© June 2026 | IJIRT | Volume 13 Issue 1 | ISSN: 2349-6002

Machine Learning–Driven Real-Time Prediction of Critical Paper Quality Properties: A Review of Soft Sensors, Process Analytics, and Industrial Deployment in Pulp and Paper Mills

Dr. Shwet Vashishtha
R.P.P.T. Laboratory, National Test House, Ghaziabad, India

***Abstract*— The pulp and paper industry is rapidly transitioning from delayed, laboratory-based quality verification toward real-time, data-driven quality control systems built on machine learning, soft sensors, and Industry 4.0 architectures [1–3]. This transition is particularly important for critical paper properties such as grammage or basis weight (g/m²), Cobb value, tensile index, burst index, tear index, ring crush strength, and related strength parameters, as these variables directly influence convertibility, end-use performance, production cost, and customer claims, while several of them are still measured offline or at relatively low sampling frequency [4–6]. Machine-learning models can bridge this gap by learning the relationships among process variables, furnish characteristics, chemical additives, environmental conditions, and delayed laboratory test results, thereby enabling soft sensors for real-time prediction and closed-loop or advisory control [1–2,7–8].**

This review critically examines the current state of data-driven quality prediction in papermaking, with particular emphasis on the online prediction of basis weight, Cobb value, and strength-related properties. Published industrial case studies and related research indicate that linear regression, support vector regression, random forests, gradient boosting methods, neural networks, and hybrid first-principles plus machine-learning models have been applied with varying degrees of success, depending on data quality, process stability, sensor availability, and target-property dynamics [7–11]. Recent evidence further suggests that hybrid and adaptive soft-sensor frameworks are particularly promising, as they combine physical consistency with the flexibility required to handle grade changes, regime shifts, and transient operating conditions [2,7,10–11].

The review organizes the field around six major themes: measurement challenges in paper quality, data infrastructure requirements, model families and feature engineering, interpretability and deployment, control-system integration, and research gaps for Indian mills. A practical framework is proposed for mills seeking to deploy machine-learning-based quality prediction models, beginning with data governance and target-property selection and extending to model maintenance, human oversight, and economic validation [1–2,5,12]. The paper concludes that machine learning has matured sufficiently to support real-time prediction of key paper properties in production environments; however, long-term value depends less on algorithmic novelty than on reliable data pipelines, robust recalibration strategies, and effective integration with process knowledge and mill operations [1–2,7,10].

***Index Terms*— Machine learning; soft sensors; paper quality prediction; basis weight (GSM); Cobb value; strength indices; papermaking process control; Industry 4.0**

I. INTRODUCTION

Quality control in papermaking has always been constrained by a mismatch between the speed of the process and the speed of many property measurements. A paper machine generates large volumes of process data in real time, but important product properties such as tensile strength, burst strength, tear strength, ring crush strength, and water absorptiveness are often confirmed through laboratory tests that arrive minutes or hours later [1,2,5]. This delay limits the ability of operators to correct deviations at the source and often leads to broader variability, conservative operating targets, excess chemical use, downgraded reels, and avoidable waste [1,2,4,8].

Online quality control systems (QCS) have long measured or estimated variables such as basis weight, moisture, caliper, and ash, yet not all critical end-use properties are directly measurable at sufficient speed or reliability across grades [2,5,6]. In practice, several mills still depend on manual or semi-automated laboratory measurements for Cobb value and many strength indicators, especially in packaging and specialty grades where the final commercial value of the paper depends on a narrow property window. This creates a strong case for data-driven soft sensors that infer hard-to-measure quality variables from readily available process signals.

The integration of real-time process data, online quality measurements, and laboratory results into a machine-learning framework for paper quality prediction is illustrated in Figure 1.

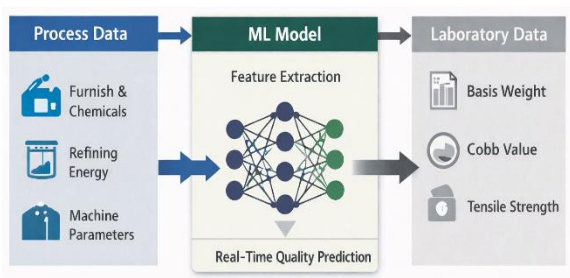

Figure 1: Conceptual framework for real-time paper quality prediction using process and laboratory data

Machine learning has emerged as a powerful approach for this task because papermaking is highly multivariate, nonlinear, and dynamic. Furnish composition, refining intensity, filler loading, wet-end chemistry, retention-aid dosage, machine speed, stock consistency, drainage behavior, drying profile, and ambient conditions can all influence final paper properties, often with complex interactions and time delays that are difficult to represent through conventional empirical equations. Modern machine-learning models can exploit these relationships when process and laboratory data are sufficiently rich and properly time-aligned, enabling accurate property prediction, earlier deviation detection, and proactive quality control.

