ISSN: 0379-5462
IPPTA: Quarterly Journal of Indian Pulp and Paper Technical Association
Vol. 38, E2, 2026, p. 146-149

IPPTA

Industrial AI for Future-Ready Paper Mills

Sneh Patel
Chief Operating Officer, JMC Paper Tech USA LLC

Abstract: *The Indian paper industry faces converging pressure from tightening policy (Extended Producer Responsibility, Business Responsibility and Sustainability Reporting and net-zero commitments), rising sustainability expectations, and the rapid maturation of industrial artificial intelligence (AI). Most mid-sized recycled-fibre and agro-based mills, however, still operate disconnected control, quality, laboratory, maintenance and enterprise systems, leaving decisions reactive and operator knowledge undocumented. This paper presents MillMind, an AI-based intelligence platform that unifies operational technology, information technology and analytics into a single operator-facing layer for paper mills. The platform is structured as a four-layer stack (connectivity and edge, data platform, AI and analytics, and decision and insight) and provides real-time monitoring, quality forecasting of burst, RCT, CMT, basis weight and moisture, predictive maintenance, per-tonne energy and resource optimisation, and capture of institutional knowledge. Early deployment at a recycled-fibre containerboard mill in Ciudad Juárez, Mexico, indicates 10–20% lower unplanned downtime, 3–7% lower steam consumption per tonne and 30–50% faster grade changeovers, together with audit-ready data for environmental reporting. The results demonstrate a practical convergence of policy compliance, sustainability performance and operating economics for mid-sized mills.*

Keywords: *Industrial AI, Predictive maintenance, Quality forecasting, Energy optimisation, Recycled fibre.*

1. Introduction

The pulp and paper sector in India is approaching an inflection point. Three forces that were until recently managed as separate concerns, namely policy, sustainability and technology, are now arriving together and demand a common response from the mid-sized recycled-fibre and agro-based mills that form the backbone of the industry.

India is among the world's largest paper producers, and a substantial share of national output comes from several hundred mid-sized recycled-fibre and agro-based mills rather than from a few large integrated producers. These mills typically run one or two machines in the 50–300 tonne-per-day range, operate on tight margins, and compete against both larger domestic producers and imports. Their scale gives them limited capital for digitalisation, yet they carry the same regulatory and customer obligations as the largest mills. This combination of constrained resources and rising expectations is what makes a modular, affordable intelligence layer relevant to them.

On the policy side, Extended Producer Responsibility under the Plastic Waste Management Rules requires auditable traceability of fibre flows and recycled content [1]; the Business Responsibility and Sustainability Reporting (BRSR) framework of the Securities and Exchange Board of India demands disclosure across detailed environmental, social and governance indicators [2]; and India's Nationally Determined Contribution and net-zero commitments are translating into sector emission-intensity targets through Perform-Achieve-Trade cycles [3]. Revised effluent norms add further pressure on process variability [4]. Each of these obligations shares one technical requirement: granular, time-stamped, machine-verified data.

On the sustainability side, brand owners increasingly place verified footprint reporting, specific water consumption below 40 m³ per tonne, and a rising renewable-energy share on their supplier contracts. Mills that cannot measure and continuously improve their environmental performance risk losing access to premium customers, while those that can substantiate it from primary process data command better prices and more durable relationships.

The third force is the maturation of industrial AI. The cost of time-series data ingestion has fallen sharply; open connectivity standards such as OPC-UA and MQTT have made integration affordable [5, 6]; and gradient-boosted and transformer-based learning methods have crossed the production-ready threshold for soft-sensor and forecasting tasks. Edge computing, lightweight time-series databases and multilingual operator interfaces are now standard components rather than research artefacts.

Yet most mid-sized mills cannot exploit these developments, because their Distributed Control System (DCS), Quality Control System (QCS), Enterprise Resource Planning (ERP), laboratory and maintenance systems operate as disconnected islands. Decisions remain reactive, energy is effectively a black box, and three decades of operator intuition leave the mill at every retirement. This paper describes MillMind, an AI-based intelligence platform engineered to unify these systems and to convert mill data into real-time decisions on quality, energy, maintenance and production.

