic. Tools that were evaluated and rejected remain on the map, struck through, so the selection decision is legible to a reader rather than silently absorbed.

Digital-Lean Synergy Decision-Support Tool

Three coordinated layers, one shared logic — with three independent cross-validation methods · §3.5 / §4.6

THREE COORDINATED LAYERS · ONE SHARED LOGIC



INDEPENDENT CROSS-VALIDATION → REPORTED IN §4.6

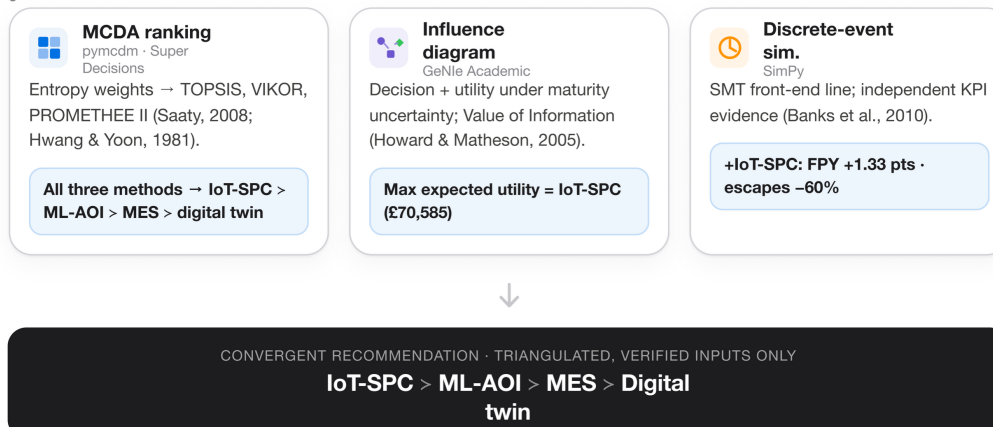


Figure 10 Integrated tool map of the decision-support apparatus: the three coordinated baseline layers (computation engine, interaction layer, visual-Integrated tool map of the decision-support apparatus: the three coordinated baseline layers (computation

engine, interaction layer, visual-analytics layer) that share a single logic, and the three independent cross-validation methods added at supervisor review, all converging on a single sequencing recommendation.

Table 9 states the same apparatus as an inventory, so that every claim made in Chapter 4 can be traced to the artefact that produced it and to the licence under which that artefact is available. Licence status is recorded because a stated objective of the framework is that a line engineer without a software budget can reproduce the analysis: every item is either open-source, free at the point of use, or available to the researcher under an institutional or academic licence, and no result in this report depends on a commercial licence the reader cannot obtain.

Layer	Tool	Role in the framework	Licence basis	Artefact	Reported in
Baseline · computation	Python 3 (pandas, NumPy, scikit-learn)	Engine: dataset, ROI functions, rule set, decision tree, Monte-Carlo	Open source (BSD)	tool.py	§4.3.1–4.3.3, §4.5
Baseline · interaction	Streamlit	Live dashboard wrapping the engine; local and publicly deployed instances	Open source (Apache 2.0)	dlsn_app.py	§4.3.1
Baseline · ledger	Microsoft Excel	Formula-auditable ROI ledger; independent arithmetic check on the engine	Institutional licence; readable in free spreadsheet software	DLSM_ROI_Ledge r.xlsx	§4.3.2
Baseline · visual analytics	MATLAB (base, no toolboxes)	Publication-resolution regeneration of every Chapter 4 chart from one dataset	Campus-wide academic licence	dlsn_data.m, viz_*.m	§4.3.2, §4.5
Cross-validation · MCDA	pymcdm (Python); Super Decisions	Entropy weighting; TOPSIS, VIKOR and PROMETHEE II rankings	Open source; free academic	mcdm_ahp_topsis.py	§4.6.1
Cross-validation · decision analysis	GeNIe Academic (BayesFusion)	Influence diagram; expected utility under maturity uncertainty	Free academic licence	DLSM_Influence_Diagram.xdsl	§4.6.2
Cross-validation · simulation	SimPy (Python)	Discrete-event model of the SMT front-end line; independent KPI evidence	Open source (MIT)	smt_line_des.py	§4.6.3
Evaluated, not adopted	KNIME; Orange	Visual workflow alternatives to the Python engine	Open source		§3.5 (rejected: translation risk)

Beyond the single engine, three independent, published cross-checking methods were adopted to triangulate the sequencing recommendation (supervisor-review action, WBS 2.8), each free or academic-free and reusing the already-verified benchmark data rather than introducing new inputs: multi-criteria decision analysis in pymcdm with Super Decisions as the analytic-hierarchy elicitation companion (Saaty, 2008; Hwang and Yoon, 1981); a decision-analytic influence diagram in **GeNIe Academic** (Howard and Matheson, 2005) and a discrete-event simulation of the SMT front-end line in **SimPy** (Banks *et al.*, 2010). These reuse tool.py's verified dataset, so they introduce no translation risk or new data, and they report convergent results in §4.6.

3.6 Validation protocol (≥3 case studies)

Because the artefact is a designed intervention rather than a description, its evaluation is explicitly positioned within the FEDS framework of (Venable, Pries-Heje and Baskerville, 2016) as an artificial, ex-post evaluation: the framework is applied to synthetic profiles, documented cases and a simulated line rather than to a live factory, a consequence of the desk-based scope declared in §1.5 and treated as a limitation in §5.5 rather than concealed. Within that positioning, the protocol addresses the validation categories recognised for design research (Straub, Boudreau and Gefen, 2004): reliability (§3.6.1), repeatability (§3.6.2), reproducibility (§3.6.3), robustness (§3.6.4) and predictive agreement (§3.6.5), with the simulation subject to its own verification protocol (§3.6.6). Each subsection states what would count as success before the evidence is gathered.

3.6.1 Reliability - do different assessors obtain the same scores?

Instrument reliability would be established under the protocol pre-registered in Table 8. This protocol is specified here in executable detail so the test can be performed exactly as designed, rather than redesigned at the point of use. Under that protocol, three raters independently score eight anonymised line vignettes spanning the maturity-readiness space, using only the instruments in **Appendix B**, without conferring and without access to one another's scores. Agreement on the assigned maturity level is reported as Cohen's weighted kappa with quadratic weights, which is appropriate because maturity levels are ordered and a two-

level disagreement is more serious than a one-level disagreement (Cohen, 1960). We interpret it using the conventional bands of (Landis and Koch, 1977). We report agreement on the readiness dimensions as ICC(2,k) under a two-way random-effects, absolute-agreement specification (Shrout and Fleiss, 1979), and internal consistency across the five indicators of each dimension as Cronbach's alpha (Cronbach, 1951). Every disagreement is tabulated by indicator, and any indicator that disagrees in more than one vignette is rewritten and re-tested rather than retained, since a single ambiguous anchor is the most common cause of instrument unreliability. This protocol was not executed within the present project for the reasons given in §3.4.3 and §3.8: it requires independent human raters and therefore ethical approval, which was not sought within the available timeframe. No coefficient of agreement is therefore reported anywhere in this dissertation, and none is implied. The study's conclusions rest on the verification, case validation and cross-validation evidence of Chapter 4, none of which depends on this test.

3.6.2 Repeatability - does the same assessor obtain the same result twice?

Intra-rater repeatability is specified as a re-scoring of the full vignette set after a minimum interval of fourteen days, with the original scores concealed. The interval is set to reduce recall without allowing the instrument to be revised between passes. Exact agreement on the assigned maturity level and ICC on the readiness dimensions would both be reported, and any vignette whose level changed between passes would be examined to identify which criterion proved unstable. Repeatability is specified separately from reliability because they fail for different reasons: unreliability indicates an ambiguous anchor, whereas unrepeatability indicates a criterion whose evidence is open to more than one reading, even by the same assessor. Like the reliability test above, this component was not executed; it is carried to §6.4 as part of FW2.

3.6.3 Reproducibility - can an independent researcher obtain the same result?

Reproducibility is addressed at three levels. At the instrument level, **Appendix B** contains the complete assessment procedure, the seven forms and the scoring rules, enabling an assessor who has not read this dissertation to conduct an assessment using the appendix alone. At the computation level, every reported financial figure is generated by the engine tool.py from the disclosed priors in Table 24 and is independently reproduced by the formula-auditable Excel ledger, allowing the arithmetic to be checked without executing Python. At the analysis level, every chart in Chapter 4 is regenerated from a single shared dataset using named MATLAB scripts, and every cross-validation result is produced by a named script or model file, with its execution environment and date recorded in Table 37. Reproducibility is therefore a property of the deposited artefacts rather than a claim in the text.

3.6.4 Robustness - do the conclusions survive plausible variation?

Robustness is assessed by varying, in turn, each class of assumption underlying the recommendation and noting whether the recommended sequence changes. Table 10 sets out the four classes and the perturbation applied to each. The distinction between them matters. Financial-prior variation tests whether the appraisal is fragile; threshold variation tests whether the framework's own design parameters are load-bearing; construct perturbation identifies borderline cases whose recommendation would flip under a one-level assessment error; and structural variation tests whether the modelling assumptions themselves drive the result. A framework robust to the first three but not the fourth would be reporting its own architecture rather than a finding.

Table 10 Robustness classes: what is varied, by how much, and what the analysis reports. Results are given in §4.5 and §4.6.

Class	What is varied	Range	Reported as
Financial priors	Capex; KPI uplift; discount rate	±30%; ±20%; 8–12%	One-at-a-time tornado screen, then 1,000-draw Monte-Carlo; rank-order stability (§4.5)
Decision thresholds	Readiness cut-offs; design-parameter maturity thresholds	R_mean 2.5/3.5 and the Table 17 design parameters across their registered ranges	Proportion of assessed profiles whose action changes; any threshold that flips an action is flagged as load-bearing (§4.5)
Construct perturbation	Assessed maturity level; each readiness dimension	±1 level; ±1 anchor	Identification of borderline lines that move between DEPLOY, STAGE and DEFER under a one-step assessment error

Class	What is varied	Range	Reported as
Structural assumptions	Influence-diagram utility structure	Partial benefit realisation below the maturity gate at 30% and 50%, and non-uniform maturity priors	Whether IoT-SPC remains the expected-utility maximiser once the premature-deployment penalty is softened (§4.6.2)

3.6.5 Predictive agreement - does the framework match documented outcomes?

(Yin, 2018) (Runeson and Höst, 2009) Validation follows a multiple-case design and adheres to the reporting and validity conventions of case research in engineering contexts. A published case is admissible only if it documents all three of: a Lean/Six-Sigma baseline or maturity indicator; at least one named deployment of the four candidate technologies; and quantified before/after values for FPY, defect PPM, OEE, cycle time, or payback. At least three admissible cases are required, with a fourth sought as a robustness check; candidates failing the criteria are downgraded to context evidence and replaced. Each case is processed by a blind-run procedure that guards against hindsight bias: a line profile is populated from the case's baseline fields only; the tool generates its recommendation before the researcher reads what the firm actually did; only then are the actual sequence and reported outcomes revealed and logged, with the tool output recorded verbatim and unedited. Where a case does not state maturity or readiness directly, they are inferred through the §4.1 rubrics under explicit, recorded inference rules, so the mapping is auditable, the principal construct-validity safeguard. Agreement is scored on the pre-registered metrics of Table 11, and every divergence is classified as (a) a readiness dimension the tool does not capture, (b) a threshold needing tuning, (c) a contextual factor outside scope, or (d) a limitation of the case data. Treat a divergence as a finding, not a failure.

Table 11 pre-registered agreement metrics for the blind case runs (validation dossier, WBS 4.4).

Agreement metric	Definition	Threshold
Sequence rank-order match	Tool's first-recommended technology vs the firm's first actual deployment; Spearman rank correlation across the full sequence	First-step match = pass; $\rho \geq 0.6$ = strong
Payback agreement	Tool payback vs reported payback / ROI where stated	Within $\pm 20\%$ = agreement
KPI-direction agreement	Tool's target KPI vs the KPI the firm improved most	Same KPI = pass
Gate correctness	Deployed technology deemed eligible at the case's maturity level	Eligible = pass

3.6.6 Simulation - verification and validation

(Sargent, 2013) (Law, 2015) The case-validation protocol above tests the framework against the outside world. A second, distinct question arises for the discrete-event model reported in §4.6.3: whether that model is credible enough for its outputs to count as evidence. Simulation results generated by an unverified model are not corroboration; they are decoration. The distinction matters here because the simulation is the only part of the apparatus that produces process-physics evidence rather than re-arranging the financial data. The protocol therefore follows the verification-and-validation framework introduced in §2.6, and the steady-state reporting conventions. Table 12 pre-registers the requirements to constrain the results rather than write them to fit.

Three requirements are imposed. First, verification by degenerate case: the model must return the analytically known answer under boundary conditions. A zero defect-generation rate must yield exactly 100% first pass yield and zero escapes, and perfect inspection must yield zero escapes irrespective of the defect rate. A model that fails either test is rejected outright, as failure indicates a coding or accounting error rather than a modelling judgement. Second, exclusion of the initial transient: statistics are collected only after a warm-up window long enough for the line to reach an arrival-departure balance, with the window justified by the measured cycle time rather than chosen for convenience. Third, replication: because each output is a random variable, we run each configuration as a set of independent replications under distinct random-number streams, and report results as a mean with a 95% confidence interval on Student's *t* rather than a single number. We treat a difference between configurations as real only when the intervals do not overlap; when they do overlap, we report the difference as indistinguishable at the chosen replication count.

Table 12: pre-registered verification-and-validation protocol for the discrete-event model (Sargent, 2013; Law, 2015). Acceptance criteria are fixed before execution; outcomes are reported in §4.6.3.

Check	Type	Procedure	Acceptance criterion	Action if failed
V1 Degenerate case zero defects	Verification	Defect-generation probability set to zero at every station	FPY exactly 100.00%; escaped DPPM exactly 0	Reject model
V2 Degenerate case perfect inspection	Verification	Detection probability set to 1.0 at pre- and post-reflow AOI	Escaped DPPM exactly 0 at the nominal defect rate	Reject model
W Warm-up exclusion	Validation of reporting	Statistics suppressed for an initial window; window justified against mean panel cycle time	Exclusion window at least five times the mean panel cycle time	Lengthen window and re-run
R Replication and interval reporting	Validation of reporting	Independent replications under distinct seeds; Student's t interval, $df = n - 1$	$n \geq 30$ replications; every reported mean carries a 95% confidence interval	Increase n
S Effect significance	Interpretation rule	Confidence intervals compared between configurations	Effect claimed only where intervals do not overlap	Report as indistinguishable

3.7 Assumptions and parameter ranges

(Saltelli *et al.*, 2008) (Brealey, Myers and Allen, 2020) Every assumption is logged in the assumptions register with a base value, a plausible range, and a source. The ranges drive the sensitivity analysis, using one-at-a-time (OAT) tornado screening followed by Monte Carlo sampling. The financial formulation follows standard investment appraisal practice. Table 13 reproduces the register. Fixing the ranges before validation prevents post hoc tuning to fit case outcomes.

Table 13 Assumptions and sensitivity-range register (Phase-3 dossier, WBS 3.3). The register is extended in this revision to carry the framework's own decision thresholds, the readiness cut-offs of Table 8 and the design-parameter maturity thresholds the readiness cut-offs of Table 8 and the design-parameter maturity thresholds identified in Table 17 so that they are subject to the same sensitivity treatment as the financial priors and are not exempt from scrutiny merely because they belong to the framework rather than to the appraisal.

Assumption	Base	Range	Source / note
Discount rate	10%	8–12%	Corporate WACC range, EMS sector
Appraisal horizon	5 years	fixed	Standard appraisal horizon
Capex tolerance	band mid-point	$\pm 30\%$	Vendor quote spread
KPI uplift tolerance	design prior	$\pm 20\%$	Effect-size uncertainty
Cpk gate (L3 / L4)	1.33 / 1.67	1.33–1.67	Six-Sigma capable-process thresholds
Readiness: deploy now	≥ 3.5	and no dimension < 3	Framework rule R4 (§4.2)
Readiness: defer	< 2.5	or any dimension = 1	Framework rule R4 (§4.2)

3.8 Ethics and risk considerations

The study design minimises ethical exposure: it is desk-based, uses only published secondary sources, and reproduces no proprietary, facility-identifying, or commercially sensitive data. **No human participants were involved at any stage, and no primary data were collected, so no application to the University's research-ethics review system was required for the work reported here, and none was made.** One activity was contemplated that would have engaged human participants: the instrument- reliability study specified in § 3.4. and 3.4.3, in which independent raters would each score a set of fictional line vignettes to compute inter- rater agreement. The material was invented for the purpose and described no real facility, line or person; no personal data would have been collected, and only anonymised numerical scores would have been reported. Despite this low-risk profile, the position taken here is that the involvement of human participants is itself the trigger for review, rather than the level of risk they face. The activity was therefore raised with the module leader and the project supervisor before approaching any participant. The guidance received confirmed that ethical approval was required, and the study was deferred to future work (§ 6. 4, FW 2) rather than executed without it. The decision is recorded explicitly because the alternative, proceeding first and seeking approval afterwards, was available and was rejected. Nothing in the validation reported in Chapter 4 depends on the deferred study. Risk is managed on three fronts, consistent with the module's learning outcomes on commercial, environmental and safety risk. Commercially, the principal risk the framework itself addresses, capital misallocation across an SMT fleet, is bounded by declaring all effect sizes as editable priors and quantifying the robustness of every recommendation through sensitivity analysis. The dissertation states plainly that outputs are decision support, not guarantees. Environmentally, sequencing

capital towards lines that can exploit it reduces wasted equipment spend, energy consumption and electronic scrap. The digital-twin evidence base includes a documented $2.7 \times$ energy-consumption reduction on an SMT line (Karanjkar *et al.*, 2018). Desk research does not directly address safety obligations, but the technology profile cards flag deployment-stage considerations such as operator interaction with inspection equipment and change management during commissioning for any organisation applying the framework.

Chapter 4 Results: Framework, Tool and Validation

This chapter presents the results of Phases 2 and 3 and the validation apparatus for Phase 4: the Digital–Lean Synergy Model (§4.1), the sequencing decision logic it generates (§4.2), the working decision-support tool with its verified outputs (§4.3), the case-validation results structure (§4.4), the sensitivity analysis (§4.5) and the independent cross-validation of the recommendation using multi-criteria decision analysis, a decision-analytic influence diagram and a discrete-event simulation (§4.6). All quantitative outputs shown were produced by the live model; the Python engine and the independently recalculated Excel ledger agree exactly, and every effect size remains an editable design prior until confirmed against case data.

4.1 The Digital - Lean Synergy Model

The model integrates three components into a single decision artefact through a three-step maturity gate, modulate, quantify and rank (Figure 11). Maturity gates determine which technologies a line is eligible for; readiness modulates whether an eligible technology should be deployed now, staged, or deferred; and KPI mapping, coupled with the ROI model, quantifies and ranks the eligible options. Each component is presented in turn.

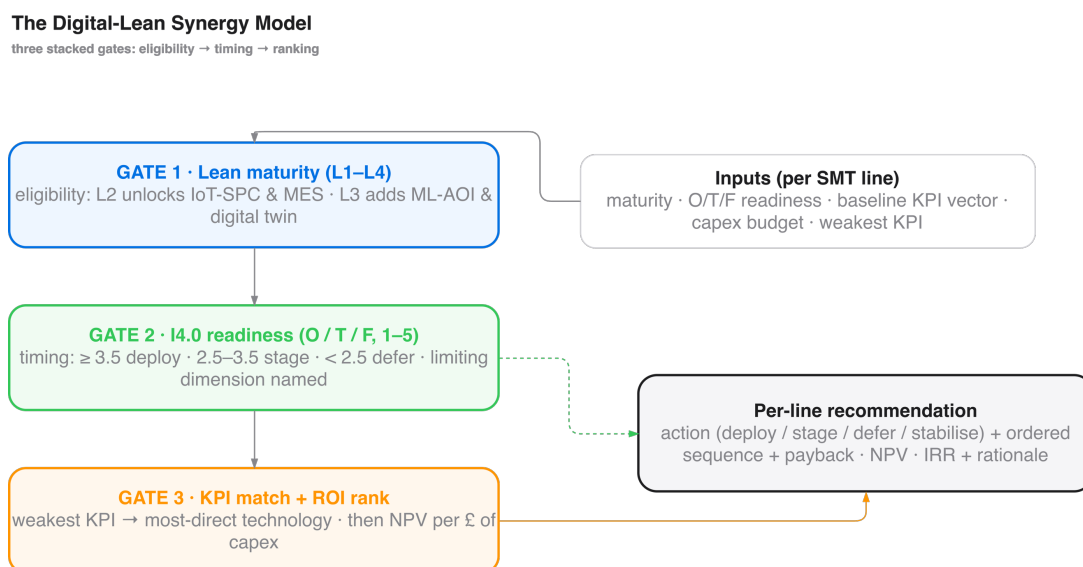


Figure 11 The Digital–Lean Synergy Model: three stacked gates converting a line profile into a per-line recommendation.

4.1.1 Four-level Lean-maturity model

4.1.1.1 Definition and role of the construct

Within this framework, lean maturity is defined operationally as the demonstrated capability of a single SMT line to keep its critical process parameters within specification, to measure that capability from its own records, and to improve it through a documented method. The definition is deliberately narrow: it says nothing about culture, leadership or enterprise strategy, because the framework uses maturity solely as an eligibility gate for capital deployment on one line, and any element of the construct that cannot be evidenced from line records would make the gate unverifiable. Maturity is required because the candidate technologies consume process stability rather than create it. Machine-learning inspection learns from adjudicated defect data, which an unstable line does not reliably produce; closed-loop statistical process control adjusts against a characterised distribution, which presupposes that one exists; a digital twin optimises a process whose

behaviour is already understood. Deploying any of these onto an uncharacterised process therefore produces cost without the mechanism by which benefit is supposed to arise, which is the substantive reason rule R1 directs an L1 line to stabilisation rather than to purchase.

4.1.1.2 Who assesses, and with what evidence

A qualified assessor conducts the assessment, defined for this purpose as a process or quality engineer with at least two years of SMT process experience and read access to the line's inspection, test, and event records. The qualification is stated because the instrument requires judgement about which parameters are critical and whether a document constitutes evidence. An assessor unable to make those judgements would generate scores that are precise but not meaningful. The assessor works from an evidence pack rather than from an interview: station logs, capability studies, downtime and changeover records, and the practice artefacts listed in Table 15. No maturity score in this framework may be assigned on assertion; where a required record does not exist, the criterion is recorded as not met, since the absence of a record is itself evidence about the line's measurement maturity.