2. Materials and Methods

2.1 Operational Gaps in Mid-Sized Mills

Field experience across more than two dozen recycled-fibre and agro-based mills in India, Mexico and South-East Asia points to six recurring gaps that explain why mid-sized mills underperform their theoretical efficiency frontier, and these gaps define the requirements the platform was built to meet. First, data silos: the DCS, QCS, ERP, laboratory and maintenance systems run as independent islands, so insights that depend on their combined state cannot be extracted within a shift. Second, reactive operations: sheet breaks and off-spec production are discovered only after the laboratory confirms them, by which time several hundred metres of product may already be lost. Third, knowledge loss: operator intuition leaves with every retirement in a high-turnover labour market. Fourth, lack of context: generic industrial-IoT platforms built for discrete manufacturing do not understand stock-prep dynamics, refining curves or dryer steam economy. Fifth, energy opacity: steam, power and chemicals are 35–45% of variable cost, yet per-tonne intensity is rarely visible in real time. Sixth, adoption friction: most industrial software ships in English and is designed for desk-based engineers rather than plant-floor operators.

2.2 System Architecture

The platform is structured as a four-layer intelligence stack that follows the ISA-95 hierarchy and the Purdue reference model [7], so that the contribution lies in the data model and algorithms rather than in architectural novelty. Sensors and machines feed the platform from the bottom up, and insight flows back to operators in real time. The four layers are summarised in Table 1.

Table 1. Four-layer architecture of the MillMind platform.

Layer	Function	Representative Components
1 — Connectivity & Edge	Acquires raw signals from every plant asset and harmonises them into a single ingestion stream.	MQTT broker, OPC-UA, PLC/DCS/QCS gateways, edge compute, lab and ERP adapters.
2 — Data Platform	Stores, contextualises and serves mill data for both control and analytical workloads.	Time-series database, relational store, contextual data model, master data management.
3 — AI & Analytics	Houses domain-specific machine-learning models and rule engines that turn data into prediction.	Quality-forecasting and predictive-maintenance models, energy engine, anomaly detection, soft sensors.
4 — Decision & Insight	Translates analytical output into actionable guidance for operators and managers.	Operator dashboards, mobile alerts, recommendation engine, role-based reports, multilingual interface.

Two design choices are central. The time-series database is treated as a first-class component, because paper-mill data is high-frequency, irregularly sampled and dominated by trends rather than discrete events. The contextual data model links operational tags to business meaning, allowing, for example, a vibration rise on a named drive to be interpreted as a press felt-roll bearing event during a specific grade transition. The architecture has been tested at more than 100,000 tag updates per second on commodity hardware, with tiered retention (full resolution for thirty days, downsampled thereafter) to manage storage growth. Edge computing is used at the connectivity layer so that signal acquisition and first-level buffering continue even when the link to the data platform is intermittent, a common condition at emerging-market sites.

2.3 Functional Modules

The platform presents as an integrated set of modules sharing one data model: Production, Quality Laboratory, Maintenance, Stores and Spares, Energy, OCC and Furnish, Sales and Despatch, Planning, and Environment-Health-Safety. A change recorded in one module propagates to the others without re-entry, eliminating spreadsheet bridges and reconciliation between subsystems. Modules are activated incrementally; a typical mill begins with Production, Quality and Maintenance and expands as integration matures. Because every module shares the same schema and interfaces, expansion requires configuration rather than re-implementation.

2.4 AI Capabilities

Five analytical capabilities run on the unified data model. Real-time monitoring ingests basis weight, moisture, freeness, vacuum, steam pressure, drive load, vibration, consistency, pH and several hundred further signals, and computes paper-industry key performance indicators such as Overall Equipment Effectiveness and steam, power and water consumption per tonne, each anchored to grade and baseline so that performance is interpretable within the shift.

Quality forecasting uses gradient-boosted models, trained on each mill's grades and furnish, to predict burst, RCT, CMT, basis weight and moisture several minutes ahead of the laboratory result. Each model consumes a feature vector assembled from stock-prep, machine and utilities tags — refining specific energy, freeness, consistency, jet-to-wire ratio, vacuum, steam pressure, dryer load and ash content. When a forecast crosses a specification or baseline threshold, a specific corrective recommendation is issued. Forecast accuracy is tracked as the residual between predicted and laboratory values for each reel; a sustained shift in residual bias or variance triggers re-calibration. This continuous-validation loop is what allows a soft sensor to be treated as an operational instrument rather than an advisory indicator.