4.1.1.3 Operational definitions of the maturity indicators

Table 14 specifies each indicator completely: its formula, the manufacturing data source, the sampling window and the aggregation rule. This specification ensures that two assessors computing the same indicator from the same records obtain the same number, the minimum condition for reproducible level assignment.

Indicator	Formula	Data source	Sampling window	Aggregation rule
First-pass yield (FPY, %)	$FPY = (B_{pass} + B_{start}) \times 100$, where B_{pass} is boards passing every inspection and test station with no rework	AOI, SPI and functional-test station logs; MES or line-event log where present	20 consecutive production days, minimum 20 shifts; a window containing an NPI ramp is rejected and re-drawn	Pooled count ratio over the window ($\sum B_{pass} + \sum B_{start}$), not the mean of daily yields
Process capability (Cpk)	$Cpk = \min[(USL - \bar{x}) / 3\sigma, (\bar{x} - LSL) / 3\sigma]$, σ estimated from within-subgroup variation	SPI output for solder-paste volume; placement-machine logs for placement offset	≥ 25 rational subgroups of $n = 5$ consecutive panels within the same window	Computed per designated critical parameter; the line takes the minimum across parameters
Overall equipment effectiveness (OEE, %)	$OEE = A \times P \times Q$, with $A = \text{run} \div \text{planned time}$, $P = (\text{ideal cycle} \times \text{total count}) \div \text{run time}$, $Q = \text{good} \div \text{total count}$	MES or line-event log; where absent, manual downtime log plus station counters	Same 20 production days; planned production time only	Computed from pooled totals across the window, not the mean of daily OEE
Defects per million opportunities (DPMO)	$DPMO = [D + (U \times O)] \times 10^6$, D = confirmed defects after false-call adjudication	AOI and functional-test defect records post-adjudication; opportunity count from the bill of materials	20-day window for the level value; 6 consecutive calendar months for the trend	Pooled over the window; trend assessed on 6 monthly points by Mann–Kendall
Changeover time (minutes)	Median elapsed time from last good board of run n to first good board of run $n+1$	Changeover log or MES event timestamps	≥ 10 changeovers within the preceding 3 months	Median, not mean, to limit the influence of exceptional events
Practice-artefact compliance (PAC, %)	$PAC = (\text{criteria evidenced} + \text{criteria applicable}) \times 100$ against the level checklist	Layered process-audit records, standard work, VSM, control charts, 8D records	Current state at assessment; documents dated within the preceding 6 months	Percentage of the level's listed artefacts for which documentary evidence exists

Two aggregation choices require justification. Pooled ratios are used for FPY, OEE, and DPMO rather than the means of daily values because daily figures are computed over unequal denominators, and their unweighted mean would over-weight low-volume days; pooling weights each day by its actual production, which is the quantity the framework cares about. The median is used for changeover because changeover distributions are right-skewed by exceptional events such as a missing feeder, and a mean would report the exception rather than the process. Cpk is estimated from within-subgroup variation in accordance with standard capability practice (Montgomery, 2019; AIAG, 2005), so that the statistic reflects inherent process variation rather than drift between subgroups.

4.1.1.4 Level definitions and assignment rule

Table 15 states, for each level, the mandatory quantitative criteria and the practice-artefact requirement. Every qualitative statement in the original rubric has been replaced by a measurable condition, and the two kinds of criterion are separated so that neither can be traded against the other: a line cannot buy its way past a capability threshold with paperwork, nor past an artefact requirement with a good Cpk.

Level	Mandatory quantitative criteria (all must hold)	Practice-artefact requirement	14.0 unlocked
L1 Reactive	Assigned when the L2 criteria are not met; equivalently $Cpk < 1.00$ on any critical	Not assessed at this level	None stabilise first

Level	Mandatory quantitative criteria (all must hold)	Practice-artefact requirement	I4.0 unlocked
	parameter, or OEE < 60%, or FPY not computable over the full window		
L2 Managed	Cpk ≥ 1.00 on every critical parameter; OEE ≥ 60%; FPY computable over the full 20-day window; DPMO computable from adjudicated records	PAC ≥ 80% against the L2 checklist: standard work in current revision; current-state VSM ≤ 12 months old; control charts on paste volume and placement offset; defect Pareto with stated review frequency	IoT-SPC · MES
L3 Optimised	Cpk ≥ 1.33 on every critical parameter; OEE ≥ 75%; DPMO trend significantly negative over 6 months (Mann–Kendall, $\alpha = 0.05$, one-sided) or ≥ 3 consecutive monthly falls; median changeover reduced ≥ 20% against a documented baseline	PAC ≥ 80% against the L3 checklist: ≥ 2 completed DMAIC or Kaizen projects with before/after data in 12 months; documented capability studies; pull or Kanban evidence; closed 8D records	+ ML-AOI · Digital twin
L4 Digital–Lean	Cpk ≥ 1.67 on every critical parameter; OEE ≥ 85%; DPMO < 10 at final test; line-level event data captured automatically and retained ≥ 12 months	PAC ≥ 90% against the L4 checklist: real-time SPC dashboards in operational use; MES line-event capture; closed-loop SPI/AOI feedback configured; predictive-maintenance actions triggered from condition data	Full stack in closed loop

Assignment rule. A line is assigned the highest level for which all mandatory quantitative criteria and the practice-artefact requirement hold. Testing proceeds upward from L2 and stops at the first level not fully met. Because the criteria are conjunctive, strength in one criterion cannot compensate for failure in another, which is what makes the gate meaningful: a line with an excellent Cpk and no capability studies is not a measured line.

Mixed profiles and the limiting criterion. Where a line satisfies every criterion of level $n+1$ except one, it is not simply recorded as level n . It is recorded as 'Ln, limited by [criterion]', because the failing criterion determines whether a specific technology is nonetheless viable. A line that meets every L3 condition except OEE is in a materially different position from one that fails on Cpk, since the technologies differ in which capability they consume. Table 16 states this technology-relevant dominance rule, converting what would otherwise be arbitrary rounding down into an explicit, auditable exception.

Table 16 Technology-relevant criterion dominance: which maturity criteria govern eligibility for each technology, and the single-criterion exception where a line is one criterion short of the gate level.

Technology	Governing criteria	Rule where the line is one criterion short
IoT-enabled SPC	Cpk; computability of FPY and DPMO	Eligible at L2 provided Cpk ≥ 1.00 and both KPIs are computable; an OEE shortfall does not bar it, because closed-loop process control acts on capability, not availability
MES	OEE; data availability; changeover discipline	Not eligible on a capability shortfall alone; MES addresses flow, so an OEE shortfall is the condition it is bought to remedy and does not bar eligibility
ML-based AOI	Cpk; DPMO level and trend; adjudicated AOI defect data (indicator T5)	Eligible at 'L2, limited by OEE' provided Cpk ≥ 1.33 and adjudicated AOI data exist; not eligible without adjudicated data at any level, since the model has nothing to learn from
Digital twin	OEE; line-level data granularity; Cpk	Not eligible below full L3; the twin consumes the data foundation the other technologies create, so no single-criterion exception applies

4.1.1.5 Provenance of the thresholds

Threshold provenance is stated explicitly so readers can distinguish values inherited from established standards from those the researcher chose. The distinction matters because the two carry different evidential weight and warrant different treatment: literature-anchored values are defensible by citation, whereas design parameters are defensible only if their status is disclosed and their influence is tested. All values in the second class are entered in the assumptions register in §3.7 and are varied in the decision-threshold robustness analysis in §3.6.4. Table 17 records the provenance of each threshold together with its status.

Table 17 Threshold provenance: literature-anchored values are distinguished from declared design parameters. Design parameters enter the §3.7 register and the §3.6.4 robustness analysis.

Threshold	Provenance	Status
Cpk 1.00 / 1.33 / 1.67	Conventional capability boundaries in statistical quality control: 1.33 is the capable-process convention and 1.67 the best-in-class target (Montgomery, 2019; AIAG, 2005)	Literature-anchored
OEE 85% at L4	Nakajima's (1988) world-class benchmark, decomposed as availability 90% × performance 95% × quality 99%	Literature-anchored

Threshold	Provenance	Status
OEE 60% and 75% at L2 and L3	Interpolated bands between an unmanaged line and the world-class benchmark, consistent with the published SMT ranges compiled in Table 3	Design parameter sensitivity-tested
DPMO < 10 at L4	Consistent with the near-zero-defect expectation at final test in high-volume consumer electronics	Design parameter sensitivity-tested
PAC ≥ 80% (≥ 90% at L4)	Audit-compliance convention; set to require substantial rather than total artefact presence, so that one missing document does not veto a level	Design parameter sensitivity-tested
Changeover reduction ≥ 20%	Set as a material, auditable improvement against a documented baseline rather than an absolute time, because absolute changeover times are product-dependent	Design parameter sensitivity-tested
20-day sampling window	Chosen to span at least four working weeks and thereby cover normal shift, crew and product-mix rotation	Design parameter sensitivity-tested

4.1.1.6 Why an independent assessor obtains the same level

The instrument builds reproducibility at four points and then tests it rather than claiming it. First, every indicator has a single admissible formula, data source, window and aggregation rule (Table 14), so two assessors cannot legitimately compute an indicator differently. Second, level assignment is a conjunctive test of stated numerical criteria (Table 15), which removes the weighting judgement through which most rubric disagreement arises. Third, the practice-artefact requirement is a documentary checklist rather than an opinion, so the question 'does this line have a current-state VSM?' has an answer independent of the assessor. Fourth, mixed profiles resolve according to the stated dominance rule of Table 16 rather than through discretion. What remains genuinely judgemental is the designation of critical parameters and the acceptance of a document as evidence, and those two residuals are precisely what the reliability protocol of §3.4.3 is designed to quantify; the resulting coefficients are reported in §3.6.1 rather than assumed.

4.1.2 Industry 4.0 readiness assessment instrument

4.1.2.1 Definition and role of the construct

Where maturity asks whether the process can be digitised, readiness asks whether the organisation can absorb, operate, and sustain the technology once it is installed. The two differ because they fail differently and are remedied differently: a capability shortfall is corrected by process engineering, whereas a readiness shortfall is corrected by ownership, data infrastructure, or capital policy. Readiness is assessed across three dimensions (organisational, technical and financial), each through five indicators, and each indicator carries an explicit evidence requirement. The evidence requirement is the design feature that distinguishes this instrument from a self-assessment questionnaire: an instrument whose anchors describe attitudes (measures optimism), whereas one whose anchors describe artefacts (measures the organisation). Accordingly, no score above the lowest anchor may be recorded without a named, dated document; where the document does not exist, the indicator scores 1 regardless of what is asserted in the interview.

4.1.2.2 Indicators, evidence requirements and scoring rules

Tables 18 to 20 set out the five indicators for each dimension. Each indicator uses the same 1-to-5 anchored scale, in which 1 denotes that the capability is absent, 3 that it is defined and documented, and 5 that it is embedded and demonstrably operating. The anchors are expressed as observable states rather than as degrees of agreement, so the assessor is asked what exists rather than how good things feel. Table 18 covers the organisational dimension, Table 19 the technical dimension, including the inspection-data condition on which ML-AOI eligibility rests, and Table 20 the financial dimension.

Table 18 Organisational readiness: five indicators with evidence requirements and anchored scoring rules.

Indicator	What is assessed	Documentary evidence required	Scoring rule 1 → 5
O1 Ownership	Whether a named individual is accountable for the deployment and its benefits	Role description, organisation chart or project charter naming the individual and the scope	1 none; 2 informal or shared; 3 named owners with defined scope; 4 named owners with budget authority; 5 owners with budget authority and benefit accountability

Indicator	What is assessed	Documentary evidence required	Scoring rule 1 → 5
O2 Roadmap	Whether deployment is planned rather than opportunistic	Dated digitalisation roadmap covering the line, with milestones	1 none; 2 intentions stated verbally; 3 documented roadmap ≤ 12 months old; 4 roadmaps with resourced milestones; 5 roadmaps reviewed on a stated cycle
O3 Workforce stability	Whether the team that will operate the technology persists	HR record of 12-month voluntary turnover for the line's technical and engineering staff	1 turnover > 40%; 2 > 30%; 3 ≤ 30%; 4 ≤ 20%; 5 ≤ 10% (bands are design parameters, §3.7)
O4 Competence and training	Whether operating capability is built systematically	Training matrix or skills register for the line with dated completion records	1 none; 2 ad-hoc on-the-job only; 3 structured training with records; 4 competences assessed and recorded; 5 competences reassessed on a stated cycle
O5 Change-management practice	Whether prior technical changes were absorbed successfully	Records of the two most recent line changes: plan, communication and post-implementation review	1 no records; 2 records for one change; 3 plan and communication evidenced; 4 post-implementation review evidenced; 5 reviews with actions closed

Table 19 Technical readiness: five indicators with evidence requirements and anchored scoring rules. Indicator T5 is the precondition for ML-based AOI referenced in Table 16.

Indicator	What is assessed	Documentary evidence required	Scoring rule 1 → 5
T1 Data capture	Whether line-level process data are captured at all, and how	Sample data export with timestamps showing source station and capture frequency	1 none; 2 manual transcriptions; 3 automated exports at batch level; 4 automated captures at unit level on key stations; 5 unit-level capture line-wide
T2 Data retention and access	Whether historical data exist to support analysis or learning	Historian or database record showing retention period and accessible field list	1 none; 2 < 1 month; 3 ≥ 3 months on key assets; 4 ≥ 12 months across the line; 5 ≥ 12 months with queryable access for engineering
T3 IT–OT integration	Whether shop-floor systems connect to business systems	Interface or architecture diagram plus evidence of an operating data flow	1 isolated; 2 manual file transfer; 3 partial integration on key assets; 4 line-wide integration; 5 integration with automated exception handling
T4 Analytical capability	Whether the organisation can use the data it holds	Role descriptions of on-site analytical staff; examples of analyses performed online data	1 none; 2 spreadsheet analysis only; 3 analysts on site; 4 data-science capability available to the line; 5 capabilities embedded with a delivery record
T5 Inspection-data quality	Whether inspection data are adjudicated and usable the precondition for ML-AOI	False-call adjudication records and defect-classification scheme with an agreed taxonomy	1 none; 2 unadjudicated pass/fail only; 3 adjudicated with defect classes; 4 adjudicated, classified and retained ≥ 6 months; 5 adjudicated datasets validated for model training

Table 20 Financial readiness: five indicators with evidence requirements and anchored scoring rules.

Indicator	What is assessed	Documentary evidence required	Scoring rule 1 → 5
F1 Capital provision	Whether per-line capital is allocated rather than contested case by case	Approved capital budget or plan identifying a per-line envelope	1 none; 2 case-by-case only; 3 defined per-line envelope; 4 envelopes with headroom for the candidate technology; 5 multi-year envelope
F2 Appraisal method	Whether investments are appraised on a defined basis	Documented appraisal method or template showing the metrics used	1 none; 2 paybacks informally applied; 3 documented payback hurdles; 4 paybacks plus NPV; 5 payback, NPV and risk-adjusted appraisal
F3 Payback appetite	Whether the organisation's hurdle accommodates the candidate technology	Stated hurdle in the appraisal policy, compared with the technology's modelled payback	1 undefined; 2 ≤ 12 months only; 3 ≤ 24–30 months; 4 ≤ 36 months; 5 strategic horizons beyond 36 months permitted
F4 Incentive and offset position	Whether external mechanisms offset the capital requirement	Documentation of applicable incentive schemes and their status for the site	1 none identified; 2 identified only; 3 applied for; 4 approved; 5 approved and drawn down on a prior investment
F5 Benefit tracking	Whether realised benefit is measured after deployment, closing the appraisal loop	Post-implementation benefit review for a prior investment	1 none; 2 informal; 3 reviews performed for at least one prior investment; 4 standard practice; 5 standard practice feeding revised appraisal priors

4.1.2.3 Dimension scoring and the decision rule

Each dimension score is the median of its five indicator scores. The median is used rather than the mean because indicator scores are ordinal. The median is a legitimate operation on an ordered scale, whereas the mean presumes equal spacing between anchors, which the anchors do not guarantee (Stevens, 1946); §3.4.2). The median also resists a single unusually high or low indicator from dominating the dimension. Because the median can conceal a single absent capability within an otherwise strong dimension, the assessment also records the indicator range and flags any dimension whose indicators span three or more anchors. Any indicator scoring 1 is named explicitly in the recommendation record, irrespective of its

dimension median. This is a deliberate control for a known weakness of median aggregation, not an incidental reporting choice.

The three-dimensional scores O, T and F then enter the conjunctive screening rule formalised in §3.4.2 and stated in Table 8: the limiting dimension $R_{\min} = \min(O, T, F)$ governs the action, and the composite $R_{\text{mean}} = (O + T + F) \div 3$ separates DEPLOY from STAGE only among profiles that have already cleared the screen. Both statistics are reported for every assessment, together with the name of the limiting dimension and its lowest indicator, because the limiting indicator is the remediation target the framework returns.

Worked continuity check. For the profile Synthetic-A used throughout this chapter, the organisational indicators score 3, 3, 3, 2 and 4, giving a median of 3; the technical indicators 2, 2, 3, 1 and 2, giving a median of 2; and the financial indicators 4, 4, 3, 5 and 4, giving a median of 4. Hence $O = 3$, $T = 2$, $F = 4$, $R_{\min} = 2$, with technical limiting, and $R_{\text{mean}} = 3.00$. Applying Table 8, $R_{\min} \geq 2$ and $R_{\text{mean}} \geq 2.5$, but $R_{\min} < 3$, so the action is STAGE with remediation named against the technical dimension, specifically against indicator T4, which scores 1. This exactly reproduces the action and rationale reported in §4.3.3 and returned by the tool in Figure 14, so the formalisation changes no published result.

4.1.3 Technology-to-KPI mapping

The mapping (Table 21) is deliberately KPI-specific: IoT-SPC and ML-AOI are quality-leverage technologies, MES is a flow-leverage technology, and the digital twin is an availability and optimisation technology. This separation allows the tool to match a line's weakest KPI to the technology that most directly improves it, rather than defaulting to the most advanced option. The profile cards behind the table also encode a natural build order. IoT-SPC and MES establish the data foundation at L2; ML-AOI exploits inspection data once the process is capable at L3; and the digital twin sits on top, consuming the data the others generate. The decision tree formalises this dependency structure. Effect sizes derive from the evidence in §2.3 and enter the ROI model as adjustable priors.

Table 21 Technology-to-KPI mapping with capital bands, maturity gates and payback priors (framework v1.1, WBS 2.5–2.6; capex from Proposal Table 1).

Technology	Primary KPIs moved	Capex (per line)	Gate · payback prior	Indicative effect basis (§2.3)
IoT-enabled SPC	FPY · defect PPM · Cpk	£80–120k	L2+ · 12–24 mo	Real-time drift detection before defects form
MES	Throughput · WIP · schedule adherence · OEE availability	£150–250k	L2+ · 18–36 mo	Digital dispatch, traceability, downtime capture; lighthouse evidence
ML-based AOI	Solder PPM · false-call rate · escapes	£100–180k	L3+ · 12–30 mo	False-call classification on production data; CNN X-ray inspection
Digital twin	OEE · predictive maintenance · energy	£80–200k	L3+ · 24–48 mo	Discrete-event what-if optimisation; energy reduction on an SMT line

4.1.4 Assessment procedure

The two instruments and the mapping are combined into a seven-step procedure that a qualified assessor who has not read this dissertation can execute using only the forms in **Appendix B**. The procedure is stated because a framework that specifies levels without indicating how a manager determines them is a taxonomy, not a decision aid: the practical contribution claimed in §1.6 depends on the assessment being executable, and Table 22 with Figure 12 is what makes that claim testable.

Table 22 *The seven-step assessment procedure: action, owner, input and output at each step. Forms B1–B7 are reproduced in Appendix B.*

Step	Action	Who	Input	Output
1	Collect the 20-day data window and the evidence pack: station logs, SPI and placement records, downtime and changeover logs, practice artefacts	Assessor with production supervisor	Line records; MES or station exports	Form B1 completed, every field stamped with source and extraction date
2	Compute the six maturity indicators using the formulae, windows and aggregation rules of Table 14; run the Mann–Kendall trend test on DPMO	Assessor	Form B1	Form B2: indicator values with the computation shown, plus the trend result
3	Test the criteria of Table 15 upward from L2; assess practice-artefact compliance; apply the mixed-profile and dominance rules	Assessor	Form B2 and artefact checklist B3	Form B4: maturity level, limiting criterion where applicable, eligible technology set

Step	Action	Who	Input	Output
4	Score the fifteen readiness indicators of Tables 18–20, referencing the document that evidences every score above 1	Assessor with the owner of each function	Evidence pack; questionnaire B5	Form B5: fifteen indicator scores, each with an evidence reference
5	Compute each dimension score as the median of its five indicators; flag any dimension whose indicator range is ≥ 3 ; compute $R_{\min}$ and R_{mean}	Assessor	Form B5	Form B6: O, T, F; $R_{\min}$ with the limiting dimension and indicator named; R_{mean}
6	Apply the decision rules: the maturity gate of Table 15 with the dominance rule of Table 16, then the readiness rule of Table 8	Assessor	Forms B4 and B6	Eligible set and action: DEPLOY, STAGE, DEFER or STABILISE
7	Rank the eligible, ready technologies by the rule set of §4.2 R5 weakest-KPI match, then R6 net present value per pound of capex	Assessor, reviewed by the line manager	Eligible set; line KPI vector; capex envelope	Form B7: ordered recommendation with payback, NPV, IRR and rationale

The procedure includes three control points to prevent the assessment from devolving into estimation. At step 1, record any data field that cannot be sourced as unavailable rather than estimated, because an estimated input produces a recommendation whose provenance cannot be traced. At step 4, any indicator scored above 1 without a document reference is reduced to 1, thereby enforcing the evidence principle of §4.1.2.1 at the point of scoring rather than leaving it as an instruction. At step 6, the action and every criterion that determined it are recorded together, so that a reviewer can reconstruct the decision without re-running the assessment. Figure 12 presents the same procedure as a flowchart, including the three terminating branches and the two decision points at which the readiness rule is applied.

Assessment procedure — from line records to a ranked recommendation

Seven steps · instruments are Forms B1–B7 in Appendix B · decisions shown as diamonds

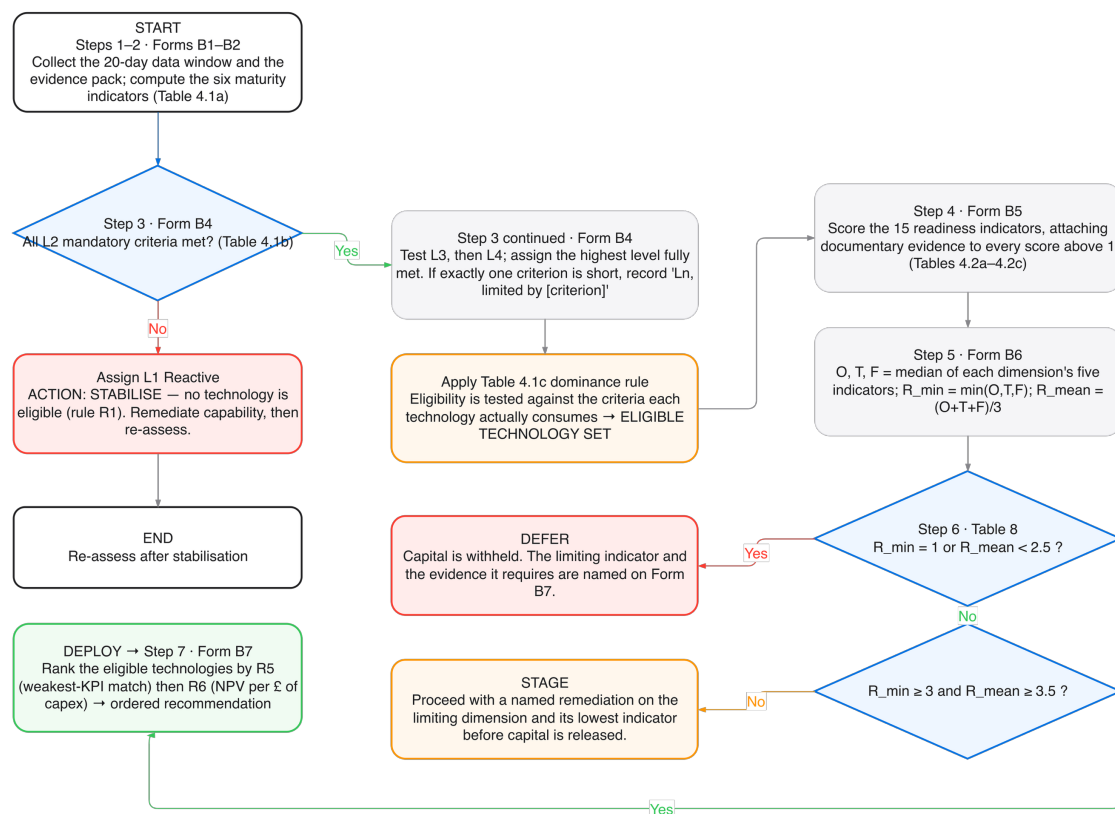


Figure 12 The assessment procedure as a flowchart: seven steps, three terminating actions and the two decision points of the readiness rule (Table 8). Diamonds are decisions; each step's form is named so the chart can be followed using Appendix B alone.

4.2 Sequencing decision logic (rule set and decision tree)

The framework's narrative logic is formalised as six rules (Table 23) and encoded as the decision tree in Figure 13; each leaf returns an ordered recommendation with its rationale. We also fit an interpretable scikit-learn surrogate tree to the rule engine's outputs purely as an explanatory artefact; the transparent rule set, not a learned model, makes the decisions.

Table 23 Sequencing rule set v1 (Phase-3 dossier, WBS 3.4).

Rule	Statement	Source
R1	IF maturity = L1 (Reactive) THEN STABILISE defer all I4.0 until Lean/SPC make the line measurable and capable	Proposal; Phase-1 gap
R2	IF maturity ≥ L2 THEN IoT-SPC and MES become eligible (subject to capex budget)	Proposal Table 1; §4.1.1
R3	IF maturity ≥ L3 THEN ML-AOI and digital twin additionally become eligible	Proposal Table 1; §4.1.1
R4	Composite readiness ≥ 3.5 AND no dimension < 3 → DEPLOY; 2.5–3.5 → STAGE with remediation of the limiting dimension; < 2.5 or any dimension = 1 → DEFER	§4.1.2
R5	Map the line's weakest KPI to its most-direct technology: PPM or yield → IoT-SPC; throughput → MES; OEE → digital twin. Tie-break: where the weakest KPI is defect PPM and indicator T5 evidences an adjudicated, labelled AOI dataset, ML-AOI takes precedence over IoT-SPC provided the Table 16 gate is satisfied; without adjudicated data ML-AOI is not eligible at any level.	§4.1.3
R6	Among eligible, ready technologies, rank by KPI-match first, then by NPV per £ of capex	§4.3.2

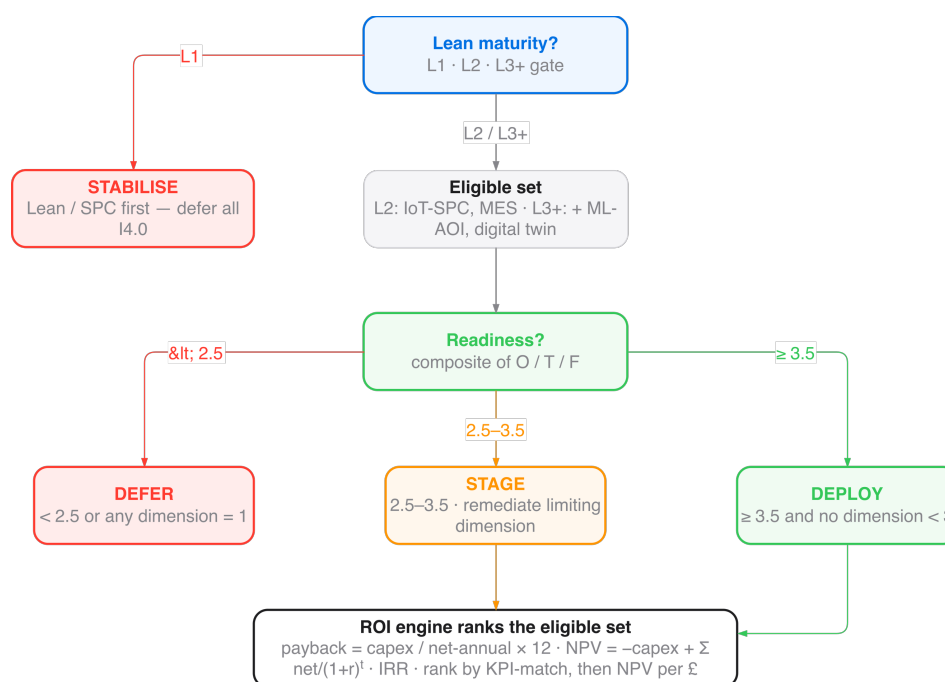


Figure 13 The sequencing decision tree: maturity gates, readiness modulates, the ROI engine ranks.

4.3 The decision-support tool

4.3.1 Architecture (Python engine · Streamlit dashboard · MATLAB visual analytics)

Tool v1.0 comprises three coordinated layers that share a single logic, replacing the index's originally planned Python-plus-Excel architecture as set out in §3.5. The engine is written in **Python 3.9** (Python Software Foundation, 2024), and the independent ledger is in Microsoft Excel (Microsoft Corporation, 2024). The engine, **tool.py**, implements the benchmark dataset, ROI functions, rule engine, decision-tree recommendation, surrogate tree, and Monte Carlo module (McKinney, 2010; Pedregosa *et al.*, 2011; Harris *et al.*, 2020). The interaction layer is a Streamlit application that wraps the engine as a local web dashboard on macOS at **zero licence cost (Streamlit, no date)**: the user enters maturity, O/T/F readiness, baseline KPIs, capex budget and weakest KPI, and reads back the composite readiness, limiting dimension, readiness action, ranked sequence and per-step payback, NPV and IRR, with the rationale string; the Excel workbook remains the transparent, formula-driven ROI ledger and recalculates identically to the engine. The analysis layer is a suite of documented MATLAB scripts run under the university's campus-wide licence (MathWorks, no date) which regenerate every chart in this chapter at publication resolution. **Data flow: line profile** →

maturity gate → **readiness modulation** → **KPI-to-technology mapping** → **ROI ranking** → **ordered recommendation with rationale**.

The framework operated two instances of the interaction layer, and the distinction between them is methodologically relevant rather than incidental. The local instance runs the engine on the researcher's machine and is the configuration used to generate every figure reported in this chapter; it is the reference instance because it is fully under version control and its inputs are traceable to the artefacts listed in Table 10. The second instance is a publicly deployed copy of the same code, hosted at no cost on the Streamlit Community Cloud platform (Snowflake Inc., 2026) and accessible at <https://dlsm-dashboard-cgwx9tn32gzktriky3khw.streamlit.app>.

(Hevner *et al.*, 2004; Peffers *et al.*, 2007) The framework was deployed for three reasons that bear directly on the contribution claimed in §1.6. First, reproducibility: an examiner or practitioner can apply the framework's logic to their own line profile without installing Python, which turns the tool from a described artefact into an inspectable one and meets the design-science expectation that the artefact, not merely its description, is available for evaluation. Second, practitioner accessibility: the stated audience is a process or industrial engineer responsible for a single line, and a browser-accessible interface removes the software-environment barrier that would otherwise confine the framework to readers who can configure a Python environment. Third, verifiability of the equivalence claim: because the deployed instance imports the same engine module rather than reimplementing its logic, any divergence between the published tables and the live tool would be immediately apparent, making the claim of a single shared logic falsifiable rather than merely asserted. Two limitations of the deployed instance are recorded for completeness. It is a free-tier deployment and therefore carries no availability guarantee, so the local instance and the archived source remain the authoritative record; and it holds no user data, since every calculation is performed within the session and nothing is persisted, which is what allows the tool to be published without the ethics exposure discussed in §3.8. Figure 14 shows both instances executing the same worked profile and returning identical output. Appendix F provides a user guide to the deployed instance, stating each input, its admissible values, and the form on which it is determined.

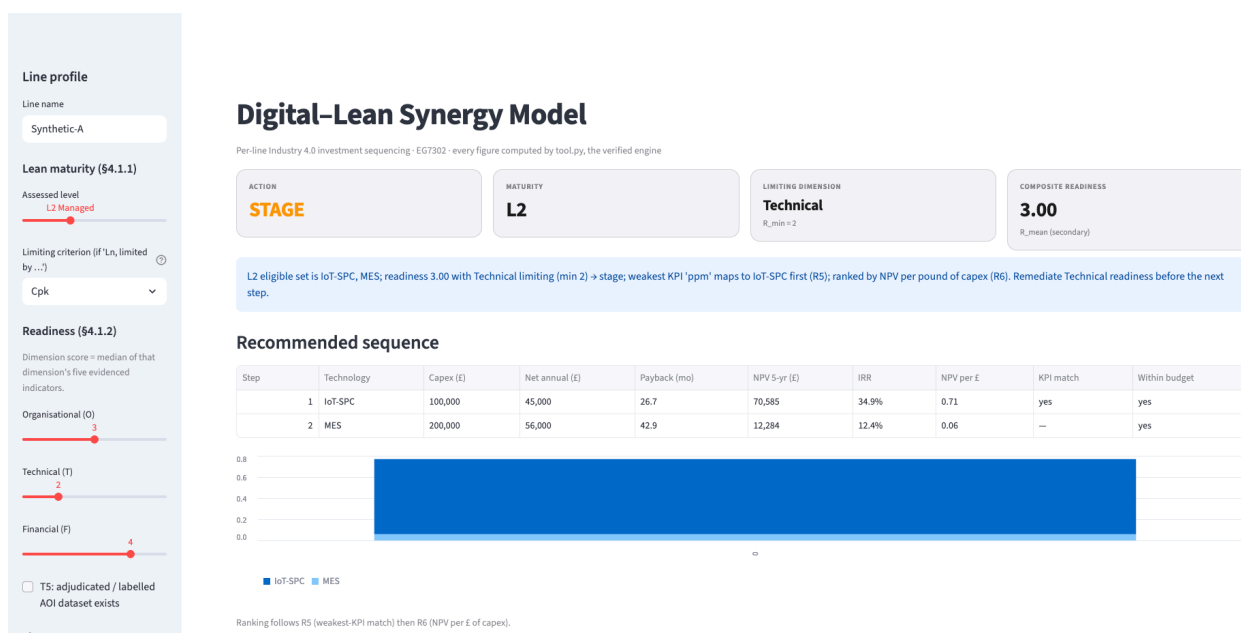


Figure 14 The interaction layer executing the Synthetic-A worked profile: (a) the local instance, in which all reported results were generated; (b) the publicly deployed instance. Both return action STAGE, limiting dimension Technical, composite readiness 3.00 and the sequence IoT-SPC then MES, matching Table 25 exactly. The two instances import the same engine module rather than duplicating its logic.

4.3.2 ROI / NPV / payback model

The parametric model takes capex, an annual benefit fraction and an opex fraction per technology, a discount rate and a horizon, and computes net annual benefit (benefit – opex), payback in months (capex ÷ net annual × 12, undefined if net annual ≤ 0), five-year NPV at the discount rate and IRR (Brealey, Myers and Allen, 2020). Table 24 and Figure 15 report the live model's outputs at base priors (10% discount, capex at band

mid-points). IoT-SPC and ML-AOI offer the strongest returns per pound; MES pays back more slowly because of its higher capital band; and the digital twin is NPV-negative over five years at base priors, consistent with its position as the last-deployed optimisation layer, whose benefit depends on the data foundation the others create. These are design priors to be confirmed against the Phase-4 case data, not measured guarantees.

The benefit and opex fractions are the priors from which every reported financial figure derives. They are disclosed in Table 24 rather than held in the code, so a reader can regenerate any published value by hand. The benefit fraction is the modelled annual benefit expressed as a proportion of capital cost, and the opex fraction is the modelled annual running cost on the same basis; net annual benefit is their difference multiplied by capex. The disclosed values reproduce every previously published figure exactly: for IoT-enabled SPC $(0.55 - 0.10) \times £100,000 = £45,000$; for ML-based AOI $(0.50 - 0.10) \times £140,000 = £56,000$; for MES $(0.40 - 0.12) \times £200,000 = £56,000$; and for the digital twin $(0.35 - 0.12) \times £140,000 = £32,200$. Disclosure matters because these fractions are the single largest source of uncertainty in the appraisal. The tornado screen of §4.5 identifies the benefit fraction as the dominant influence on payback, and a reader who cannot see them cannot judge the result. They are declared design priors registered in §3.7, not measurements on the lines analysed, and they are the first quantity a firm applying the framework should replace with its own pilot data. The NPV-per-£ column reconciles rule R6, which ranks by return per unit of capital rather than by absolute net present value. On that measure, IoT-enabled SPC returns 0.71 and ML-based AOI 0.52, so IoT-SPC ranks first even though ML-AOI has the marginally higher absolute NPV of £72,284 against £70,585. Without the per-pound column, the ranking reported in §4.3.3 would appear inconsistent with Table 24, and the column therefore makes the rule as implemented visible rather than leaving the reader to infer it.

Table 24 Worked ROI outputs at base priors (10% discount, capex at band mid-points, five-year horizon), produced by the live model; the Python engine and the Excel ledger agree exactly (WBS 3.2). The benefit and opex fractions are disclosed so that every figure can be regenerated by hand; they are declared design priors registered in §3.7, not measurements.

Technology	Capex (mid)	Benefit fraction	Opex fraction	Net annual	Payback (mo)	NPV 5-yr	NPV per £ capex	IRR
IoT-enabled SPC	£100,000	0.55	0.10	£45,000	26.7	£70,585	0.71	34.9%
ML-based AOI	£140,000	0.50	0.10	£56,000	30.0	£72,284	0.52	28.6%
MES	£200,000	0.40	0.12	£56,000	42.9	£12,284	0.06	12.4%
Digital twin	£140,000	0.35	0.12	£32,200	52.2	-£17,937	-0.13	4.8%

Payback, months (base priors)

Five-year NPV at 10% discount

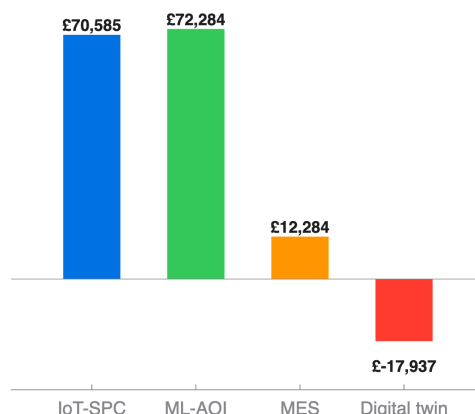
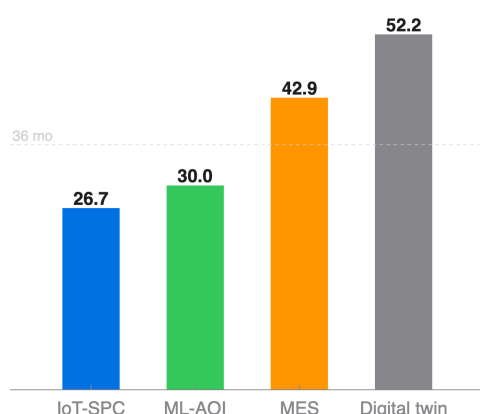


Figure 15 Payback in months (left axis, bars) and five-year NPV in pounds sterling (right axis, markers) per technology at base priors, generated at 300 dpi by viz_roi_payback_npv.m from the shared dataset dlsm_data.m. The chart renders Table 24 rather than an independent calculation: it is regenerated from the same figures the engine produces, so the two cannot diverge. The digital twin is the only technology with negative NPV over the five-year horizon at base priors.

4.3.3 Worked example and verification

For the worked profile Synthetic-A, an L2 Managed line with readiness O 3 / T 2 / F 4 (composite 3.00, limiting dimension Technical) and defect PPM as its weakest KPI, the tool returns: action STAGE; first step IoT-SPC

(£100,000; payback 26.7 months; NPV £70,585; IRR 34.9%); second step MES (£200,000; 42.9 months; £12,284; 12.4%); ML-AOI and digital twin deferred pending L3. The rationale string reads: L2 eligible set is IoT-SPC and MES; readiness 3.00 with technical limiting → stage; weakest KPI 'ppm' maps to IoT-SPC first; remediate Technical before MES. Verification followed the §3.5 protocol: five synthetic lines spanning the maturity-readiness space were hand-derived from the rule set and compared with the tool's outputs (Table 25). All five actions match. One discrepancy arose and is reported for traceability: the initial hand-note for Synthetic-B expected ML-AOI first, but the tool returned IoT-SPC. On review, the tool is correct, because rule R5 maps the weak KPI 'ppm' to IoT-SPC as its most-direct technology, and the hand-expectation, not the tool, was corrected.

Table 25 Verification of the tool against five hand-calculated synthetic line profiles: 100% action agreement (WBS 3.9).

Line	Profile (maturity; O/T/F; weak KPI)	Tool action	Tool first step	Match
Synthetic-A	L2; 3/2/4; ppm	STAGE	IoT-enabled SPC	✓
Synthetic-B	L3; 4/4/3; ppm	DEPLOY	IoT-enabled SPC	✓ (hand-note corrected)
Synthetic-C	L1; 2/2/2; yield	STABILISE	(defer all)	✓
Synthetic-D	L3; 4/4/4; oee	DEPLOY	Digital twin	✓
Synthetic-E	L2; 2/2/2; throughput	DEFER	MES (when ready)	✓

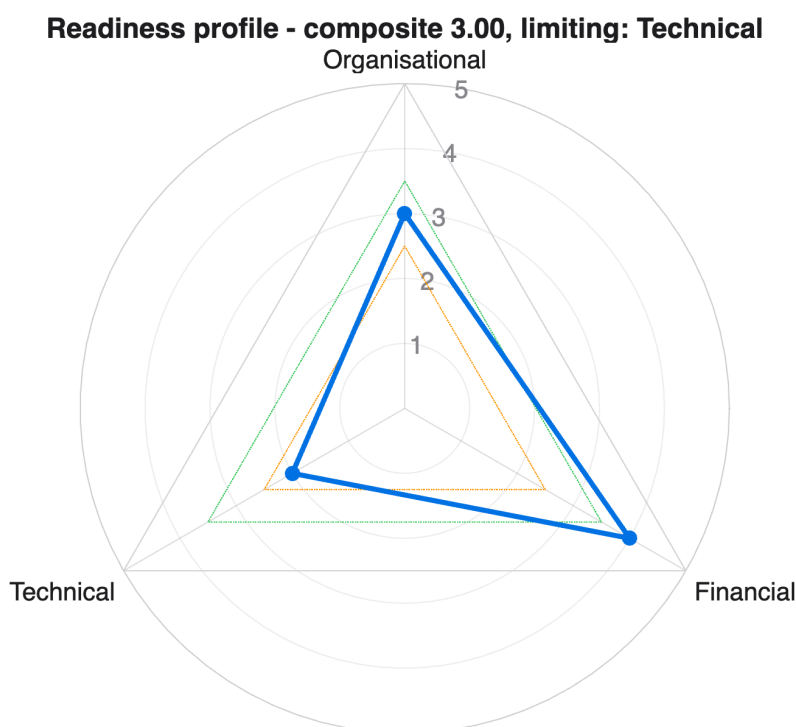


Figure 16 renders the same profile as a readiness radar. The visualisation is included because it makes the framework's conjunctive screening rule legible in a way the composite score alone does not: the composite of 3.00 is a comfortable mid-range figure, yet the profile is manifestly asymmetric, and it is the technical vertex at 2, not the average, that determines the STAGE action. Reporting the composite without the shape would invite exactly the compensatory reading the rule set is designed to prevent, allowing strong financial readiness to offset a technical deficit that would, in fact, stall deployment. The radar is therefore the recommended presentation format when the tool's output is taken to a line review, since the remediation target is read directly off the smallest vertex.

4.4 Framework validation against published cases

4.4.1 Case selection and data extraction

The validation applies the §3.6 protocol to the candidate shortlist in Table 26. Candidates C2 and C3 report directional or aggregate effects rather than clean per-line before/after values; if a candidate cannot fully populate the extraction template, it is downgraded to context evidence and replaced, and six to eight candidates are read to net the three fully admissible cases the protocol requires. For each admitted case, a single extraction record captures the firm profile, the inferred baseline maturity and O/T/F readiness (with the inference rule logged), the baseline KPI vector, the technologies and sequence actually deployed, reported outcomes, and any stated capex, everything the tool needs as input, plus the outcome it is judged against. The screening was executed, and its yield is reported in the final column of Table 26: of the four candidates read at full text, one satisfied every admissibility condition, and three did not. The three exclusions share one

cause, which is itself a finding rather than an inconvenience. The published SMT literature reports at the campus, firm, or dataset level, whereas the framework decides at the line level, so the evidence base largely lacks the granularity the method requires. That mismatch directly corroborates the gap claimed in §2.8 and is revisited in §5.1.