Predictive maintenance monitors vibration, bearing temperature, motor current, vacuum and dryer load to flag bearing wear, motor-insulation degradation and vacuum-system fouling days or weeks ahead of failure. The first version covers the asset classes responsible for most unplanned downtime in recycled-fibre mills — section drives and gearboxes, vacuum systems, refining lines, press-section bearings and rolls, dryer-cylinder bearings and condensate systems, and pulpers and screens — each with a dedicated model family and mill-specific calibration. Energy and resource optimisation measures steam, power, water and chemicals per tonne against grade-specific baselines. Institutional knowledge capture logs every parameter change, quality deviation, maintenance intervention and grade change together with its operating context, building a searchable record that survives operator turnover.

2.5 Deployment and Implementation

The platform supports on-premise, hybrid and cloud deployment, all sharing the same data model and interfaces; the choice depends on a mill's security, sovereignty and connectivity requirements. Implementation proceeds through discovery, connectivity, configuration, model training and continuous operation, with most mills reaching core functionality within four to six months of kickoff. Operator interfaces are multilingual and designed for plant-floor use, which is essential for adoption where operators have limited screen time. Access is role-based, with a tamper-evident audit trail of operator actions and recommendation acceptances and encryption of data in transit and at rest; the on-premise option keeps all operational data inside the mill perimeter.

3. Results and Discussion

3.1 Case Site

The platform is being deployed at a greenfield recycled-fibre containerboard mill in Ciudad Juárez, Mexico, which serves the cross-border and Mexican packaging markets. The mill produces Test Liner and Fluting Medium in the 100–240 GSM range and Core Board in the 200–400 GSM range on a 100% recovered-fibre (OCC) furnish, with a design capacity of 58,500 tonnes per year. Because it is a recycled-fibre mill operating under demanding customer and regulatory expectations, it is a close analogue of the typical mid-sized Indian mill, and its results translate directly to that context. Deployment covers all modules; the connectivity layer integrates the DCS, QCS, weighbridge, laboratory system and ERP, with a time-series store and an MQTT broker for sensor ingestion.

3.2 Operating Results

Even at an early stage of ramp-up, the deployment produces measurable benefits, summarised in Table 2. In February 2026 the mill produced 2,431 tonnes at approximately 93.5% saleable output while progressing toward design capacity. The indicative ranges are deliberately conservative: mills that integrate the platform into daily shift reviews and management cycles achieve outcomes toward the upper bound, whereas those that treat it as a one-time installation capture far less. For a typical 200-tonne-per-day mill, cumulative savings on energy, downtime and quality losses recover the investment within roughly twelve to eighteen months, after which institutional learning continues to compound.

Table 2. Indicative operating impact observed during early deployment.

Indicator	Indicative Impact	Mechanism
Unplanned downtime	10–20% reduction	Predictive maintenance shifts interventions into planned windows; chronic failure modes are addressed through root-cause learning.
Steam consumption per tonne	3–7% reduction	Real-time visibility of dryer load, pocket ventilation and condensate return closes habitual inefficiencies.
Grade changeover time	30–50% faster	AI-guided transitions using mill-specific best historical settings replace trial-and-error tuning.
Institutional knowledge	Continuous capture	Operator actions, quality deviations and interventions are logged with context and remain searchable.

Quality-forecasting performance over the same period is shown in Figure 1 and summarised in Table 3. Across the five reported properties the forecasts track the laboratory closely enough to act on within the shift, with mean absolute errors that stay comfortably inside normal specification bands. Figure 1 illustrates this for ring crush test (RCT), where the model anticipates two grade transitions across forty consecutive reels.