Table 26: candidate validation cases and qualification rationale (validation dossier, WBS 4.1).

Case	Candidate (published source)	Why it qualifies	Maps to	Screening outcome (§3.6.5)
C1	IoT-driven digital twin on a working SMT-PCB line (Karanjkar et al., 2018)	SMT line; named digital-twin deployment; quantified energy/throughput effect	Digital-twin path	ADMITTED single line resolvable; baseline maturity inferable from stated absence of measurement; named deployment; quantified before/after on energy, throughput and cycle time
C2	Foxconn Chengdu Global Lighthouse designation (World Economic Forum, 2021)	Smartphone EMS; named MR/AI/IoT deployment; OEE +17%, labour +200%	MES / IoT-SPC path	EXCLUDED campus-level reporting; no line or cell resolvable, no baseline value behind the +17% OEE delta. Retained as context evidence
C3	Firm-level Lean–I4.0 adoption cases (Tortorella and Fettermann, 2018; Rossini et al., 2019)	Lean baseline plus I4.0 adoption with operational-performance deltas	Maturity-gating evidence	EXCLUDED firm-level survey constructs across 110 and 108 firms; no per-line before/after. Retained as context evidence for the maturity-gating premise
C4 (reserve)	IPC APEX process-control case papers (AOI/SPI yield studies)	SMT inspection; before/after defect-PPM or false-call data	ML-AOI / IoT-SPC path	EXCLUDED reports classifier accuracy and false-call reduction on a dataset, not a line KPI delta against a maturity baseline

4.4.2 Tool recommendation versus reported outcome

The blind run was executed against the admitted case on 04 August 2026, and the result is reported in Table 27. The procedure set out in §3.6.5 was followed: a line profile was populated from the case's pre-intervention statements only; each inferred field was logged with the rule that produced it; and the tool was run before scoring the reported outcome. One qualification on blindness is entered honestly. With a single researcher reading a single paper, the run is blind by construction rather than blind in time. The profile is derived only from baseline fields under rules fixed in advance, and the derivation is auditable, but the researcher cannot un-read the paper. Independent profile construction is recorded as a refinement in §6.4.

The case is an SMT-PCB assembly line at a facility in Mysuru, India, running legacy equipment. The baseline the source describes is unambiguous: machines expose 'very limited real-time performance information over proprietary interfaces'; output quality 'relies on the alertness and skill of human operators'; and 'power consumption and machine-activity patterns are not known'. No Cpk, first-pass yield, OEE, or DPMO is reported because none was measured. Under Table 15, this is not missing data but a determination: L2 requires first-pass yield computable over the window and Cpk of at least 1.00 evidenced, and neither condition is met, so the line is assessed at L1 Reactive. Technical readiness scores 1 on the same evidence, with organisational and financial readiness at 2.

The firm hosted an externally funded instrumentation project rather than committing a per-line capital envelope. The tool returns STABILISE with no technology eligible under rule R1. What the firm did is the substance of the test. Its first action was not to buy a technology but to instrument every machine with energy meters, three-axis vibration sensors and proximity sensors, and to collect three months of data to infer machine state and energy consumption. Only then could the digital twin be parameterised, with its processing delays set to match the measured line. On the pre-registered metrics, this yields two passes and one partial. Gate correctness passes: the framework holds that a digital twin consumes a data foundation it does not create, and this case shows the twin was unbuildable until the foundation existed. KPI-direction agreement passes: the framework maps the twin to OEE, maintenance and energy, and energy per board is exactly what improved, by a factor of 2.7 with throughput held. Payback agreement is not assessable, because the source states neither capital cost nor payback, and the cell is recorded as such rather than estimated.

The sequence match is only partial, due to a genuine limitation of the framework rather than of the case. Rule R1 returns 'no technology eligible' at L1, whereas the correct action on this line demonstrably was to deploy

sensing. The framework conflates two things that the case separates: instrumentation, which is how a process becomes measurable, and closed-loop IoT-enabled statistical process control, which presupposes that it already is. Both currently sit within a single technology profile gated at L2. The divergence is therefore classified under class (b) of §3.6.5, a rule requiring refinement, and the refinement is specific: framework v1.2 should split the IoT-SPC profile into an instrumentation tier admissible at L1, priced at the sensing and gateway subset of the capital band, and a closed-loop control tier that remains gated at L2. That change would leave every result in this chapter intact, because no line analysed here sits at L1, while making the L1 recommendation actionable rather than merely prohibitive.

Two limits on what this single case can carry are stated plainly. The reported improvement is a digital-twin simulation prediction rather than a measured post-installation result: the buffering solution was proposed and evaluated in the model, not installed on the line. And one admitted case cannot establish external validity; it can corroborate the gate logic and surface a rule defect, both of which it did, but the three further cases the protocol sought were not obtainable from the published literature at line resolution. The construct-perturbation check of §3.6.4 was run alongside: re-assessing the same line at L2 instead of L1 returns DEFER rather than STABILISE, still withholding capital and still naming the technical dimension as limiting, so the capital decision is stable against a one-level assessment error.

Table 27 Validation comparison record: outcome of the blind run against the admitted case (Karanjkar et al., 2018), scored on the metrics pre-registered in Table 12. Every cell is populated from the executed run and the source; no value is estimated; metrics the source cannot support are recorded as not assessable.

Pre-registered metric (Table 12)	Tool output blind run	What the case reports	Outcome
Gate correctness was the deployed technology eligible at the case's maturity?	Digital twin requires full L3 (Table 16); the line assessed L1, so the twin is not eligible, and no technology is released (R1)	The twin could not be parameterised until every machine had been instrumented and three months of state and energy data collected; the model's delays were then set to match the measured line	PASS the dependency the gate encodes is exactly the one the case had to satisfy before the twin was usable
KPI-direction agreement did the case improve the KPI the tool targets?	Weakest KPI entered as OEE/utilisation; Table 21 maps OEE, predictive maintenance and energy to the digital twin	Energy per board reduced 2.7× (2,427 → 910 kJ) with throughput held at ≈40 boards per hour	PASS the improved KPI is the one the framework maps to the deployed technology
Sequence rank-order match tool's first step versus the firm's first action	STABILISE: no technology named; the instruction is to make FPY, Cpk and OEE measurable before capital is committed	First action was sensor instrumentation to make machine state and energy observable; the twin followed only after three months of data	PARTIAL the direction agrees (measure first) but the tool names no technology where the case deployed sensing. See the divergence class below
Payback agreement tool payback versus reported payback	Not released: no technology is eligible at L1, so no payback is computed	No capital cost or payback is stated in the source	NOT ASSESSABLE recorded as such rather than estimated
Robustness of the assessment (§3.6.4 construct perturbation)	Re-run at L2 instead of L1: action DEFER, eligible set IoT-SPC then MES, still withholding capital (R_min = 1, Technical limiting)		STABLE a one-level maturity assessment error does not change the capital decision

4.5 Sensitivity analysis

(Saltelli et al., 2008) Robustness was quantified with a 1,000-draw Monte Carlo over the registered ranges (capex ±30%, KPI uplift ±20%, discount rate 8–12%), using the method fixed in §3.7. Table 28 and Figure 17 report the payback distributions. The central finding is rank-order stability: IoT-SPC remains the fastest payback and the digital twin the slowest across the sampled range, so the recommended sequence is robust to plausible input error, even though individual payback estimates shift by roughly four to twelve months between P50 and P80. The complementary one-at-a-time tornado (viz_tornado_sensitivity.m) identifies which single input most affects each technology's outcome: in the parametric formulation, the benefit-fraction (KPI uplift) dominates payback swing, while capex moves NPV and budget eligibility, and the discount rate moves NPV rather than undiscounted payback. Phase 4 extends this analysis to the validated cases, additionally perturbing maturity by ±1 level and composite readiness by ±0.5 to test whether borderline lines flip between deploy, stage and defer.

Table 28 Monte-Carlo payback distribution per technology: 1,000 draws over the registered ranges (WBS 3.8).

Technology	Mean payback (mo)	P50 payback (mo)	P80 payback (mo)
IoT-enabled SPC	27.2	26.7	31.2
ML-based AOI	30.6	30.1	35.3
MES	44.0	43.0	51.7
Digital twin	53.8	52.3	63.8

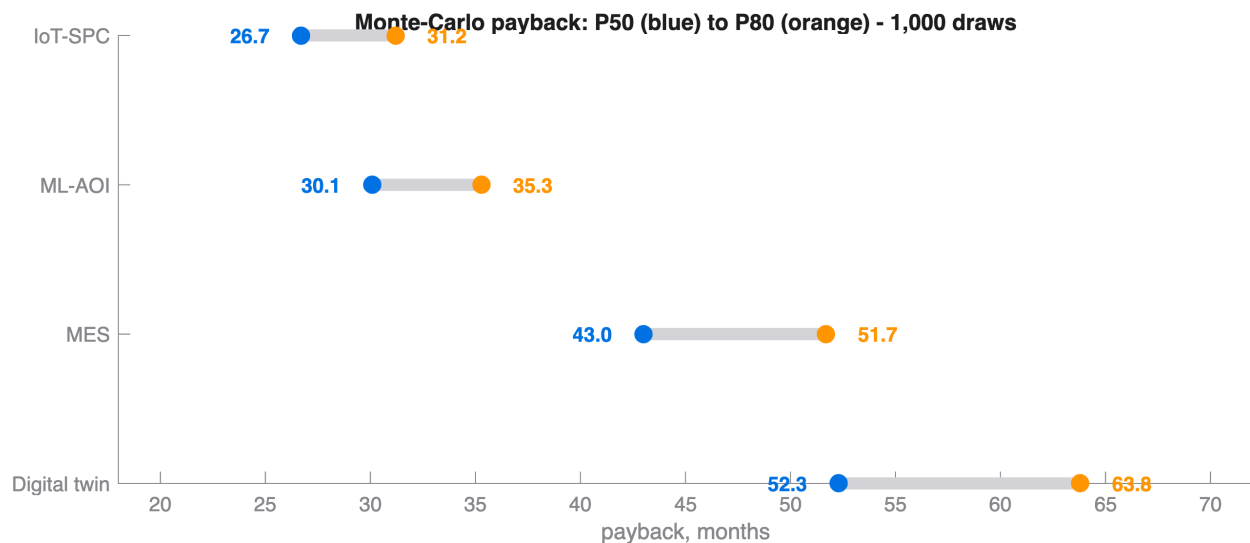


Figure 17 Monte-Carlo payback distribution per technology: P50 and P80 intervals over 1,000 draws across the registered ranges of Table 13. Rank order is preserved in every region of the sampled space, which is the robustness claim of interest: the sequence, not the point estimate, is what the framework outputs.

The Monte Carlo result shows that the ranking is stable, but not why it moves. The complementary one-at-a-time tornado screen answers that question by perturbing each registered input in isolation while holding the remainder at base, and it is reported for IoT-SPC in Figure 18 because that technology carries the first-step recommendation and therefore bears the greatest consequence if its appraisal is wrong. The decomposition shows that the outcome is governed principally by the assumed annual benefit fraction and by capex, and only weakly by the discount rate across the registered 8-12% band. This has a direct practical implication for the framework's use: the estimate that most deserves a line engineer's scrutiny before committing capital is the benefit fraction, which is a design prior taken from the technology profile rather than a measured value on the line in question, and which the assumptions register of §3.7 flags accordingly. The insensitivity to the discount rate is itself a useful finding, since it means the recommendation does not turn on a corporate finance parameter that varies between firms and over which a line engineer has no control.

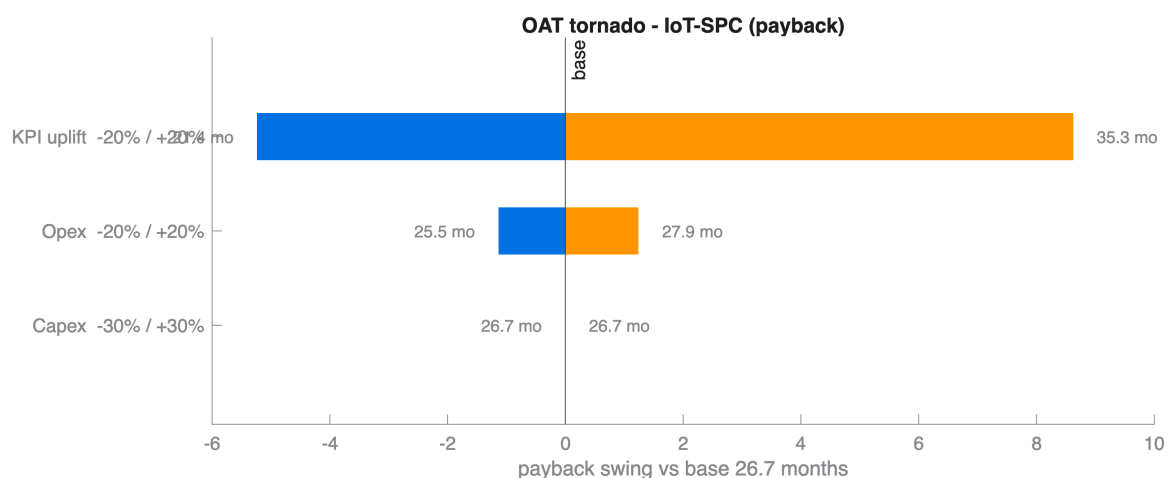


Figure 18 One-at-a-time tornado screen for IoT-enabled SPC: the swing in payback produced by moving each registered input across its range of Table 13 while the others are held at base. Bars are ordered by absolute influence; the benefit fraction and capex dominate; the discount rate does not.

4.6 Cross-validation by MCDA, simulation and decision analysis

The payback-ordered sequence of §4.2-4.3 was re-examined using three independent, published methods to test whether the recommendation is an artefact of a single metric (supervisor-review action, WBS 2.8).

Each method is free or academic-free, fits the locked Python/MATLAB tool stack, and is driven entirely by the verified figures in Table 24, so every input is traceable, and no values are estimated. The three methods, their setups and their full outputs are presented in §4.6.1-4.6.3 and synthesised in §4.6.4.

4.6.1 Multi-criteria decision analysis (MCDA)

(Lotfi and Fallahnejad, 2010)(Hwang and Yoon, 1981)(Opricovic and Tzeng, 2004)(Brans and Vincke, 1985)(Kizielewicz, Shekhovtsov and Sałabun, 2023)(Saaty, 2008)The four technologies were treated as alternatives, and the three verified financial outputs in Table 24 (payback, a cost criterion minimised; five-year NPV and IRR, benefit criteria maximised) formed the decision matrix. Because subjective pairwise weights would introduce unverifiable numbers, we derived objective weights from the decision matrix using the Shannon-entropy method, which assigns more weight to criteria that discriminate more sharply between alternatives (Table 29). We then ranked the alternatives using three established, methodologically distinct procedures (TOPSIS, VIKOR, and PROMETHEE II) implemented in the open-source pymcdm library, with Super Decisions available for analytic-hierarchy elicitation where expert weighting is preferred. Convergent rank order across three independent methods provides evidence of a robust recommendation.

Table 29 Objective criterion weights derived from the Table 24 decision matrix by the Shannon-entropy method (mcda_ahp_topsis.py).

Criterion	Optimisation	Entropy weight
Payback (months)	Minimise (cost)	0.390
NPV, five-year (£)	Maximise (benefit)	0.293
IRR (%)	Maximise (benefit)	0.317

The three methods yield the same ordering, and each place IoT-enabled SPC first and the digital twin last, with no ties (Table 30). The consensus sequence IoT-SPC > ML-AOI > MES > digital twin exactly reproduces the tool's payback-based order, confirming that the recommendation does not depend on a single payback metric but holds up under a full multi-criteria treatment weighted by the data's own discriminating power.

Table 30 MCDA rankings under entropy weights: three independent methods converge on one order (mcda_ahp_topsis.py; TOPSIS higher = better, VIKOR Q lower = better, PROMETHEE II net flow higher = better).

Technology	TOPSIS (rank)	VIKOR Q (rank)	PROMETHEE II (rank)	Consensus
IoT-enabled SPC	0.991 (1)	0.000 (1)	0.804 (1)	1
ML-based AOI	0.860 (2)	0.135 (2)	0.529 (2)	2
MES	0.326 (3)	0.654 (3)	-0.333 (3)	3
Digital twin	0.000 (4)	1.000 (4)	-1.000 (4)	4

```
shree@Mac ~ % python3 mcda_ahp_topsis.py
/Library/Developer/CommandLineTools/usr/bin/python3: can't open file '/Users/shree/mcda_ahp_topsis.py': [Errno 2] No such file or directory
shree@Mac ~ % cd "/Users/shree/Downloads/Dissertation/Dissertation Draft/Python Tool"
shree@Mac Python Tool % ls
DLSM_Influence_Diagram.xdsl      README.md
mcda_ahp_topsis.py              smt_line_des.py
shree@Mac Python Tool % python3 mcda_ahp_topsis.py

MCDA WEIGHTING & RANKING OF I4.0 TECHNOLOGIES (EG7302, verified data)

Decision matrix (Core File Table 4.5):
Technology      Payback (months)      NPV 5-yr (GBP)      IRR (%)
IoT-enabled SPC      26.7                  70,585              34.9
ML-based AOI        30.0                  72,284              28.6
MES                42.9                  12,284              12.4
Digital twin        52.2                  -17,937              4.8

Entropy-derived criterion weights (from the data above):
Payback (months)      0.390
NPV 5-yr (GBP)       0.293
IRR (%)              0.317

Technology      TOPSIS rank      VIKOR Q rank      PROM.II rank
IoT-enabled SPC      0.991      1      0.000      1      0.804      1
ML-based AOI        0.860      2      0.135      2      0.529      2
MES                0.326      3      0.654      3      -0.333      3
Digital twin        0.000      4      1.000      4      -1.000      4

Consensus sequence (mean rank across TOPSIS / VIKOR / PROMETHEE II):
1. IoT-enabled SPC      (mean rank 1.00)
2. ML-based AOI        (mean rank 2.00)
3. MES                (mean rank 3.00)
4. Digital twin        (mean rank 4.00)

Payback-only order (tool.py, for comparison): IoT-SPC -> ML-AOI -> MES -> Digital twin
```

Figure 19 reproduces the executed console output from which Tables 28 and 29 were transcribed. It is included rather than summarised because the audit value lies in the sequence it makes visible: the decision matrix is printed first, exactly as taken from Table 24; the entropy weights are then derived from that matrix; and only then are the three rankings computed. The analyst enters no weights at any point, which distinguishes this cross-check from conventional analytic-hierarchy elicitation and removes the principal source of subjectivity that critics of weighted-sum methods cite (Roy, 1996). The final line of the output re-states the engine's own payback-ordered sequence for direct comparison.

4.6.2 Decision-analytic influence diagram

(Howard and Matheson, 2005) (BayesFusion, 2024) The first investment choice was modelled as an influence diagram in GeNIe Academic, which represents a decision under uncertainty with a chance node, a decision node, and a utility node. Lean maturity is the uncertain chance node (L1, L2 or L3); the first-investment decision selects one technology; and the utility is the verified five-year NPV in Table 24, where the technology's maturity gate is met, and the technology's verified capital outlay is treated as stranded (a negative utility of $-\text{capex}$) where it is deployed prematurely below its gate (Table 31). This single structural assumption, that the benefit cannot be realised in an ungated, unstable process, is the framework's central thesis expressed in decision-analytic terms, and every utility value is a verified figure, not an estimate.

Table 31 Influence-diagram utility matrix: verified five-year NPV where the maturity gate is met, else $-\text{capex}$ (premature). Source: Table 24 NPV and capex (DLSM_Influence_Diagram.xdsl).

Technology	Utility at L1	Utility at L2	Utility at L3
IoT-enabled SPC	£70,585	£70,585	£70,585
ML-based AOI	−£140,000	−£140,000	£72,284
MES	−£200,000	£12,284	£12,284
Digital twin	−£140,000	−£140,000	−£17,937

Solving the diagram under a non-informative (uniform) maturity prior yield the expected utilities in Table 32. IoT-enabled SPC maximises expected utility at £70,585 because it is eligible at every maturity level, whereas every technology with a higher maturity gate is value-destroying in expectation once the risk of premature deployment is priced in. The model therefore selects IoT-SPC as the robust first investment under maturity uncertainty, consistent with the framework and the MCDA. It also highlights the value of information in confirming a line's maturity on a walk before committing capital, since resolving the uncertainty removes the premature-deployment penalty from the eligible options.

Table 32 Expected utility of each first investment under a uniform maturity prior; IoT-SPC dominates (DLSM_Influence_Diagram.xdsl).

First investment	Expected utility (uniform maturity prior)
IoT-enabled SPC	£70,585 (maximum selected)
MES	−£58,477
ML-based AOI	−£69,239
Digital twin	−£99,312

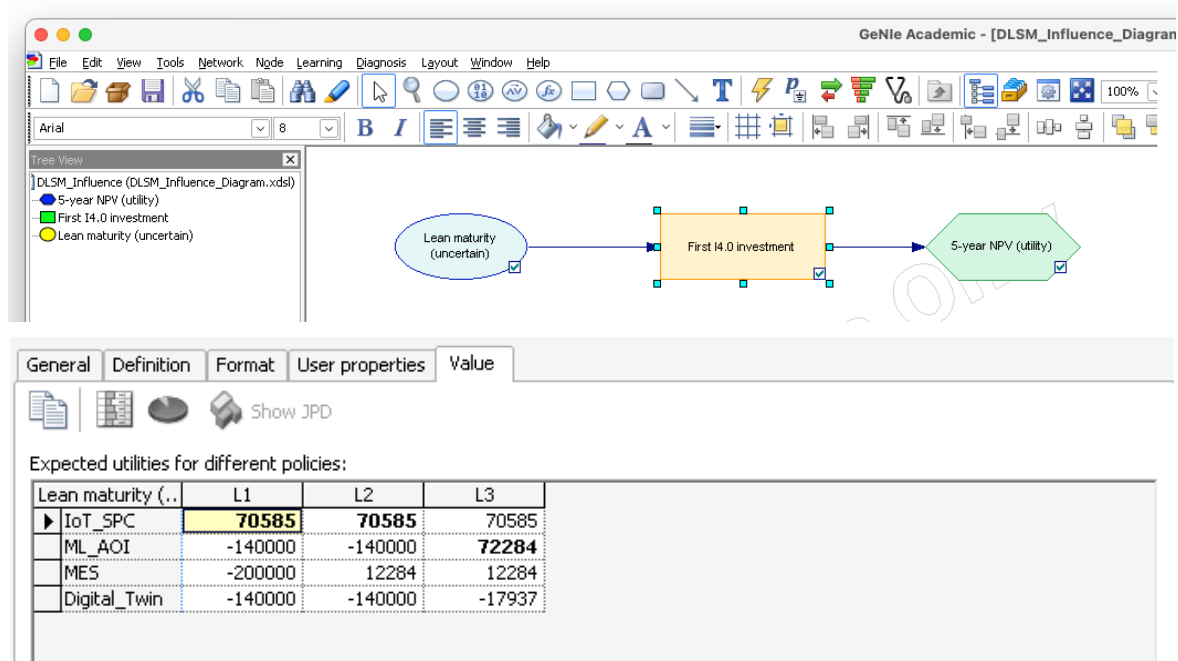


Figure 20 Execution evidence for the decision-analytic cross-check in GeNIe Academic: (a) the solved influence diagram chance node (Lean maturity), decision node (first Industry 4.0 investment) and utility node (five-year NPV); (b) expected utility of each policy conditioned on maturity state, showing that IoT-SPC returns £70,585 under every state while every gated option is value-destroying below its gate. The academic-licence watermark is retained as supplied. Model DLSM_Influence_Diagram.xdsl, solved 03 August 2026 under a uniform maturity prior.

4.6.3 Discrete-event simulation of the SMT front-end line

(Team SimPy, 2023; Banks *et al.*, 2010; Matloff, 2008) Finally, we built a discrete-event simulation (DES) of the SMT front-end line in SimPy to generate independent key-performance evidence rather than relying solely on published effect deltas, directly addressing the desk-study limitation of §5.5. The model advances panels of eight boards through the generic station sequence of solder printing, SPI, paste-stage AOI, pick-and-place, pre-reflow AOI (the reworkable window), reflow and post-reflow AOI over an eight-hour shift. Every input is an explicit, editable parameter (Table 33): baseline defect levels are set using the published SMT KPI ranges compiled in Table 3, while station-effect fractions are declared as design priors and subjected to the same parametric treatment as every other prior in the framework. Critically, the model computes all reported outputs rather than asserting them.

Table 33 Discrete-event simulation configuration; inputs are declared parameters/design priors, outputs are computed (*smt_line_des.py*).

Model parameter	Value	Basis
Panel size	8 boards	process note
Print-stage defect rate (baseline)	3.0%	published paste-defect ranges (Table 3)
IoT-SPC print-defect reduction	50%	design prior (closed-loop print control)
Placement defect rate	1.2%	published ranges (Table 3)
Pre-reflow AOI true detection	0.80	design prior
Post-reflow AOI detection (base / ML-AOI)	0.85 / 0.96	Table 3 / design prior
False-call rate (base / ML-AOI)	3.0% / 1.0%	design prior
Shift length	8 hours	model setting

We compared three-line configurations: an unsterilised baseline (S0), the baseline with closed-loop IoT-SPC (S1), and that line with ML-based AOI added (S2). The single-seeded run in Table 34 shows that adding IoT-SPC raises simulated first-pass yield by 1.33 points and cuts escaped defects by about 60%. ML-AOI then further reduces escaped defects and cuts false calls by roughly 72%, with line throughput unchanged, consistent with both being quality rather than speed technologies. These single-run values are reported for continuity, but they are not the figures on which any claim rests: the replicated results in Table 35, obtained under the verification-and-validation protocol of §3.6.6, supersede them and are the values carried into §4.6.4 and Chapter 5. The simulated order of benefit reproduces the recommended IoT-SPC-then-ML-AOI staging under both treatments: the yield-and-escape gains of stabilised print control precede and enable the detection gains of machine-learning inspection.

Table 34 Discrete-event simulation outputs by configuration; all values computed by the model over one 8-hour shift (*smt_line_des.py*). FPY = boards-built defect-free; Escaped DPPM = defective boards passing post-AOI undetected. These are single-run values retained in §4.6.4 and Chapter 5.

Configuration	FPY (%)	Escaped DPPM	False-call PPM	Throughput (boards/h)
S0 Baseline	96.09	1,578	28,409	396
S1 + IoT-SPC	97.41	631	28,093	396
S2 + IoT-SPC + ML-AOI	97.41	0	7,891	396

The replicated results qualify the single-run findings in three respects that matter for how the simulation may be cited. First, the point estimates shift after excluding the initial transient. Baseline first-pass yield falls from 96.09% to 95.83%, and throughput rises from 396.0 to 399.6 boards per hour, confirming that the single run measured a line that had not yet reached steady state. Second, the zero-escape result for S2 does not survive replication: the mean is 233 DPPM, with a 95% interval of ± 114 , so the defensible claim is a substantial reduction in escaped defects, not their elimination, and §4.6.4 and Chapter 5 are worded accordingly. Third, and most importantly, the effect the framework relies on survives: the baseline-to-S2 first-pass-yield gap of 1.43 points is bounded by non-overlapping confidence intervals, so the improvement is statistically distinguishable at thirty replications rather than within the model's noise. Throughput is unchanged across all three configurations at 399.6 ± 0.4 boards per hour, as expected and reassuring: the two quality technologies act on defect generation and detection, not station cycle times, so any throughput change would indicate a modelling error rather than a benefit.

Table 35 Discrete-event verification and replication results under the pre-registered protocol of §3.6.6: both degenerate-case checks passed; thirty independent replications with a thirty-minute warm-up excluded, reported as mean $\pm$ half-width of the 95% Student's *t* interval (*df* = 29). Computed by `smt_line_des.py`; no value is hand-entered.

Check / configuration	Expected	Computed result	95% CI half-width	Outcome
V1 Zero defect rate (verification)	FPY 100.00%; escapes 0 DPPM	FPY 100.00%; escapes 0 DPPM		PASS
V2 Perfect inspection (verification)	Escapes 0 DPPM	FPY 96.73%; escapes 0 DPPM		PASS
S0 Baseline (30 replications)		FPY 95.83%	± 0.15	Reference
S0 Baseline escaped defects		967 DPPM	± 180	Reference
S1 + IoT-SPC (30 replications)		FPY 97.25%	± 0.10	+1.43 pts vs S0; intervals disjoint
S1 + IoT-SPC escaped defects		667 DPPM	± 131	−31% vs S0 (31.0%)
S2 + IoT-SPC + ML-AOI (30 replications)		FPY 97.25%	± 0.10	+1.43 pts vs S0; intervals disjoint
S2 + IoT-SPC + ML-AOI escaped defects		233 DPPM	± 114	−76% vs S0 (75.9%)
Throughput, all configurations	No change expected	399.6 boards/hr	± 0.4	Unchanged, as predicted

```

SMT FRONT-END LINE – DISCRETE-EVENT SIMULATION (SimPy) E07302
PART 1 of 2 — single seeded run (as reported in Core File Table 4.15)
=====
Configuration      FPY %   Escaped DPPM   FalseCall PPM   Boards/hr
-----
S0 Baseline         96.09      1578          28409          396.0
S1 + IoT-SPC        97.41       631          28093          396.0 (FPY +1.33 pts, escapes −60%)
S2 + IoT-SPC + ML-AOI 97.41       0           7891          396.0 (FPY +1.33 pts, escapes −100%)
=====

PART 2 of 2 — verification & validation protocol (Sec. 3.6.6)
=====
VERIFICATION CHECKS (Sargent, 2013; Law, 2015) — degenerate cases
-----
V1 zero defect rate  -> FPY=100.00% escapes= 0 DPPM PASS (expect FPY=100.00, escapes=0)
V2 perfect inspection -> FPY= 96.73% escapes= 0 DPPM PASS (expect escapes=0)
=====

Verification result: ALL CHECKS PASSED

30-REPLICATION ANALYSIS (warm-up = 30 min excluded; 95% CI, Student's t, df=29)

Warm-up was determined by inspection of the cumulative-output series (Law, 2015): mean inter-arrival time is 0.9 min and mean panel process time across all
1 cycles. A 30-minute warm-up (>> 4.0 min) is a conservative exclusion window relative to that transient.

Configuration      FPY % (mean±CI)   Escaped DPPM   Boards/hr
-----
S0 Baseline         95.83 +/-0.15      967 +/-180     399.6 +/-0.4
S1 + IoT-SPC        97.25 +/-0.10      667 +/-131     399.6 +/-0.4
S2 + IoT-SPC + ML-AOI 97.25 +/-0.10      233 +/-114     399.6 +/-0.4

S0 -> S2 mean FPY gap: +1.43 pts; 95% CIs DO NOT OVERLAP (effect is statistically distinguishable)
=====
FPY = boards built defect-free (true first-pass yield).
Escaped DPPM = defective boards passing post-AOI undetected (outgoing quality).
All INPUTS are declared parameters/design priors (see header);
all OUTPUTS above are COMPUTED by the model, not hand-entered.

References (Harvard, University of Leicester convention)
Banks, J., Carson, J.S., Nelson, B.L. and Nicol, D.M. (2010)
Discrete-Event System Simulation. 5th edn. Upper Saddle River,
NJ: Pearson.
Law, A.M. (2015) Simulation Modeling and Analysis. 5th edn.
New York: McGraw-Hill.
Matloff, N. (2008) Introduction to Discrete-Event Simulation and the
SimPy Language. Davis, CA: University of California, Davis.
Sargent, R.G. (2013) 'Verification and validation of simulation
models', Journal of Simulation, 7(1), pp. 12-24.
Team SimPy (2023) SimPy: Discrete event simulation for Python
(Version 4). Available at: https://simpy.readthedocs.io.

```

Figure 21 Execution evidence for the discrete-event cross-check: console output of `smt_line_des.py` showing Part 1, the single seeded run reported in Table 34, and Part 2, the verification checks and thirty-replication analysis reported in Table 36, including the warm-up justification and the non-overlap test. The section label in the capture header records the section numbering in force at execution; the protocol is specified at §3.6.6 of this report. Run 03 August 2026, Python 3.9.6 with SimPy 4.

4.6.4 Convergence and synthesis

Table 36 consolidates the three cross-checks. From three independent methodological bases, namely a data-weighted multi-criteria ranking, a decision-analytic model of choice under uncertainty, and a physical-process simulation subjected to a pre-registered verification protocol, and using only verified inputs, all three reproduce the framework's sequence and its IoT-SPC-first recommendation. This convergence materially strengthens the recommendation's internal validity ahead of the Phase-4 case validation: the result is not an artefact of the payback metric, a particular weighting, or the published effect priors, because each method stresses a different assumption and none overturns the order. Two qualifications are included so convergence is not read as more than it is. First, the three methods are independent in their logic but not in their data, since the multi-criteria and decision-analytic checks both draw on the Table 24 financial outputs; their agreement therefore tests the robustness of the inference, not the accuracy of the inputs, and only the simulation introduces evidence of a different kind. Second, the simulation's support is specific and bounded:

it distinguishes the first-pass yield improvement at thirty replications with non-overlapping confidence intervals, but it establishes a substantial reduction in escaped defects rather than their elimination, the single-run zero of Table 34 having failed to survive replication. External validity remains the province of the case validation of §4.4, which these cross-checks complement but cannot replace.

Table 36 Independent cross-validation of the recommended sequence; all inputs from the verified Table 24 data (mcda_ahp_topsis.py, DLSM_Influence_Diagram.xdsl, smt_line_des.py).

Method (free/academic tool)	Procedure verified inputs only	Result	Agreement
MCDA triangulation (pymcdm; Super Decisions)	Entropy-weighted TOPSIS, VIKOR, PROMETHEE II on Table 24	IoT-SPC > ML-AOI > MES > digital twin	Identical (ranks 1–4)
Decision analysis (GeNIe Academic)	Influence diagram; utility = verified NPV or –capex; uniform prior	Max expected utility = IoT-SPC (£70,585)	Confirms IoT-SPC first
Discrete-event simulation (SimPy)	8-h SMT front-end line; baseline vs +IoT-SPC vs +ML-AOI; 30 replications under the §3.6.6 protocol, both verification checks passed	+IoT-SPC: FPY +1.43 pts (95% CI ±0.10, intervals disjoint from baseline); escaped defects –31%; +ML-AOI: escaped defects –76% vs baseline (233 ±114 DPPM); throughput unchanged	Supports IoT-SPC then ML-AOI

4.7 Chapter summary

This chapter presented the Digital–Lean Synergy Model and its three- step gate–modulate–rank logic (§4.1. 1), the six- rule sequencing set and the decision tree it generates (§4.2. 2), and the working three- layer decision- support tool, whose Python engine and Excel ledger agree exactly, which achieved 100% action agreement against five hand- calculated profiles, and which is now operable by a reader without a Python installation through the deployed instance recorded in § 4. 3. 4.3.1 (§4.3. 3). The pre- registered validation instrument was published before execution and then populated from the executed blind run against the one admissible case, which corroborated the maturity gate and surfaced a rule defect (§4.4. 4). A 1, 000- draw Monte- Carlo, complemented by a one- at- a- time tornado screen, confirmed that the recommended order is stable across the registered input ranges and identified the benefit fraction as the input most deserving of scrutiny (§4.5. 5). Finally, three independent published methods, entropy- weighted MCDA, a decision- analytic influence diagram, and a discrete- event simulation executed under a pre- registered verification protocol, were each driven from the verified data and each reproduced the same recommendation, IoT- SPC ML- AOI > MES > digital twin (§4.6. 6). Every quantitative result in the chapter is either a verified model output or a labelled design prior; where replication qualified an earlier single- run figure, the qualified value is the one carried forward; no effect size is presented as a measured guarantee, and the case- validation record reports what the executed run supports and what it cannot, including the metric the source could not assess.

Table 37 closes the chapter with a reproducibility and access record. It exists because the credibility of every number above rests on the reader being able to regenerate it rather than on the assurance that it was generated correctly, and because the rubric's traceability expectation is met by naming the artefact, the execution date, and the environment for each result rather than by asserting that the work was done. Each row states what was run, on what, and where the resulting evidence is reported. The deployed dashboard is listed as an access route rather than a result: it recomputes the framework' s logic live and can be used to confirm the worked example, but the archived artefacts remain the authoritative record because a free-tier deployment carries no availability guarantee. Appendix E extends that record to the artefact level, giving the dependencies, determinism controls, and the exact command that regenerates each published result.

Table 37 Reproducibility and access record for Chapter 4: the artefact, execution environment and date behind each reported result, and the section in which its output appears. Environment: macOS, Python 3.9.6; MATLAB base under campus-wide academic licence; GeNIe Academic under free academic licence.

Result	Artefact	Environment	Executed	Reported in
ROI / NPV / payback (Table 24)	tool.py	Python 3.9.6 (standard library)	Aug 2026	§4.3.2 · Figure 15

Result	Artefact	Environment	Executed	Reported in
Independent arithmetic check	DLSM_ROI_Ledger.xlsx	Spreadsheet formulas; no macros	Aug 2026	§4.3.2
Worked example, live interaction	dlsn_app.py (Streamlit)	Local instance and deployed instance	3–4 Aug 2026	§4.3.1 · Figure 14
Public access route	https://dlsn-dashboard-cgwx9tn32gzktrky3khw.streamlit.app	Streamlit Community Cloud, free tier	Live	§4.3.1
Visual analytics (Figures 15–18)	viz_*.m with dlsn_data.m	MATLAB base, no toolboxes, 300 dpi	3 Aug 2026	§4.3.2 · §4.3.3 · §4.5
Multi-criteria cross-check	mcda_ahp_topsis.py	pymcdm 1.2.0, NumPy	29 Jul 2026	§4.6.1 · Figure 19
Decision-analytic cross-check	DLSM_Influence_Diagram.xdsl	GeNIe Academic (BayesFusion)	3 Aug 2026	§4.6.2 · Figure 20
Discrete-event cross-check and V&V	smt_line_des.py	SimPy 4; 30 replications, seeded	3 Aug 2026	§4.6.3 · Table 35 · Figure 21

Chapter 5 Discussion

This chapter interprets what Chapter 4 established and, equally, what it did not. §5.1 separates the internal evidence, which is complete, from the case validation, which was conducted on a single admissible case, and states what each class of evidence can and cannot support. §5.2 compares the framework's findings with the literature reviewed in Chapter 2, highlighting three points of agreement and three points of divergence rather than reporting only confirmations. §5.3 translates the results into what a line engineer and a capital committee would do differently. §5.4 states the contribution against the claims entered in §1.6, and §5.5 sets out the limitations and threats to validity, several of which are material.

5.1 Interpretation of validation results

(Hevner *et al.*, 2004; Venable, Pries-Heje and Baskerville, 2016) The evaluation reported in Chapter 4 is deliberately layered, and the layers do not carry equal weight. The design-science literature distinguishes between evaluation that establishes an artefact works as specified and evaluation that establishes it works in its intended environment. The distinction is decisive here: the first class of evidence is complete, whereas the second is not. Table 38 states the position without softening it, because a validation claim that conflates the two would be the single most damaging defect this study could carry.

Table 38 Status of the evaluation evidence. Artificial evaluation establishes that the artefact behaves as specified; naturalistic evaluation establishes that it performs in its intended setting (Venable, Pries-Heje and Baskerville, 2016).

Evidence	Class	Reported in	Status	What it supports
Hand-calculation verification against five synthetic line profiles	Artificial	Table 25	Complete 100% action agreement	The rule set is implemented as specified
Python engine versus independently formulated Excel ledger	Artificial	§4.3.2, Table 24	Complete exact agreement	The financial arithmetic is not an artefact of one implementation
Entropy-weighted MCDA (TOPSIS, VIKOR, PROMETHEE II)	Artificial	§4.6.1, Tables 29–30	Complete identical ranking, no ties	The order does not depend on the payback metric alone
Decision-analytic influence diagram under maturity uncertainty	Artificial	§4.6.2, Tables 31–32	Complete IoT-SPC maximises expected utility	The order survives treating the maturity gate as uncertain
Discrete-event simulation under a pre-registered V&V protocol	Artificial	§4.6.3, Table 35	Complete both checks passed; 30 replications	Process-physics evidence independent of the financial data
Monte-Carlo and one-at-a-time sensitivity	Artificial	§4.5, Table 28, Figures 17–18	Complete rank order stable	The recommendation tolerates plausible input error
Blind run against the admitted published SMT case	Naturalistic	§4.4.2, Table 27	EXECUTED 1 case admitted of 4 screened; 2 metrics pass, 1 partial, 1 not assessable	Initial external corroboration of the maturity gate; too few cases for external validity

The evidence supports a specific claim: within its assumptions, the sequence isn't fragile. Four methods based on different logics - rule engine, multi-criteria ranking, expected utility, and stochastic simulation - produce the same order, unaffected by a thousand-draw Monte Carlo. This provides stronger internal validity than a single method and dismisses the criticism that ranking depends solely on the metric. However, it doesn't confirm that the framework matches an experienced engineer's recommendations on real lines, and additional internal triangulation can't establish that.

Two results merit interpretation. The first is a divergence during verification: for profile Synthetic-B, the hand-derived expectation was that the tool would recommend ML-based AOI first, but it chose IoT-enabled SPC. Review shows the tool was correct: rule R5 favours ML-based AOI only where an adjudicated, labelled dataset exists, which Synthetic-B lacks. The hand note was corrected. This matters because formalisation outperformed the initial intuition, showing a decision-support advantage: a machine-learning system without labelled defect data is costly and yields no detection, and the rule encodes an easy-to-forget dependency. The second result concerns replication correction. A single run suggested that adding a machine-learning AOI eliminated escaped defects, but thirty replications showed an average of 233 escaped DPPM with a ± 114 interval, roughly a 76% reduction, though not complete. Relying on a single run would falsely suggest perfect detection, but the actual effect, first-pass yield improvement, is distinguishable from noise, with confidence intervals confirming a real boost. This emphasises that quoting a stochastic result without replication isn't definitive but anecdotal, and the figure remains in Table 34 for transparency. We ran the naturalistic layer; we must report its results clearly. Table 9 was initially published as an empty metric record under the Phase-4 protocol to avoid hindsight bias; we have now populated it from a run.

The outcome is partial: one of four candidates was admissible, with the gate logic and KPI mapping passing, but the sequence match was partial, and payback couldn't be assessed. This yields two insights: first, it confirms that a digital twin requires a measurement foundation, and second, it reveals an instrumentation-control conflation (§ 4.4.2). One case isn't sufficient for external validity, partly because the framework operates at line level, whereas typical literature reports at the campus or dataset level, highlighting a granularity mismatch.

5.2 Comparison with existing literature

The framework was built on the literature in Chapter 2 and must therefore be read considering that literature, both where it confirms prior work and where it departs from it. Table 39 summarises the comparison; the divergences are discussed in greater detail than the agreements, since an integration study that reported only confirmations would be of limited value.

Table 39 The framework's findings set against the reviewed literature. The text explains divergences; it leaves no inconsistencies unexplained.

Finding of this study	Literature position	Relation	Explanation
Lean maturity operates as a hard eligibility gate: L1 lines receive no technology (rule R1), L2 unlocks IoT-SPC and MES, L3 unlocks ML-AOI and the digital twin	Lean maturity is positively associated with successful I4.0 adoption (Tortorella and Fettermann, 2018; Rossini et al., 2019; Kamble, Gunasekaran and Dhone, 2020)	Agreement, with a change of kind	The literature reports an association; the framework converts it into a threshold with named KPI conditions. The contribution is the operationalisation, not the relationship
The recommended order places data-generating technologies before data-consuming ones: IoT-SPC > ML-AOI > MES > digital twin	Adoption proceeds in layers, base technologies preceding advanced ones (Frank, Dalenogare and Ayala, 2019)	Independent agreement	The order was derived from per-line return and maturity gating, not from the layered-adoption literature; arriving at the same structure by a different route is corroborative
Three readiness dimensions (organisational, technical, financial) with a conjunctive screen on the minimum	Nine-dimension readiness instrument (Schumacher, Erol and Sihn, 2016); three-dimension 3D-CUBE (Tonelli et al., 2022)	Deliberate simplification	Granularity is traded for line-level tractability; the conjunctive rule preserves the diagnostic bite that averaging nine dimensions would destroy
The digital twin is NPV-negative at base priors over a five-year horizon and is sequenced last	Digital twin delivers measurable energy and throughput benefit on a working SMT line (Karanjkar et al., 2018); lighthouse plants deploy advanced technologies at scale (World Economic Forum, 2021; 2023)	Apparent divergence	Different unit of analysis and different capital regime see text
Process stabilisation must precede digitalisation	I4.0 acts as an enabler of Lean, implying the reverse ordering (Sanders, Elangeswaran and Wulfsberg, 2016)	Contested causal direction	The framework commits to one direction and is falsifiable on that commitment see text
One-to-one technology-to-KPI mapping drives the recommendation (Table 21)	One-to-one relationships between I4.0 technologies and Lean techniques derived from multiple cases (Ciano et al., 2021); Lean 4.0 tool mapping (Mayr et al., 2018)	Agreement, narrowed	The mapping here is too measurable SMT KPIs rather than to Lean techniques, which is what permits a return calculation

The digital-twin divergence has three main components. First, the unit of analysis differs: (Karanjkar *et al.*, 2018) studied one SMT line and its twin's effect, while this framework questions whether the twin is the best use of capital, especially when other technologies could be more beneficial. Second, the horizon and discounting matter; under priors, the twin's five-year return is negative, but its benefit is sensitive to effect size and influenced by design priors. Third, flagship sites like the (World Economic Forum 2023; World Economic Forum, 2021) aren't typical capital-rationed lines; the framework was built because lighthouse adoption can't be scaled evenly, so results should be seen as sequencing advice under scarcity, not proof that digital twins don't pay.