3.3 Energy as the Principal Cost and Emissions Lever

Because steam, power, water and chemicals dominate the variable cost of recycled-fibre containerboard, energy is both the largest cost-reduction opportunity and the largest single contributor to Scope 1 and Scope 2 emissions; the same intervention therefore improves operating margin and sustainability performance at once. The platform answers, for every tonne of saleable paper, how much of each resource was consumed and where, and compares it against a grade-specific model baseline. Four optimisation loops act on this envelope: a steam-economy loop comparing actual dryer steam to a predicted baseline for the current grade and ambient conditions; a power-efficiency loop benchmarking each drive against its design curve; a water-balance loop that flags fresh-water creep as the recycle loop is closed; and a chemical-dosing loop that ties additive consumption to incoming furnish

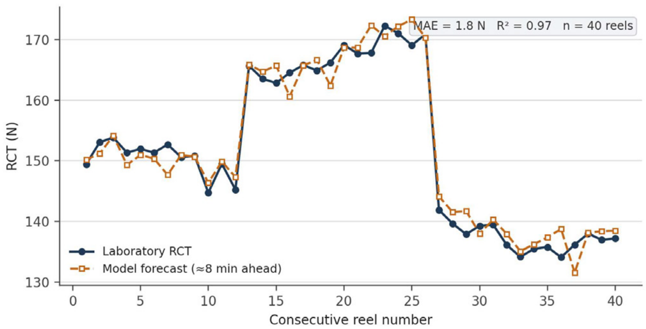

Figure 1. Model forecast versus laboratory ring crush test (RCT) across forty consecutive reels spanning three grades; the soft sensor tracks grade transitions about eight minutes ahead of the laboratory (MAE 1.8 N, R^2 0.97).

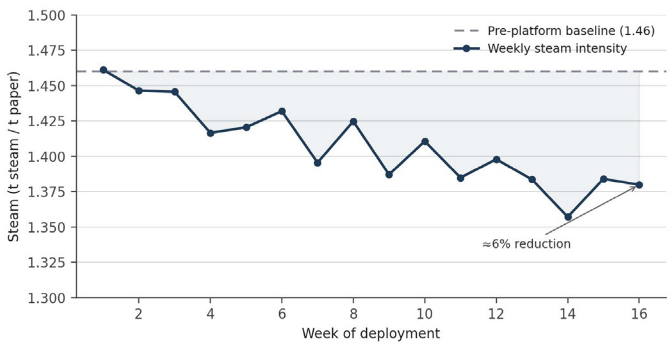

Figure 2. Weekly dryer-steam intensity at the case site against the pre-platform baseline of 1.46 t steam per tonne, showing an approximately six per cent reduction over sixteen weeks as the steam-economy loop is tuned to each grade.

Table 3. Quality-forecast accuracy by property during early deployment; errors are mean absolute deviations from the laboratory value (n = 40–120 reels per property).

Quality property	Forecast lead time	Mean absolute error	R^2
Burst strength	6–10 min	9.1 kPa	0.93
Ring crush (RCT)	6–10 min	1.8 N	0.97
Concora medium (CMT)	6–10 min	5.7 N	0.92
Basis weight	1–2 min	1.1 g/m²	0.97
Moisture	1–2 min	0.15 %	0.96

quality, replacing fixed dosage tables with condition-responsive dosing.

The cumulative effect of these loops at the case site is traced in Figure 2, which compares weekly dryer-steam intensity with the pre-platform baseline.

3.4 Convergence with Policy and Sustainability

A central result is that the same data fabric that runs the mill also produces, as a by-product of normal operation, the data required for major policy and sustainability obligations. Table 4 maps each driver to its underlying data requirement and to the platform capability that supplies it. Because every reported figure is machine-derived, time-stamped and traceable to its underlying tags, the mill becomes audit-ready: an EPR auditor, a BRSR assurer or a customer assessor can be answered from live data rather than from reconciliations prepared after the fact. The same continuous baseline-and-achievement accounting positions the mill to participate in emerging carbon-market mechanisms [3][8].

Table 4. Mapping of policy and sustainability drivers to platform capabilities.

Driver	Data Requirement	Platform Capability
EPR — recycled-content traceability	Lot-level OCC sourcing, yield per bale, recycled-content percentage, despatch traceability.	OCC & Furnish and Sales & Despatch modules with the contextual data model.
BRSR — ESG disclosure	Environmental, social and governance key performance indicators.	Energy, EHS and HR modules with the audit trail.
Scope 1 and 2 emissions	Boiler fuel, on-site combustion and purchased energy with grid-emission factors.	Energy module with source-attributed accounting.
Effluent norms	Continuous effluent volume, pH, BOD, COD and TSS with threshold alarming.	Connectivity layer and EHS module with the alerting engine.
Specific water consumption	Per-tonne fresh-water draw and recycle-loop accounting.	Energy module with grade-specific baselines.
PAT / carbon-market readiness	Specific energy baseline and achievement data with verifiable provenance.	Energy module with audit trail and archive.

3.5 Discussion, Limitations and Future Work

The case indicates that the principal barrier to digitalisation in mid-sized mills is not algorithmic but structural: disconnected systems and undocumented knowledge. Unifying operational, laboratory and enterprise data on a contextual model, rather than deploying generic industrial-IoT dashboards that are not engineered around paper-machine physics, is what converts data into operational decisions. The economic case rests less on any single optimisation than on the compounding effect of capturing knowledge from every shift, which is difficult for a competitor to replicate.

Beyond this structural barrier, two adoption barriers deserve explicit attention. The first is a skills gap: mid-sized mills rarely employ dedicated data engineers or data scientists, and plant teams accustomed to manual logs and spreadsheet reporting may initially distrust model-driven guidance. Durable benefit therefore depends as much on training, change management and visible early wins as on model accuracy; the platform addresses this through multilingual plant-floor interfaces, role-based views and explainable recommendations rather than opaque scores. The second is adoption cost: connectivity hardware, integration effort and subscription or licence fees compete with maintenance and production capital for scarce funds, and returns accrue only after several months of data accumulation and model training. The platform lowers this hurdle through incremental module activation, which spreads expenditure over time and lets each phase be justified on demonstrated benefit, and through on-premise or hybrid deployment on commodity hardware. Recognising these barriers as organisational as much as technical is what separates mills that capture the full benefit from those that do not.

Two limitations should be stated. First, the operating results reported here come from a single, early-stage greenfield deployment; the indicative ranges will need confirmation across a wider set of mills and over longer periods, particularly at brownfield sites with legacy instrumentation. Second, model performance depends on data quality and on organisational follow-through; mills that do not act on alerts capture only a fraction of the available benefit, so the platform is best understood as a socio-technical intervention rather than a purely technical one. Future work follows two directions. The first is a domain-specific assistant that allows operators to query mill history and standard procedures in natural language. The second is selective closed-loop control, in which well-validated recommendations are permitted to act on the control system within defined envelopes. Both extend the same data fabric rather than replacing it.

Conclusion

Policy compliance, sustainability performance and operating economics can no longer be pursued as separate tracks in the mid-sized paper mill. This paper has argued that the viable response is convergence on a single, machine-verifiable data fabric, and has described MillMind as one realisation of that approach: a four-layer architecture with integrated modules and five AI capabilities, tuned to the realities of recycled-fibre mills. Early evidence from a containerboard mill in Mexico points to materially lower unplanned downtime and steam consumption, faster grade changeovers, improved retention of institutional knowledge, and audit-ready environmental data, with payback indicatively within twelve to eighteen months. Mills that begin building their data and AI fabric now are likely, within three to five years, to hold a structural advantage in cost, compliance and customer trust.

Acknowledgements

The author thanks the management and the engineering and operations teams at Balaji JMC Paper Mill for supporting the deployment that anchors this work, and the development partners who contributed to the analytics layer.

References

1. Ministry of Environment, Forest and Climate Change, Government of India. Plastic Waste Management Rules and Extended Producer Responsibility framework notifications (2022).
2. Securities and Exchange Board of India. Business Responsibility and Sustainability Reporting (BRSR) framework, SEBI Circular SEBI/HO/CFD/CMD-2 (2021).
3. Bureau of Energy Efficiency, Government of India. Perform, Achieve and Trade (PAT) Scheme — Pulp and Paper Sector (2023).
4. Central Pollution Control Board. Comprehensive Industry Document on Pulp and Paper, Environment (Protection) Rules (2015).
5. OPC Foundation. OPC Unified Architecture Specification, Part 1: Overview and Concepts (2017).
6. OASIS. MQTT Version 5.0 Specification (2019).
7. International Society of Automation. ANSI/ISA-95 Enterprise-Control System Integration, Parts 1–5 (2010).
8. World Resources Institute and World Business Council for Sustainable Development. Greenhouse Gas Protocol: Corporate Accounting and Reporting Standard, Revised Edition (2004).