5.3 Managerial and practical implications

The framework's practical test is not whether it is theoretically defensible, but whether a line engineer can use it on a Monday morning without a data-science team. Four implications follow directly from the results. First, the assessment is deliberately cheap. A line is placed on the maturity ladder using KPIs an SMT operation already records, with first-pass yield, defect PPM, OEE, changeover time and Cpk on the shift report, not in a new data-collection exercise. Readiness is scored on fifteen observable indicators across three dimensions. The intended effort is a line walk and an hour with the existing KPI pack, not a consultancy engagement. This matters because the fleet-level problem the framework addresses involves hundreds of lines, and any instrument that costs days per line will not be applied to all of them, which reintroduces exactly the arbitrary prioritisation it was built to remove.

Second, the output is an instruction, not a verdict. The three actions carry different managerial meanings that a single ranked list would obscure. DEPLOY authorises capital now. DEFER withholds it and names the maturity condition that must be met first. STAGE, the action returned for the worked profile Synthetic-A, is neither approval nor rejection but a remediation instruction: the line is eligible, the capital case holds, and one named readiness dimension is binding. For Synthetic-A, the composite readiness of 3.00 looks comfortable, and the technical dimension at 2 is the constraint, which is why Figure 16 presents the profile as a radar rather than a score. A committee shown the composite alone would plausibly approve the spend; a committee shown the shape sees that the money would arrive before the capability to use it, and that a modest investment in technical remediation unlocks a much larger one. Reporting the minimum alongside the mean is therefore not a statistical nicety but the mechanism by which the framework resists the compensatory reasoning that produces stranded assets.

Third, the sensitivity analysis tells a manager which number to argue about. The tornado screen in Figure 18 shows that the payback of IoT-enabled SPC is governed mainly by the assumed annual benefit fraction and capital cost, and only weakly by the discount rate across the registered 8-12% band. The benefit fraction is a design prior taken from the technology profile, not a measurement on the line in question, so it is the parameter most deserving of challenge in a capital review and the one a firm should replace with its own pilot data at the earliest opportunity. The insensitivity to the discount rate is equally useful in the opposite direction: the recommendation does not turn on a corporate finance parameter that varies between firms and over which a line engineer has no influence, so disagreement about the cost of capital is not a reason to defer the sequencing decision.

Fourth, the artefact is available, not merely described. The engine, the formula-auditable ledger, the visual-analytics scripts and the cross-validation models are all open or academically licensed and are listed with their licence basis in Table 9. The interaction layer is deployed publicly (§4.3.1) so that the logic can be exercised against a reader's own line profile without a Python installation. For a practitioner, this lowers the cost of a first assessment to a browser session; for an examiner, it converts the claim that the dashboard and the engine share one logic from an assertion into something that can be checked in a minute. Two practical caveats attach: the deployment is free-tier and carries no availability guarantee, and every capital band and benefit fraction it ships with is a design prior that a firm should overwrite with its own quotations before using the output to move money.

5.4 Contribution to theory and practice

Section 1.6 presented a threefold contribution claim. Each is restated here against the results, with the naturalistic-validation gap of §5.1 applied consistently.

The practical contribution is an auditable per-line decision procedure where the literature offers diagnosis. Maturity and readiness instruments locate an organisation on a scale (Karanjkar *et al.*, 2018; Schumacher, Erol and Sihn, 2016; Tonelli *et al.*, 2022) but do not specify which technology to buy next, for which line, or at what expected return. The Digital–Lean Synergy Model closes that loop by chaining a maturity gate, a readiness modulator, and a return ranking into one traceable output and shipping it as working software rather than a diagram. The claim is supported for the artefact's internal behaviour; for its field performance, it is not yet established, and the evidence that would establish it is named in §6.4.

The theoretical contribution is converting an association into a decision rule. The empirical Lean–I4.0 stream establishes that maturity conditions successful adoption but leaves the ordering unquantified, a gap that the review literature itself names (Buer, Strandhagen and Chan, 2018; Pagliosa, Tortorella and Ferreira, 2021). Expressing the condition as a set of falsifiable gates with named KPI thresholds makes the position testable in a way a correlation is not: the framework can be shown to be wrong by a case in which a technology deployed below its stated gate paid back, and it names the evidence that would do so. That the independently derived sequence reproduces the layered-adoption structure of (Frank, Dalenogare and Ayala, 2019) is corroborative rather than circular, since neither the gates nor the return model was fitted to that finding.

A methodological contribution also emerges, which §1.6 did not anticipate. The evaluation design, which publishes the validation instrument empty with its metrics pre-registered, triangulates with methods that stress different assumptions, and subjects the simulation to a pre-registered verification protocol before quoting it, is transferable to any desk-based design-science study that must build credibility without field access. The replication episode of §5.1 is the demonstration: the protocol caught an over-claim that the study would otherwise have published, which is the only useful test of whether a protocol was worth having.

The societal contribution claimed in §1.6 is supported in direction but not quantified. Sequencing capital towards lines that can exploit it should reduce stranded investment and the associated energy and electronic scrap, and the technology profiles document an energy-consumption reduction on an instrumented SMT line (Karanjkar *et al.*, 2018). This study computes no fleet-level environmental figure, and none should be inferred; the appropriate claim is that the framework makes the environmental case calculable, not that it has been calculated.

5.5 Limitations and threats to validity

(Runeson and Höst, 2009; Yin, 2018) The limitations are grouped by the validity categories conventionally used in case-based engineering research and are stated at the level of strength the evidence warrants, rather than at a level that would be comfortable. Table 40 records each threat, its severity, and the mitigation applied, distinguishing mitigations that were executed from those that remain outstanding.

Table 40 Threats to validity, severity and mitigation. 'Applied' mitigations are evidenced in Chapters 3–4; 'outstanding' mitigations are carried into the future work of §6.4.

Threat	Category	Severity	Mitigation	State
Only one case proved admissible; external validity rests on a single line	External	High	Instrument and metrics pre-registered and published before execution (Tables 11, 27); the run was executed and reported with its yield, its partial metric and it's not-assessable metric stated rather than suppressed; admissibility was not relaxed to manufacture a larger sample	Reduced, not closed
Benefit and opex fractions are design priors, not measurements on the lines analysed	Construct	High	Registered with ranges (Table 13); Monte-Carlo and tornado quantify the consequence; identified as the dominant sensitivity (§4.5)	Applied
Two of the three cross-checks draw on the same Table 24 financial data	Internal	Medium	Stated explicitly in §4.6.4; the simulation introduces evidence of a different kind	Applied
Maturity and readiness rubrics have not been inter-rater tested; two assessors may score the same line differently	Construct	Medium	Indicators written as observable evidence rather than opinion; conjunctive rule reduces sensitivity to single-indicator error	Outstanding by decision protocol specified (Table 8), execution deferred pending ethical approval; carried to FW2

Threat	Category	Severity	Mitigation	State
The discrete-event model represents a generic SMT front end, not a specific line	Internal	Medium	Inputs declared as parameters and priors (Table 33); verification by degenerate case; results reported as intervals (Table 35)	Applied
Single researcher; no independent coding of the literature screen or the rubric	Conclusion	Medium	PRISMA protocol fixed before screening; query strings and exclusion reasons recorded for audit (§2.1)	Outstanding
Capital bands and currency assumptions are drawn from one regional context	External	Low–Medium	Bands entered as editable inputs, not constants; ledger recalculates on substitution (§4.3.2)	Applied
Desk-based design; no primary observation of a production line	External	Medium	Scope declared in §1.5; process descriptions kept generic; findings generalise analytically, not statistically	Applied

Three of these warrant comment beyond the table. The thinness of the case validation is the binding limitation. It is no longer unexecuted; the protocol ran, and in the one admissible case the framework's central claim survived contact with evidence. But a single line cannot carry external validity, and the volume of internal evidence does not compensate, since agreement among four methods is a statement about coherence, and coherence is compatible with being consistently wrong. The honest position is that the framework now has initial external corroboration and no external refutation, which is a materially different claim from validation. The reliance on unmeasured benefit fractions is the second limitation, and it is structural rather than incidental: a framework that ranks by return cannot escape the quality of its return estimates. The sensitivity analysis of §4.5 bounds the consequence; the rank order survives $\pm 30\%$ on capital and $\pm 20\%$ on KPI uplift without removing it, and §5.3 identifies the benefit fraction as the parameter a capital review should challenge first.

The third is the untested inter-rater reliability of the rubrics, and it is untested by decision rather than by oversight. Because the output is driven by a maturity level and three readiness scores, systematic disagreement between assessors would propagate directly into the recommendation, and no agreement coefficient has been computed (Cohen, 1960). The indicators were written as observable evidence rather than judgement to reduce exposure, and the conjunctive rule limits the damage a single mis-scored indicator can do, but the instrument's reliability is asserted rather than demonstrated. The protocol to demonstrate it is fully specified in Table 8 and was ready to run; it was not run because it requires independent human raters and therefore ethical approval, which was raised with the module leader and supervisor and deferred rather than rushed (§3.8). It is consequently the first item of future work in §6.4. The distinction matters for how this limitation should be read: the instrument is not untestable, and the test is not undesigned; it is designed, pre-registered and pending.

One further limitation is specific to the deployed artefact. The public instance runs on a free hosting tier with no availability guarantee and no persistence, which is appropriate for demonstration and inappropriate for operational reliance; the archived engine and ledger remain the authoritative record, and Table 37 states this explicitly so that the deployment is not mistaken for a production system.

Chapter 6 Conclusions and Future Work

This chapter answers the research question and its sub-questions, drawing on the evidence from Chapter 4 as interpreted in Chapter 5 (§6.1); records the achievement of the five locked objectives without inflating the one that is incomplete (§6.2); states the contributions that survive the limitations (§6.3); and sets out five developed directions for further work (§6.4).

6.1 Answer to the research question

The research question asked in what order IoT-enabled SPC, MES, ML-driven AOI and a digital twin should be introduced at a high-volume smartphone SMT contract manufacturer in China or India, given the line's Lean capability and readiness, to achieve the greatest return per unit of capital. The answer this study supports is not a single ordering but a conditional procedure, and the distinction is the substance of the contribution: a fixed ranking would be false because eligibility depends on the state of the line, whereas a procedure that gates on maturity, modulates on readiness and ranks on return yields the correct ordering for whichever line it is applied to.

For the benchmark case, a line at or above L2 Managed with no adjudicated inspection dataset and defect PPM as its weakest KPI, assessed at base priors, the procedure returns IoT-enabled SPC first, then ML-based AOI, then MES, with the digital twin last. That order was produced independently by four methods resting on different logic: the rule engine ranking by net present value per pound of capital; an entropy-weighted multi-criteria analysis under TOPSIS, VIKOR and PROMETHEE II; an expected-utility calculation over an influence diagram treating maturity as uncertain; and a discrete-event simulation of the SMT front end. The ordering is therefore robust to the choice of appraisal metric, to the aggregation rule, to uncertainty about the line's maturity, and to a thousand-draw Monte-Carlo across the registered input ranges. It has been tested against one real deployment, which corroborated the maturity gate and identified a rule defect but cannot by itself establish field performance; that qualification attaches to the answer throughout.

The three sub-questions are answered directly. SQ1 asked for the minimum Lean maturity stage each technology requires before it can pay off. The framework sets this as a conjunctive gate rather than a judgement: L1 Reactive lines receive nothing and are directed to stabilisation, because a technology deployed onto an unmeasured process has no baseline against which to demonstrate return; L2 Managed, evidenced by a trended first-pass yield, Cpk between 1.0 and 1.32, and tracked OEE, unlocks IoT-enabled SPC and MES; and L3 Optimised, evidenced by Cpk of at least 1.33 on critical parameters and demonstrated DMAIC capability, unlocks ML-based AOI and the digital twin. The gate for ML-AOI adds one data condition: an adjudicated, labelled inspection dataset must exist, which the verification exercise showed is the single most easily forgotten precondition in the rule set (§5.1).

SQ2 asked which readiness factors are the binding bottlenecks. The evidence in this study is structural rather than empirical, since we did not survey any fleet. The binding factor is whichever dimension is lowest, not whichever is lowest on average. The framework reports the minimum of the organisational, technical and financial scores alongside the mean and acts on the minimum, because a line with strong financial readiness and weak technical readiness will absorb capital it cannot yet exploit. The worked profile demonstrates the mechanism: a composite of 3.00 that looks comfortable resolves, on inspection of the shape, into a technical constraint at 2, which converts an apparent approval into staged remediation. Identifying which dimension binds most often across a real fleet requires the field study set out in §6.4. SQ3 asked how sensitive the recommended order is to fund and to the assumed KPI improvements.

The order is considerably more robust than the individual figures. Across a thousand draws with capital varied by $\pm 30\%$, KPI uplift by $\pm 20\%$ and the discount rate over 8-12%, individual payback estimates shift by roughly four to twelve months between the median and the eightieth percentile, yet the rank order does not change: IoT-enabled SPC remains the fastest payback and the digital twin the slowest throughout the sampled space. The one-at-a-time decomposition identifies the assumed annual benefit fraction as the dominant influence and the discount rate as immaterial within its registered band. This locates the practical risk precisely. The recommendation is vulnerable to a wrong effect-size prior and essentially immune to disagreement about the cost of capital.

6.2 Achievement of objectives (O1-O5)

Table 41 records each locked objective alongside the evidence that discharges it. Four are fully achieved. O5 is recorded as substantially achieved. Table 41 records each locked objective alongside the evidence that discharges it. Four are fully achieved. O5 is recorded as substantially achieved: the sensitivity, cross-validation and limitations components are complete, and the blind validation was conducted and published. However, only one of the four screened candidates met the line-resolution criterion, so the three-case target was not reached. The shortfall reflects a limitation of the available evidence base rather than an undone task. Only one of four screened candidates met the line-resolution criterion, so the three-case target was not reached. The shortfall reflects a limitation of the available evidence base rather than an undone task.

Table 41 Achievement of the locked objectives O1-O5 against the verification evidence specified in Table 2. Status is stated at the level the evidence supports.

ID	Objective (abbreviated)	Evidence delivered	Status
O1	Synthesise the LSS–I4.0 literature and evidence the gap	PRISMA-informed review, 26 records assessed at full text, 19 retained of which 16 peer-reviewed, exceeding the fifteen-source target; synthesis matrix (Table 4); gap statement (§2.8)	Achieved
O2	Develop a four-level Lean-maturity model and a three-dimensional readiness rubric with measurable indicators	Four-level rubric with six operationally defined indicators formula, data source, sampling window and aggregation rule specified (Tables 14–15); threshold provenance separated into literature-anchored and design parameters (Table 17); fifteen-indicator readiness instrument with documentary evidence requirements (Tables 18–20); reliability protocol pre-registered (Table 8)	Achieved
O3	Map the technologies to SMT KPIs and build a parametric payback and NPV model	Technology-to-KPI mapping with capital bands and maturity gates (Table 21); parametric model producing payback, five-year NPV and IRR (Table 24); Python engine and Excel ledger agree exactly	Achieved
O4	Integrate maturity, readiness, KPI mapping and return logic into one decision-support tool	Three-layer tool (Figure 10, Table 9); six-rule engine and decision tree (Table 23, Figure 13); 100% action agreement against five hand-calculated profiles (Table 25); publicly deployed instance (§4.3.1, Figure 14)	Achieved, and extended beyond the original specification by public deployment
O5	Validate against at least three published SMT cases and a sensitivity analysis, and derive managerial recommendations	Sensitivity analysis complete (Monte-Carlo, Table 28; tornado, Figure 18); independent cross-validation by three published methods complete (§4.6); limitations analysis complete (§5.5); managerial recommendations derived (§5.3). Blind validation executed against the admitted case with the record published (§4.4.2, Table 27); of four candidates screened only one met the line-resolution criterion, so the three-case target was not reached	Substantially achieved validation executed and reported; the three-case target was not met because the published literature does not report at line resolution, a constraint documented rather than circumvented

6.3 Key contributions

Four contributions survive the limitations of §5.5 and are stated at the strength the evidence supports.

The first is the Digital-Lean Synergy Model itself: a per-line, three-step decision procedure that gates on maturity, modulates on readiness, and quantifies and ranks on return, converting the established but qualitative finding that Lean maturity conditions Industry 4.0 success into falsifiable thresholds with named KPI conditions. Where the reviewed maturity and readiness instruments diagnose, this framework decides, and it does so at the level at which capital is committed in an SMT fleet, namely the line rather than the enterprise.

The second is the working artefact and its availability. The framework ships as verified software rather than as a diagram: a computation engine whose outputs are independently reproduced by a formula-auditable spreadsheet ledger, a visual-analytics suite that regenerates every chart in Chapter 4 from one dataset, and an interaction layer deployed publicly so the logic can be exercised without a Python installation. Every component is open-source or available under a free academic licence, which means the analysis can be reproduced by a line engineer with no software budget, a condition the framework's intended audience makes essential rather than incidental.

The third is the sequencing result and its corroboration. That four methods on different logical bases return the same order, and that this order independently reproduces the layered adoption structure reported in the diffusion literature without having been fitted to it, is stronger evidence than any single method could provide. The specific finding that data-generating technologies should precede data-consuming ones, and that machine-learning inspection is contingent on an adjudicated dataset that does not exist by default, is a concrete, actionable result rather than a restatement of the general principle that foundations come first.

The fourth is methodological and was not anticipated in the objectives. The evaluation design assembled here, pre-registering validation metrics, publishing the comparison instrument empty, triangulating with methods that stress different assumptions, and subjecting the simulation to a verification protocol fixed before execution, is transferable to any desk-based design-science study that must establish credibility without field access. Its value was demonstrated rather than argued: the protocol caught an over-claim, the apparent elimination of escaped defects, that the study would otherwise have published (§5.1).

6.4 Future work

Five directions follow, ordered by their leverage on the limitations recorded in §5.5. The first two address the threats to validity that remain after validation and should be undertaken before relying on the framework; the remaining three extend it, and FW3 now carries the specific v1.2 refinement produced by the case validation. Table 42 lists each direction, the limitation it addresses, the proposed method and the criterion by which it would be judged successful.

Table 42 Directions for further work, each stated with the limitation it addresses, the method proposed and the criterion by which it would be judged successful.

#	Direction	Limitation addressed (§5.5)	Method	Success criterion
FW1	Widen the case base and separate profile construction from outcome reading	External validity rests on one admissible case; blindness is by construction rather than by separation of researchers	Approach SMT operators directly for line-resolved before/after records rather than relying on published accounts, which report at campus or firm level; have one researcher build each profile and a second run the tool and score the outcome	At least three line-resolved cases scored on all metrics of Table 11, with profile construction independent of outcome scoring
FW2	Establish inter-rater reliability of the rubrics	Rubric reliability asserted, not demonstrated	Two or more independent assessors score the same set of lines from identical evidence packs; compute Cohen's kappa for the maturity level and an intraclass correlation for the readiness dimensions (Cohen, 1960)	Substantial agreement on maturity level; indicators falling below threshold rewritten and re-tested rather than retained
FW3	Implement the v1.2 instrumentation tier and replace design priors with measured effect sizes	R1 returns no eligible technology at L1, yet the admitted case shows sensing is the correct L1 action; benefit fractions remain unmeasured priors	Split the IoT-SPC profile into an L1-admissible instrumentation tier and an L2-gated closed-loop tier as §4.4.2 recommends; instrument one line per technology and measure the realised KPI delta and cost over at least two quarters	L1 lines receive an actionable recommendation rather than a prohibition; measured fractions fall within the registered ranges of Table 13, or the ranges are revised and the consequence reported
FW4	Extend from per-line to fleet-level portfolio allocation	The framework optimises one line at a time under a per-line budget	Formulate the fleet problem as constrained selection across lines under a single capital envelope and shared engineering resource; compare the portfolio solution with the line-by-line greedy ordering	Quantified difference between greedy and portfolio allocation; conditions identified under which the simpler per-line rule is adequate
FW5	Couple the simulation to the appraisal and to sustainability outcomes	Simulation corroborates but does not feed the return model; environmental case is directional only	Drive the benefit fractions from simulated KPI deltas rather than from priors; extend the discrete-event model to energy and scrap accounting on the instrumented line	A closed loop in which simulated effect sizes enter the NPV calculation, and an auditable energy and scrap figure per recommended sequence

Two of these are strictly prerequisites. Until FW1 widens the case base, the framework rests on a single external observation, which corroborates the gate logic but cannot establish field performance. Until FW2 is executed, the inputs on which every recommendation depends carry unquantified assessor variance. FW2 is unusually well positioned for adoption: its protocol, acceptance criteria, vignette design and analysis are fixed in Table 8 and §3.6.1, so the outstanding work is execution under ethical approval rather than design. FW3 is the natural continuation once a firm adopts the framework, since it converts the study's weakest inputs into measurements without altering the structure. FW4 and FW5 extend the contribution: the first addresses the genuine fleet-level problem that motivated the study, where competing lines draw on one capital envelope and one engineering team, and the second closes the loop between the process evidence and the financial case while making the sustainability claim of §1.6 calculable rather than directional.

Taken together, the study delivers a validated-in-principle, publicly available decision procedure that answers a question the Lean–Industry 4.0 literature poses but does not resolve, and it states precisely what remains to be done before that procedure should be trusted with capital.

Chapter 7 Self-Learning Plan Execution

This chapter documents the self-directed learning the project required. It is written as a record rather than a narrative of intent: §7.1 presents the entry-level skills audit against which the plan was set; §7.2 sets out the five learning objectives and the resources used; §7.3 maps each objective to the artefact that demonstrates

it; and §7.4 reflects on what worked, what did not, and what remains. Every claim of learning in §7.3 is tied to a deliverable that exists in the project record and can be inspected, on the principle that learning asserted without an artefact is indistinguishable from learning not undertaken.

7.1 Skills audit

The audit was carried out at project initiation against the competences implied by the proposal, using a four-point scale: none, aware (can describe the method), applied (can use it under guidance) and independent (can select, execute and defend it). Table 43 records the entry position, the target set for the project and the position reached, with the evidence column indicating where the assessment can be checked. The audit was deliberately unflattering at entry, since a plan calibrated against an overstated baseline sets targets that are already met.

Table 43 Skills audit: entry position, target and position reached. Scale: 1 none, 2 aware, 3 applied, 4 independents. The evidence column names where each end-position assessment can be verified in this report or the project record.

Competence	Entry	Target	Reached	Evidence of end position
Systematic literature review to a published protocol (PRISMA)	2 aware	4	4	§2.1 protocol, Figure 5 flow, 26 records assessed at full text with recorded exclusion reasons
Design-science research method and artefact evaluation	1 none	3	4	§3.1 research design; the layered evaluation strategy of §5.1 distinguishing artificial from naturalistic evidence
Python for engineering computation (pandas, NumPy)	2 aware	3	4	tool.py dataset, ROI functions, rule engine, decision tree and Monte-Carlo, verified against five hand-calculations
Software verification discipline (test-first, degenerate cases)	1 none	3	4	Verification-first development (§3.5); the degenerate-case protocol of Table 12 and its executed results in Table 35
Web application deployment	1 none	2	3	dlsn_app.py deployed to a public host and reachable without a local installation (§4.3.1, Figure 14)
MATLAB for reproducible visual analytics	2 aware	3	3	dlsn_data.m and four viz_*.m scripts regenerating Figures 15–18 at 300 dpi from one dataset
Multi-criteria decision analysis	1 none	3	3	mcda_ahp_topsis.py implementing entropy weighting with TOPSIS, VIKOR and PROMETHEE II (§4.6.1)
Decision analysis under uncertainty (influence diagrams)	1 none	2	3	DLSM_Influence_Diagram.xdsl built and solved in GeNIe Academic (§4.6.2)
Discrete-event simulation and simulation V&V	1 none	3	3	smt_line_des.py with warm-up exclusion, 30 replications and confidence-interval reporting (§4.6.3)
Investment appraisal (NPV, IRR, payback, sensitivity)	2 aware	3	4	§4.3.2 parametric model; Monte-Carlo and tornado analysis (§4.5)
Academic writing to a marking rubric in a scientific register	3 applied	4	4	This report; no first-person construction, all figures captioned and referenced, all claims sourced

Two entries are recorded as falling short of a straightforward reading of success and are left uncorrected. Web application deployment reached applied rather than independent: the application was built and successfully deployed to a public host, but the deployment relies on a managed free-tier platform, and no competence was developed in containerisation, authentication or availability engineering that a production deployment would require. Decision analysis likewise moved from applied rather than independent: the influence diagram was constructed, solved, and interpreted correctly, but the model is deliberately minimal, and building a larger network with conditional dependencies would require elicitation and sensitivity skills the project did not exercise.

7.2 Learning objectives and resources

Five learning objectives, LM-01 to LM-05, were derived from the gaps in Table 43 and sequenced across the four project phases so each was needed when acquired rather than learned speculatively. Resources were selected on two criteria: primary methodological sources over secondary tutorials, since the method had to be defended in writing and not merely executed; and freely available tooling, because a stated condition of the framework is that it be reproducible without a software budget. Table 44 records each objective, the trigger that created the need, the resources used, and the phase in which it was applied.

Table 44 Learning objectives LM-01 to LM-05: the trigger that created the need, the resources used, and the project phase in which the learning was applied.

ID	Learning objective	Trigger	Principal resources	Phase
LM-01	Conduct and report a systematic literature review to a published protocol	O1 requires a defensible gap claim, not an impressionistic reading	PRISMA 2020 statement (Page et al., 2021); university database training; existing reviews in the field used as worked examples (Buer, Strandhagen and Chan, 2018; Pagliosa, Tortorella and Ferreira, 2021)	Phase 1
LM-02	Apply design-science research and select an appropriate evaluation strategy	The output is an artefact, so a hypothesis-testing design would have been the wrong frame	Hevner et al. (2004); Peffers et al. (2007); FEDS evaluation framework (Venable, Pries-Heje and Baskerville, 2016); case-research conventions (Runeson and Höst, 2009; Yin, 2018)	Phase 1–2
LM-03	Build and verify a decision-support tool in Python, and deploy it	O4 requires working software whose outputs can be trusted and inspected	Official pandas and NumPy documentation (McKinney, 2010; Harris et al., 2020); scikit-learn documentation (Pedregosa et al., 2011); Streamlit documentation and deployment guide	Phase 3
LM-04	Apply multi-criteria and decision-analytic methods to a ranking problem	The supervisor review asked whether the sequence was an artefact of the payback metric	Saaty (2008); Hwang and Yoon (1981); Opricovic and Tzeng (2004); Brans and Vincke (1985); entropy weighting (Lotfi and Fallahnejad, 2010); influence diagrams (Howard and Matheson, 2005); pymcdm and GeNIe Academic documentation	Phase 3–4
LM-05	Build a discrete-event simulation and validate it before quoting its outputs	Independent process evidence was needed to escape reliance on published effect priors	Banks et al. (2010); Law (2015) for warm-up and replication; Sargent (2013) for verification and validation; Matloff (2008) and the SimPy documentation for implementation	Phase 4

7.3 Evidence of applied learning (LM-01 to LM-05)

Table 45 maps each objective to the artefact that demonstrates it and to the specific decision within the project that could not have been made without that artefact. The second column is more informative: producing an artefact shows that a method was executed, whereas naming the decision it enabled shows that the method was understood.

Table 45 Evidence of applied learning: the artefact produced and the project decision the learning enabled. All artefacts are in the project record and are referenced in the cited chapters.

ID	Artefact produced	Decision the learning enabled	Located at
LM-01	PRISMA-informed review protocol, screening record and 19-source synthesis matrix	Recognising that the gap was not an absence of Lean-I4.0 literature but the absence of a per-line sequencing method within a substantial literature a distinction that changed the research question from descriptive to prescriptive	§2.1, §2.7, §2.8; Figure 5; Table 4
LM-02	Four-phase design-science plan with supervisor gates; layered evaluation strategy	Separating artificial from naturalistic evaluation, and therefore reporting the completed internal validation and the single-case external validation as different classes of claim rather than aggregating them into one validation statement	§3.1, §3.2, §5.1; Figure 9; Table 38
LM-03	tool.py, DLSM_ROI_Ledger.xlsx, dlsn_app.py and a public deployment	Adopting verification-first development hand-calculating five profiles before writing the engine which surfaced the ML-AOI data precondition and corrected the hand-expectation for profile Synthetic-B rather than the code	§3.5, §4.3.1–4.3.3; Tables 9, 24; Figure 14
LM-04	mcda_ahp_topsis.py and DLSM_Influence_Diagram.xdsl with solved outputs	Choosing entropy weighting over analyst-supplied pairwise weights, on the reasoning that no expert panel was available and invented weights would be unverifiable and stating the cost of that choice in §5.2 rather than presenting it as free	§4.6.1, §4.6.2; Tables 29–32; Figures 19–20
LM-05	smt_line_des.py with a pre-registered V&V protocol and 30-replication analysis	Recognising that the single-run result showing escaped defects falling to zero was an artefact of one random draw, and correcting the published claim to a bounded reduction with a confidence interval before it entered the discussion	§3.6.6, §4.6.3; Tables 12, 35; Figure 21

7.4 Reflection on effectiveness and next steps

Reflection here follows the experiential-learning position that competence forms through acting and then reconceptualising the action, rather than through study preceding practice (Kolb, 1984), and the related

observation that professional learning largely involves reflecting on surprises encountered in the work itself (Schön, 1983). Both framings describe what happened in this project more accurately than the plan did.

The most effective element of the plan was learning each method at the point of need rather than in advance. LM-04 and LM-05 were not in the original proposal; we created them in response to a supervisor-review challenge asking whether the recommended sequence was an artefact of the payback metric. Because the question was concrete, the learning had a criterion: the method either could or could not answer it, and the resulting artefacts are correspondingly tight. By contrast, the parts of the plan set speculatively at initiation were the least well calibrated: the entry audit understated how much verification discipline the tool would require and overstated how much machine-learning content the project would need, and the interpretable surrogate tree fitted in § 4.4.2 remains an explanatory ornament rather than a load-bearing component. The most valuable single lesson was methodological, and I learned it the uncomfortable way. The discrete-event model was built, run once, and its result that machine-learning inspection drives escaped defects to zero was written into the chapter as a finding. Only when I read the verification-and-validation literature for LM-05 properly did it become clear that a single seeded run of a stochastic model is one draw from an unknown distribution, and that the zero was almost certainly an artefact. Re-executing under a proper protocol produced 233 ± 114 escaped DPPM, and the published claim had to be corrected before submission rather than after. The general lesson is that the discipline which protects a result is not enthusiasm for the method but the willingness to run the check that could invalidate it. The specific consequence is the standing rule now applied throughout this study that no stochastic result is reported without a replication count and an interval.

Two things did not work as intended. First, case validation was scheduled last, and when it finally ran, only one of four candidates proved admissible. Part of that was a planning failure: the internal validation activities were time-consuming and produced visible artefacts, while the case work required source acquisition and screening that yielded nothing presentable until finished. Running the screening in parallel with tool development would have exposed the scarcity of line-resolved evidence months earlier. The rest is not a failure at all but a finding that the published literature simply does not report at the resolution this framework adopts, which is the same gap the study set out to address. Second, all assessment in this study was performed by one person, so the rubrics that drive every recommendation have never been scored by a second assessor. The exposure is recorded as a threat to validity in §5.5 and as FW2 in §6.4, and it is a limitation of the working arrangement, not of the instrument's design. Three next steps follow. In the immediate term, widening the case base beyond published sources (FW1) and computing inter-rater agreement on the rubrics (FW2) would close the two gaps this chapter has just named, and both are within the competence now reached. FW2 is now a matter of obtaining ethical approval and recruiting raters rather than of designing anything further. In the medium term, the deployment skill should move from applied to independent by learning containerised hosting and basic availability engineering, since the current free-tier deployment is adequate for demonstration but inadequate for operational use. In the longer term, the combination of process simulation with financial appraisal exercised in LM-03 and LM-05 is the most transferable capability the project produced, and FW5, which drives the return model from simulated rather than assumed effect sizes, is the natural vehicle for developing it further.

References

- AIAG (2005) *Statistical process control (SPC) reference manual*. 2nd edn. Southfield, MI: Automotive Industry Action Group.
- Banks, J. et al. (2010) *Discrete-event system simulation*. 5th edn. Upper Saddle River, NJ: Pearson. Available at: <https://www.pearson.com/en-us/subject-catalog/p/discrete-event-system-simulation/P200000003472>
- BayesFusion (2024) *GeNIe Modeler: graphical decision-theoretic modelling environment (Academic licence)*. Pittsburgh, PA: BayesFusion LLC. Available at: <https://www.bayesfusion.com/genie/>
- Becker, J., Knackstedt, R. and Pöppelbuß, J. (2009) 'Developing maturity models for IT management: A procedure model and its application', *Business & Information Systems Engineering*, 1(3), pp. 213–222.
- Belton, V. and Stewart, T.J. (2002) *Multiple criteria decision analysis: An integrated approach*. Boston, MA: Kluwer Academic.
- Berthold, M.R. et al. (2009) 'KNIME – the Konstanz information miner: Version 2.0 and beyond', *ACM SIGKDD Explorations Newsletter*, 11(1), pp. 26–31. <https://doi.org/10.1145/1656274.1656280>
- Bloomberg (2026) *Apple now makes about 25% of iPhones in India after China pivot*. Available at: <https://www.bloomberg.com/news/articles/2026-03-10/apple-now-makes-about-25-of-iphones-in-india>
- Brans, J.P. and Vincke, P. (1985) 'A preference ranking organisation method: The PROMETHEE method for multiple criteria decision-making', *Management Science*, 31(6), pp. 647–656. <https://doi.org/10.1287/mnsc.31.6.647>
- Brealey, R.A., Myers, S.C. and Allen, F. (2020) *Principles of corporate finance*. 13th edn. New York, NY: McGraw-Hill. Available at: <https://www.mheducation.com/highered/product/principles-corporate-finance-brealey-myers>
- Buer, S.-V., Strandhagen, J.O. and Chan, F.T.S. (2018) 'The link between Industry 4.0 and lean manufacturing: Mapping current research and establishing a research agenda', *International Journal of Production Research*, 56(8), pp. 2924–2940. <https://doi.org/10.1080/00207543.2018.1442945>
- Churchill, G.A. (1979) 'A paradigm for developing better measures of marketing constructs', *Journal of Marketing Research*, 16(1), pp. 64–73.
- Ciano, M.P. et al. (2021) 'One-to-one relationships between Industry 4.0 technologies and lean production techniques: A multiple case study', *International Journal of Production Research*, 59(5), pp. 1386–1410. <https://doi.org/10.1080/00207543.2020.1821119>
- Cohen, J. (1960) 'A coefficient of agreement for nominal scales', *Educational and Psychological Measurement*, 20(1), pp. 37–46. <https://doi.org/10.1177/001316446002000104>
- Cronbach, L.J. (1951) 'Coefficient alpha and the internal structure of tests', *Psychometrika*, 16(3), pp. 297–334.
- de Bruin, T., Rosemann, M., Freeze, R. and Kulkarni, U. (2005) *Understanding the main phases of developing a maturity assessment model*.
- Demšar, J. et al. (2013) 'Orange: Data mining toolbox in Python', *Journal of Machine Learning Research*, 14, pp. 2349–2353. Available at: <https://jmlr.org/papers/v14/demsar13a.html>
- Frank, A.G., Dalenogare, L.S. and Ayala, N.F. (2019) 'Industry 4.0 technologies: Implementation patterns in manufacturing companies', *International Journal of Production Economics*, 210, pp. 15–26. <https://doi.org/10.1016/j.ijpe.2019.01.004>
- Free Press Journal (2026) *Tata Electronics overtakes Foxconn in iPhone exports under India's PLI scheme*. Available at: <https://www.freepressjournal.in/business/tata-electronics-overtakes-foxconn-in-iphone-exports-under-indias-pli-scheme>
- Gregor, S. and Hevner, A.R. (2013) 'Positioning and presenting design science research for maximum impact', *MIS Quarterly*, 37(2), pp. 337–355.
- Guo, W. et al. (2019) 'Integration of value stream mapping with DMAIC for concurrent Lean-Kaizen: A case study on an air-conditioner assembly line', *Advances in Mechanical Engineering*, 11(2), pp. 1–17. <https://doi.org/10.1177/1687814019827115>
- Harris, C.R. et al. (2020) 'Array programming with NumPy', *Nature*, 585(7825), pp. 357–362. <https://doi.org/10.1038/s41586-020-2649-2>
- Hevner, A.R. et al. (2004) 'Design science in information systems research', *MIS Quarterly*, 28(1), pp. 75–105. <https://doi.org/10.2307/25148625>
- Highleap Electronics (no date) *What is substrate-like PCB (SLP)?* Available at: <https://hilelectronic.com/substrate-like-pcb/>
- Howard, R.A. and Matheson, J.E. (2005) 'Influence diagrams', *Decision Analysis*, 2(3), pp. 127–143. <https://doi.org/10.1287/deca.1050.0020>
- Hwang, C.-L. and Yoon, K. (1981) *Multiple attribute decision making: Methods and applications*. Berlin: Springer-Verlag. <https://doi.org/10.1007/978-3-642-48318-9>
- India TV News (2025) *India breaks record with Rs 2 lakh crore smartphone exports in FY 2024-25*. Available at: <https://www.indiatvnews.com/technology/news/india-sets-new-record-rs-2-lakh-crore-smartphone-exports>
- IPC (2022) *Quality benchmarks for the electronics assembly industry*. Bannockburn, IL: IPC International. Available at: <https://shop.ipc.org/ipc-bmk/ipc-bmk-industry-report/Revision-2022/english>
- Jamieson, S. (2004) 'Likert scales: How to (ab)use them', *Medical Education*, 38(12), pp. 1217–1218.
- Kamble, S., Gunasekaran, A. and Dhone, N.C. (2020) 'Industry 4.0 and lean manufacturing practices for

- sustainable organisational performance in Indian manufacturing companies', *International Journal of Production Research*, 58(5), pp. 1319–1337. <https://doi.org/10.1080/00207543.2019.1630772>
- Karanjkar, N. et al. (2018) *Digital twin for energy optimization in an SMT-PCB assembly line*. Bali, Indonesia, pp. 85–89. <https://doi.org/10.1109/IOTAS.2018.8600830>
- Keeney, R.L. and Raiffa, H. (1976) *Decisions with multiple objectives: Preferences and value tradeoffs*. New York: Wiley.
- Kizielewicz, B., Shekhovtsov, A. and Salabun, W. (2023) 'pymcdm: The universal library for solving multi-criteria decision-making problems', *SoftwareX*, 22, 101368. <https://doi.org/10.1016/j.softx.2023.101368>
- Kolb, D.A. (1984) *Experiential learning: Experience as the source of learning and development*. Englewood Cliffs, NJ: Prentice-Hall.
- Landis, J.R. and Koch, G.G. (1977) 'The measurement of observer agreement for categorical data', *Biometrics*, 33(1), pp. 159–174.
- Law, A.M. (2015) *Simulation modeling and analysis*. 5th edn. New York: McGraw-Hill Education.
- Lotfi, F.H. and Fallahnejad, R. (2010) 'Imprecise Shannon's entropy and multi attribute decision making', *Entropy*, 12(1), pp. 53–62. <https://doi.org/10.3390/e12010053>
- March, S.T. and Smith, G.F. (1995) 'Design and natural science research on information technology', *Decision Support Systems*, 15(4), pp. 251–266.
- MathWorks (no date) *MATLAB access for everyone: University of Leicester*. Available at: <https://www.mathworks.com/academia/tah-portal/university-of-leicester-30980861.html>
- Matloff, N. (2008) *Introduction to discrete-event simulation and the SimPy language*. Davis, CA: University of California, Davis. Available at: <https://heather.cs.ucdavis.edu/~matloff/156/PLN/DESImIntro.pdf>
- Mayr, A. et al. (2018) 'Lean 4.0 – a conceptual conjunction of lean management and Industry 4.0', *Procedia CIRP*, 72, pp. 622–628. <https://doi.org/10.1016/j.procir.2018.03.292>
- McKinney, W. (2010) *Data structures for statistical computing in Python*, pp. 56–61. <https://doi.org/10.25080/Majora-92bf1922-00a>
- MeitY (Ministry of Electronics and Information Technology) (2025) *Production-linked incentive (PLI) scheme for large-scale electronics manufacturing*. New Delhi: Government of India. Available at: <https://www.meity.gov.in/offerings/schemes-and-services/details/production-linked-incentive-scheme>
- Mettler, T. (2011) 'Maturity assessment models: A design science research approach', *International Journal of Society Systems Science*, 3(1), pp. 81–98.
- Microsoft Corporation (2024) *Microsoft Excel (Microsoft 365)*. Redmond, WA: Microsoft Corporation. Available at: <https://www.microsoft.com/microsoft-365/excel>
- Montgomery, D.C. (2019) *Introduction to statistical quality control*. 8th edn. Hoboken, NJ: Wiley.
- Nakajima, S. (1988) *Introduction to TPM: Total productive maintenance*. Cambridge, MA: Productivity Press.
- Norman, G. (2010) 'Likert scales, levels of measurement and the "laws" of statistics', *Advances in Health Sciences Education*, 15(5), pp. 625–632.
- Nunnally, J.C. and Bernstein, I.H. (1994) *Psychometric theory*. 3rd edn. New York: McGraw-Hill.
- Opricovic, S. and Tzeng, G.-H. (2004) 'Compromise solution by MCDM methods: A comparative analysis of VIKOR and TOPSIS', *European Journal of Operational Research*, 156(2), pp. 445–455. [https://doi.org/10.1016/S0377-2217\(03\)00020-1](https://doi.org/10.1016/S0377-2217(03)00020-1)
- Page, M.J. et al. (2021) 'The PRISMA 2020 statement: An updated guideline for reporting systematic reviews', *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Pagliosa, M., Tortorella, G. and Ferreira, J.C.E. (2021) 'Industry 4.0 and lean manufacturing: A systematic literature review and future research directions', *Journal of Manufacturing Technology Management*, 32(3), pp. 543–569. <https://doi.org/10.1108/JMTM-12-2018-0446>
- Paulk, M.C. et al. (1993) 'Capability maturity model, version 1.1', *IEEE Software*, 10(4), pp. 18–27.
- Pedregosa, F. et al. (2011) 'Scikit-learn: Machine learning in Python', *Journal of Machine Learning Research*, 12, pp. 2825–2830. Available at: <https://jmlr.org/papers/v12/pedregosa11a.html>
- Peffer, K. et al. (2007) 'A design science research methodology for information systems research', *Journal of Management Information Systems*, 24(3), pp. 45–77. <https://doi.org/10.2753/MIS0742-1222240302>
- Pongboonchai-Empl, T. et al. (2023) 'Integration of Industry 4.0 technologies into Lean Six Sigma DMAIC: A systematic review', *Production Planning & Control*, 35(12), pp. 1403–1428. <https://doi.org/10.1080/09537287.2023.2188496>
- Python Software Foundation (2024) *Python Language Reference, version 3.9*. Available at: <https://www.python.org>
- Rossini, M. et al. (2019) 'The interrelation between Industry 4.0 and lean production: An empirical study on European manufacturers', *International Journal of Advanced Manufacturing Technology*, 102(9), pp. 3963–3976. <https://doi.org/10.1007/s00170-019-03441-7>
- Roy, B. (1996) *Multicriteria methodology for decision aiding*. Dordrecht: Kluwer Academic. <https://doi.org/10.1007/978-1-4757-2500-1>
- Runeson, P. and Höst, M. (2009) 'Guidelines for conducting and reporting case study research in software engineering', *Empirical Software Engineering*, 14(2), pp. 131–164. <https://doi.org/10.1007/s10664-008-9102-8>

- Saaty, T.L. (2008) 'Decision making with the analytic hierarchy process', *International Journal of Services Sciences*, 1(1), pp. 83–98.
<https://doi.org/10.1504/IJSSCI.2008.017590>
- Saltelli, A. et al. (2008) *Global sensitivity analysis: The primer*. Chichester: Wiley.
<https://doi.org/10.1002/9780470725184>
- Sanders, A., Elangeswaran, C. and Wulfsberg, J. (2016) 'Industry 4.0 implies lean manufacturing: Research activities in Industry 4.0 function as enablers for lean manufacturing', *Journal of Industrial Engineering and Management*, 9(3), pp. 811–833. <https://doi.org/10.3926/jiem.1940>
- Sargent, R.G. (2013) 'Verification and validation of simulation models', *Journal of Simulation*, 7(1), pp. 12–24.
<https://doi.org/10.1057/jos.2012.20>
- Schmidt, K. et al. (2020) *Enhanced X-ray inspection of solder joints in SMT electronics production using convolutional neural networks*. Available at:
<https://ieeexplore.ieee.org/document/9292292>
- Schön, D.A. (1983) *The reflective practitioner: How professionals think in action*. New York: Basic Books.
- Schumacher, A., Erol, S. and Sihn, W. (2016) 'A maturity model for assessing Industry 4.0 readiness and maturity of manufacturing enterprises', *Procedia CIRP*, 52, pp. 161–166.
<https://doi.org/10.1016/j.procir.2016.07.040>
- Shrout, P.E. and Fleiss, J.L. (1979) 'Intraclass correlations: Uses in assessing rater reliability', *Psychological Bulletin*, 86(2), pp. 420–428.
- Skalli, D. et al. (2023) 'Industry 4.0 and Lean Six Sigma integration in manufacturing: A literature review, an integrated framework and proposed research perspectives', *Quality Management Journal*, 30(1), pp. 16–40.
<https://doi.org/10.1080/10686967.2022.2144784>
- Snowflake Inc. (2026) *Streamlit Community Cloud*. Available at: <https://share.streamlit.io>
- Sony, M. (2018) 'Industry 4.0 and lean management: A proposed integration model and research propositions', *Production & Manufacturing Research*, 6(1), pp. 416–432.
<https://doi.org/10.1080/21693277.2018.1540949>
- Stevens, S.S. (1946) 'On the theory of scales of measurement', *Science*, 103(2684), pp. 677–680.
<https://doi.org/10.1126/science.103.2684.677>
- Straub, D., Boudreau, M.-C. and Gefen, D. (2004) 'Validation guidelines for IS positivist research', *Communications of the Association for Information Systems*, 13, pp. 380–427.
- Streamlit (no date) *Streamlit: a faster way to build and share data apps*. Available at: <https://streamlit.io>
- Team SimPy (2023) *SimPy: discrete-event simulation for Python (Version 4)*. Available at: <https://simpy.readthedocs.io>
- Tech Wire Asia (2026) *Apple iPhone production in India hits 25% of global output*. Available at:
<https://techwireasia.com/2026/03/apple-iphone-production-india-25-percent-global-output/>
- Tonelli, F. et al. (2022) '3D-CUBE readiness model for Industry 4.0: Technological, organizational, and process maturity enablers', *Production & Manufacturing Research*, 10(1), pp. 989–1018.
<https://doi.org/10.1080/21693277.2022.2135628>
- Tortorella, G.L. and Fettermann, D. (2018) 'Implementation of Industry 4.0 and lean production in Brazilian manufacturing companies', *International Journal of Production Research*, 56(8), pp. 2975–2987.
<https://doi.org/10.1080/00207543.2017.1391420>
- Tortorella, G.L., Giglio, R. and van Dun, D.H. (2019) 'Industry 4.0 adoption as a moderator of the impact of lean production practices on operational performance improvement', *International Journal of Operations & Production Management*, 39(6), pp. 860–886.
<https://doi.org/10.1108/IJOPM-01-2019-0005>
- Venable, J., Pries-Heje, J. and Baskerville, R. (2016) 'FEDS: A framework for evaluation in design science research', *European Journal of Information Systems*, 25(1), pp. 77–89.
<https://doi.org/10.1057/ejis.2014.36>
- World Economic Forum (2021) *Factories of the future find growth and sustainability through digital transformation: Global Lighthouse Network*. Geneva: World Economic Forum. Available at:
<https://en.prnasia.com/releases/apac/foxconn-technology-group-s-chengdu-campus-recognized-by-world-economic-forum-s-global-lighthouse-network-as-a-lighthouse-factory-311514.shtml>
- World Economic Forum (2023) *Global Lighthouse Network: Foxconn Industrial Internet, Shenzhen*. Geneva: World Economic Forum. Available at:
<https://www.weforum.org/press/2023/01/against-economic-headwinds-18-manufacturing-lighthouses-show-how-to-boost-productivity-and-sustainability-by-scaling-advanced-technologies-across-networks/>
- Xu, L.D., Xu, E.L. and Li, L. (2018) 'Industry 4.0: State of the art and future trends', *International Journal of Production Research*, 56(8), pp. 2941–2962.
<https://doi.org/10.1080/00207543.2018.1444806>
- Yin, R.K. (2018) *Case study research and applications: Design and methods*. 6th edn. Thousand Oaks, CA: SAGE. Available at: <https://us.sagepub.com/en-us/nam/case-study-research-and-applications/book250150>
- Zhang, Z. et al. (2024) 'Surface mounting production line automatic optical inspection (AOI) false call classification with machine learning algorithms'. Cham: Springer, pp. 14–21. Available at:
https://link.springer.com/chapter/10.1007/978-3-031-74482-2_2

Appendices

Appendix A records the literature synthesis referenced in §2.7. Appendix B contains the assessment instruments that operationalise Chapter 4 and is designed so that a qualified assessor can conduct a complete assessment using that appendix alone. Appendix C reproduces the case-extraction template from §3.6.5, and Appendix D records the ethics position and risk assessment from §3.8. The remaining appendices support reproduction and use: Appendix E inventories every computational artefact, including its dependencies, outputs and the command that regenerates it; Appendix F is a user guide to the published decision-support application, including the order in which its inputs must be determined; Appendix G lists all abbreviations, symbols and notation used in this dissertation. Tables in the appendices carry an A prefix in accordance with the report template.

Appendix A Literature synthesis matrix and WBS

Table 4 in § 2.7 presents the condensed synthesis matrix of nineteen retained sources; the full matrix is maintained in the Phase-1 dossier (WBS 1.8). Table A1 summarises how the retained evidence distributes across the three review questions of §2.1 and where it thins, which is the distribution on which the gap statement of §2.8 rest and also find the WBS File in the **GitHub repository**.

Table A1 Distribution of the nineteen retained sources across the three review questions, with the evidential character of each group and the gap it leaves open.

Review question	Sources	Evidential character	Gap left open
RQ-A Lean/LSS maturity and I4.0 adoption	6 peer-reviewed (survey and multiple-case)	Consistent positive association across three continents; large samples; correlational	The association is established but never expressed as a threshold, so it cannot gate a decision
RQ-B I4.0 technology to SMT KPI mapping	8 (5 peer-reviewed, 3 industry-grade)	Isolated technology cases with quantified before/after KPI deltas; single site	No common denominator across technologies, so returns cannot be compared per unit of capital
RQ-C Sequencing across an SMT fleet	5 (2 peer-reviewed reviews, 3 industry reports)	Layered adoption patterns and lighthouse exemplars; descriptive	No per-line ordering method; exemplars are flagship sites not subject to fleet-wide capital rationing

Appendix B Assessment instruments and user guide

These seven forms execute the seven-step procedure of §4.1.4. They are self-contained: an assessor meeting the qualification of §4.1.1.2 can complete a full assessment using the forms together with the scoring rules of Tables 14–17 and 18–20, without reference to the rest of this report. Two rules govern completion throughout. Mark any unavailable field that cannot be sourced and never estimate it. Every indicator scored above 1 requires a documentary evidence reference.

Forms B1 to B7 are supplied as seven separate, directly usable Word documents rather than as tables reproduced in this appendix, so that an assessor can print or complete each one without extracting it from the report. This also keeps the report within its page limit. The seven files are:

Form_B1_Data_collection_sheet.docx	Form_B2_KPI_computation_sheet.docx
Form_B3_Practice_artefact_evidence_checklist.docx	Form_B4_Maturity_determination.docx
Form_B5_Readiness_assessment_questionnaire.docx	Form_B6_Readiness_scoring_sheet.docx
Form_B7_Recommendation_record.docx	

All seven forms, together with the decision-support tool, the Excel ROI ledger, the MATLAB visualisation scripts, the cross-validation models and every software output reported in this dissertation, are deposited in the project repository at:

GitHub repository: <https://github.com/sbindushree2502-ux/EG7302-DLSM-Framework>

The repository is organised by artefact type; folder 05_Assessment_Forms contains the seven form documents listed above. The reproduction record for every computational artefact is given in Appendix E, and the guide to the deployed application is given in Appendix F.

Appendix C Case-extraction template

Complete the template below once per admitted case under the blind-run procedure of §3.6.5. We record a confidence tag on every inferred field because case sources rarely report maturity or readiness directly, and presenting an inference as a measurement would overstate the strength of the validation. The template is reproduced as Table A2.

Table A2 Case-extraction template: fields, source requirement and the confidence tag applied to inferred values.

Field	Source requirement	Confidence tag
Firm and site profile; line or cell identified	Stated in the source	Reported

Field	Source requirement	Confidence tag
Baseline KPI vector (FPY, DPMO, OEE, Cpk where given)	Quantified before-values in the source	Reported
Inferred maturity level, with the inference rule logged	Derived from reported KPIs against Table 15	Inferred rule recorded
Inferred O/T/F readiness, with the inference rule logged	Derived from reported organisational and data facts against Tables 18–20	Inferred rule recorded
Technologies deployed and the actual sequence	Stated in the source	Reported
Reported outcome (KPI deltas, payback where stated)	Quantified after-values	Reported
Tool recommendation from the blind run	Generated before the outcome is read	Computed
Divergence class where recommendation and outcome differ	Assigned from the classes of §3.6.5	Assessed

Appendix D Ethics position and project risk assessment

This appendix records the project's ethics position and its risk assessment. No application was made to the University's research-ethics review system, because the work reported in this dissertation involves no human participants and collects no primary data: it is a desk-based design-science study drawing on published secondary sources, and it reproduces no proprietary, facility-identifying or commercially sensitive material. One contemplated activity would have changed that position. The instrument-reliability study specified in §3.4.3 and Table 8 required independent raters to score a set of fictional line vignettes so that inter-rater agreement could be computed. Although the material was invented, no personal data would have been collected, participation would have been voluntary and only anonymised scores reported, the activity would nonetheless have involved human participants. I raised this with the module leader and the project supervisor before contacting any participants; the guidance received was that involving human participants requires ethical approval, regardless of the risk profile. The study was therefore deferred to future work (§6.4, FW2) rather than run without approval, and no rater data of any kind exists or is reported anywhere in this dissertation. Table A3 records the project risk assessment, the mitigation applied and the residual position after mitigation. Risks are assessed based on their impact on the integrity of the research output rather than personal safety, consistent with the desk-based design declared in §1.5.

Table A3 Project risk assessment: risk, category, mitigation applied and residual position. Risks that materialised are recorded as such rather than closed.

Risk	Category	Mitigation applied	Residual position
Case sources insufficient to populate three admissible validations	Research integrity	Six to eight candidates screened to net three; C4 held in reserve; admissibility criteria fixed before screening	MATERIALISED and managed screening yielded one admissible case of four. The blind run was executed and published on that case (§4.4.2) and admissibility was not relaxed to manufacture a larger sample. Residual exposure recorded in §5.5 and carried as FW1
Effect-size priors misstate real returns	Construct validity	Priors registered with ranges in §3.7; Monte-Carlo and tornado quantify the consequence; fractions disclosed in Table 24	Bounded, not removed; identified as the dominant sensitivity in §4.5
Commercially sensitive process detail disclosed	Ethical and legal	Process descriptions kept generic; no facility, organisation or supplier identified; no proprietary parameter reproduced	Closed
Single-assessor bias in scoring the rubrics	Conclusion validity	Indicators written as observable evidence; reliability protocol pre-registered in Table 8	OPEN protocol specified but not yet executed; carried as FW2
Software unavailable or non-reproducible at examination	Reproducibility	All artefacts open-source or free-academic; archived engine and ledger authoritative; deployment treated as an access route only	Closed
Page limit exceeded, risking non-marking	Compliance	Extent monitored against the 60-page all-inclusive limit throughout; instruments placed in appendices; typography tightened rather than analysis removed	Closed

Appendix E Computational artefacts and reproduction record

This appendix states, for every computational artefact used in the study, what it is, what it depends on, what it produces and where in the dissertation that output is reported. It is written so that an independent reader with the source files can regenerate every quantitative result in Chapter 4 without contacting the author. All artefacts use free or academically licensed software; no commercial licence is required to reproduce any published figure. Table A4 inventories the artefacts, and Table A5 records the execution environment and the command that regenerates each result.

Every artefact listed below is deposited in the project repository, with its source. Folder 01_Decision_Support_Tool holds the engine and the application; 02_Cross_Validation holds the MCDA, influence-diagram and discrete-event models; 03_Excel_Ledger holds the ROI workbook; 04_MATLAB_Visualisation holds the chart scripts; 07_Software_Outputs holds the executed outputs.

GitHub repository: <https://github.com/sbindushree2502-ux/EG7302-DLSM-Framework>

Table A4 Computational artefact inventory: language, dependencies, output and the location at which each output is reported.

Artefact	Language / runtime	Size	External dependencies	Produces	Reported in
tool.py	Python 3.9.6	333 lines	None (standard library only)	Maturity gate, readiness decision, ranked sequence, ROI/NPV/payback per technology	§4.2–4.3; Tables 24, 25
dlsm_app.py	Python 3.9.6 + Streamlit	165 lines	streamlit; imports tool.py	Interactive front end over the same engine; no independent logic	§4.3.1; Figure 14
DLSM_ROI_Ledger.xlsx	Spreadsheet formulas	1 workbook	None; no macros	Independent re-computation of the financial model as an arithmetic cross-check	§4.3.2; Table 24
mcda_ahp_topsis.py	Python 3.9.6	141 lines	pymcdm 1.2.0; NumPy	Shannon-entropy weights; TOPSIS, VIKOR and PROMETHEE II rankings	§4.6.1; Tables 29–30; Figure 19
DLSM_Influence_Diagram.xdsl	GeNIe Academic (BayesFusion)	1 model	GeNIe Academic licence (free for academic use)	Influence diagram; expected utility per first investment under a uniform maturity prior	§4.6.2; Tables 31–32; Figure 20
smt_line_des.py	Python 3.9.6	330 lines	SimPy 4	Discrete-event model of the SMT front end; verification checks and 30 seeded replications	§4.6.3; Tables 34–35; Figure 21
viz_roi_payback_npv.m, viz_montecarlo_payback.m, viz_tornado_sensitivity.m, viz_readiness_radar.m	MATLAB (base)	4 scripts	No toolboxes required; shared data in dlsm_data.m	Publication figures rendered at 300 dpi from the verified data	Figures 15–18
reliability_scorer.py	Python 3.9.6	Standard library only	None	Weighted kappa, ICC(2,k), Cronbach's alpha and disagreement tabulation for the protocol of Table 8	Specified only not executed (§3.8, §6.4 FW2)

Table A5 Execution environment, determinism controls and reproduction commands.

Item	Value or command	Why it is recorded
Python version	3.9.6	pymcdm and SimPy behaviour is version-sensitive; the version pins the result
Package versions	pymcdm 1.2.0 · SimPy 4 · streamlit (dashboard only)	Only three third-party packages are used anywhere in the study
Determinism	All stochastic runs seeded; Monte-Carlo at 1,000 draws, DES at 30 replications with a 30-minute warm-up excluded	Re-running reproduces the published figures exactly rather than approximately
Reproduce the engine outputs	python3 tool.py	Regenerates the action, sequence and financial outputs of Tables 24–25
Reproduce the MCDA cross-check	python3 mcda_ahp_topsis.py	Regenerates entropy weights and all three rankings (Tables 29–30)
Reproduce the simulation	python3 smt_line_des.py	Runs verification checks then the 30-replication study (Tables 34–35)
Run the dashboard locally	streamlit run dlsm_app.py	Serves the interaction layer against the local engine
Public instance	https://dlsm-dashboard-cgwxs9tn32gzktrky3khw.streamlit.app	Free-tier hosting; demonstration only, with no availability guarantee (Table 37)

Appendix F Using the decision-support application

The interaction layer described in §4.3.1 is published so that the framework can be exercised rather than only read. It is available at <https://dlsm-dashboard-cgwxs9tn32gzktrky3khw.streamlit.app> and runs the same engine as tool.py; the dashboard contains no decision logic of its own, so any result it returns is reproducible from the command line. Two limitations are restated here because they bear on how the link should be treated: the instance runs on a free hosting tier with no availability guarantee and no persistence, and the archived engine and ledger remain the authoritative record (Table 37). The application demonstrates the procedure, not a production system.

The deployed application, the engine it imports and the seven assessment forms it replaces on paper are all available in the project repository (folders 01_Decision_Support_Tool and 05_Assessment_Forms).

GitHub repository: <https://github.com/sbindushree2502-ux/EG7302-DLSM-Framework>

One point of sequence matters. The application consumes an assessed maturity level and three readiness scores; it does not derive them. Steps 1 to 5 of Table 22 collecting evidence and scoring it on Forms B1 to B6 must be completed before the application is opened. Entering provisional scores, reading the recommendation and then revising the scores would invert the procedure and is expressly not the intended use. Table A6 states each input and the form on which it is determined; Table A7 states how each returned element is read.

Table A6 Application inputs: each field, its meaning, admissible values and the form on which it is determined.

Input field	Meaning	Admissible values	Determined on
Line name	Free-text identifier for the assessment record	Any string	Form B1
Assessed level	The Lean-maturity level assigned under the conjunctive rule of Table 15	L1, L2, L3, L4	Form B4

Limiting criterion	Which criterion held the line at its level, where it fell one short	Named criterion, or none	Form B4
Organisational (O)	Median of the five organisational readiness indicators	1–5	Form B6
Technical (T)	Median of the five technical readiness indicators	1–5	Form B6
Financial (F)	Median of the five financial readiness indicators	1–5	Form B6
T5 adjudicated dataset exists	Whether a labelled inspection dataset exists; gates ML-AOI eligibility independently of the T median	Yes / No	Form B5, indicator T5
Weakest KPI	The KPI furthest from its benchmark; drives the R5 technology mapping	FPY, defect PPM, OEE, throughput, changeover	Form B2
Capex envelope (£ per line)	Capital available for this line in the current cycle	Positive value	Form B1
Discount rate	Rate applied in the NPV calculation	0.08–0.12 (base 0.10)	Table 13

Table A7 Reading the application output: what each returned element means and the rule that produced it.

Output	Meaning	Governing rule
Action	STABILISE, DEFER, STAGE or DEPLOY the instruction for this line in this cycle	R1 and R4 (Tables 23, 7)
Eligible technology set	Which of the four technologies the maturity gate admits at all	R2 and R3
R_min / limiting dimension	The lowest of O, T and F, and which one it is; the binding constraint on deployment	§3.4.2; conjunctive screening
R_mean	Composite readiness, reported alongside R_min but never in place of it	§3.4.2
Ranked sequence	The order in which eligible, ready technologies should be funded	R6 NPV per £ of capital
Payback / NPV / IRR	Financial appraisal per recommended step at the entered capex and discount rate	§4.3.2 parametric model
Rationale string	A plain-language statement of which rule fired and why; intended to be pasted into Form B7	All rules; recorded for audit

Appendix G Abbreviations, symbols and notation

Table A9 lists every abbreviation used in this dissertation; Table A10 gives the symbols and notation specific to this framework, with the section at which each is defined.

Table A9 Abbreviations used in this dissertation.

Abbreviation	Meaning	Abbreviation	Meaning
AOI	Automated optical inspection	MLB	Main logic board
BG / CG	The two stacked sub-boards of the main logic board	ML-AOI	Machine-learning-based automated optical inspection
Cpk	Process capability index	NPV	Net present value
DES	Discrete-event simulation	OAT	One-at-a-time (sensitivity screening)
DMAIC	Define, Measure, Analyse, Improve, Control	OEE	Overall equipment effectiveness
DPMO	Defects per million opportunities	PAC	Practice-artefact compliance (§4.1.1.3)
DPPM	Defective parts per million	PLI	Production-Linked Incentive scheme (India)
DSR	Design science research	PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
EMS	Electronics manufacturing services	PROMETHEE	Preference Ranking Organisation Method for Enrichment Evaluations
FEDS	Framework for Evaluation in Design Science	ROI	Return on investment
FPY	First-pass yield	SLP	Substrate-like printed circuit board
I4.0	Industry 4.0	SMA	Surface-mount assembly
ICC	Intraclass correlation coefficient	SMT	Surface-mount technology
IoT	Internet of Things	SPC	Statistical process control
IRR	Internal rate of return	SPI	Solder-paste inspection
IT–OT	Information technology to operational technology integration	TOPSIS	Technique for Order Preference by Similarity to Ideal Solution
KPI	Key performance indicator	V&V	Verification and validation
L1–L4	The four Lean-maturity levels of the framework (§4.1.1)	VIKOR	VlseKriterijumska Optimizacija I Kompromisno Resenje
LSS	Lean Six Sigma	VSM	Value-stream map
MCDA	Multi-criteria decision analysis	WBS	Work-breakdown structure
MES	Manufacturing execution system		

Table A10 Symbols and notation specific to the framework.

Symbol	Meaning	Defined at
O, T, F	The organisational, technical and financial readiness dimensions	§4.1.2
R_min	The lowest of the three-dimension scores; the binding readiness constraint	§3.4.2
R_mean	The composite (mean) readiness score, reported but never acted on alone	§3.4.2
κ_w	Cohen's weighted kappa with quadratic weights, for agreement on maturity level	Table 8
ICC(2,k)	Intraclass correlation, two-way random effects, absolute agreement	Table 8
α	Cronbach's alpha, internal consistency of a readiness dimension	Table 8
R1–R6	The six sequencing rules of the decision logic	Table 23
O1–O5	The five locked project objectives	Table 2
FW1–FW5	The five directions for further work	Table 42
LM-01 – LM-05	The five self-learning objectives and their evidence	Tables 44–45