This review article focuses on data-driven quality control models for the real-time prediction of key paper properties, with particular emphasis on basis weight, Cobb value, and strength indices. The paper aims to synthesize available research and industrial evidence, compare major modeling approaches, identify implementation requirements, and outline a practical roadmap for mills, especially in the Indian context, where digitalization is increasing but data quality and instrumentation levels vary widely across operations.

A major challenge in modern paper mills is that process data are often distributed across multiple platforms, including distributed control systems (DCS), quality control systems (QCS), laboratory information systems (LIMS), and production historians. Integrating these heterogeneous data streams into a unified predictive framework is essential for the development of reliable soft sensors and machine-learning-based quality models. The ability to convert real-time process signals into early estimates of final product quality has the potential to reduce off-spec production, improve raw-material utilization, and support economically optimized machine operation [4,7,9].

1.1 Why these properties matter

1.1.1 Basis weight (GSM)

Basis weight, commonly expressed as grammage or GSM (g/m²), remains the most fundamental quality variable in papermaking because it directly influences product yield, raw material consumption, production cost, convertibility, printability, and several downstream strength properties such as tensile strength, burst strength, stiffness, and ring crush performance [3,4,13,16]. Even a small deviation in basis weight can lead to significant economic loss due to excess fiber consumption, downgraded reels, off-spec production, and customer complaints, particularly in packaging, printing, and specialty paper grades.

Basis Weight = Mass of paper (g) / Area (m²)

In industrial papermaking practice, even a deviation of ±1–2 g/m² can significantly affect fiber consumption, cost per square meter, and the final mechanical performance of the product.

The major control points and process variables influencing basis-weight development across the paper machine are shown in Figure 2.

Figure 2: Major process stages and control points influencing basis weight across the paper machine, including stock preparation, headbox consistency, drainage, pressing, drying, and final reel measurement.

From a process-control perspective, basis weight is strongly governed by stock flow rate, headbox consistency, machine speed, drainage behavior, and sheet formation dynamics. Although most modern mills employ online Quality Control Systems (QCS) for final basis-weight measurement, predictive models remain highly valuable for upstream process sections where earlier estimation can reduce startup losses, improve machine stability, and support anticipatory control actions such as stock-flow adjustment and headbox consistency correction [3,4,12].

Recent studies on tissue and other paper-machine applications have shown that hybrid soft-sensor models, integrating process knowledge with machine-learning algorithms, can significantly improve early-stage basis-weight prediction, particularly during transient operating conditions, startup phases, and regime shifts [12]. This makes basis weight an ideal target variable for real-time predictive quality control, process optimization, and closed-loop control in modern paper mills.

1.1.2 Cobb value

Cobb value, generally measured according to standardized timed water-absorption tests such as ISO 535 / TAPPI T 441, is a critical quality parameter for packaging, converting, and specialty paper grades because it directly influences dimensional stability, printability, adhesive performance, and resistance to moisture under humid service conditions [16,17]. In practical mill operations, higher or unstable Cobb values can lead to increased converting waste, poor glue bonding, print smudging, dimensional distortion, reduced runnability, and a higher risk of customer complaints, making it economically significant even when it is not measured continuously online.

The Cobb test is generally expressed as the mass of water absorbed per unit area of paper in a specified time, usually 60 seconds, and is reported in g/m².

Cobb Value = Mass of water absorbed (g)/Area (m^2)

The principal process and material factors influencing Cobb value across the paper machine and surface treatment stages are illustrated in Figure 3.

Figure 3: Key process and material variables influencing Cobb value, including sizing chemistry, porosity, drying, and surface structure.

In paper science terms, Cobb value is strongly influenced by internal sizing efficiency, surface sizing, sheet porosity, fiber bonding density, ash content, drying conditions, and surface structure. The performance of sizing agents such as AKD (alkyl ketene dimer), ASA (alkenyl succinic anhydride), and rosin-based systems, along with starch pickup at the size press, plays a decisive role in controlling water penetration and surface absorptiveness [5,6,17].

For packaging and board grades, Cobb value is particularly critical because excessive water absorption adversely affects adhesive bonding, carton stiffness, compression strength, and dimensional integrity during storage and transport. Conversely, an excessively low Cobb value may negatively influence print ink absorption and glue anchorage in converting applications.

Because Cobb depends simultaneously on surface chemistry, pore structure, drying history, and furnish-related characteristics, it is an excellent candidate for machine-learning-based soft-sensor development, provided that process signals and delayed laboratory test results are accurately synchronized in time [9,11]. Key predictive variables may include size dosage, starch addition, ash content, caliper, moisture profile, drying temperature, machine speed, and porosity-related proxies, which together can support real-time prediction and early process correction.

1.1.3 Strength indices

Strength-related properties such as tensile strength, burst strength, tear strength, compressive strength, ring crush strength, and wet tensile strength are central to grade qualification, product certification, and customer acceptance, particularly for packaging papers, paperboards, kraft papers, and specialty grades [15,16]. These indices directly govern end-use performance during printing, converting, stacking, transportation, and final application.

In industrial practice, strength parameters are often the deciding factors for acceptance of packaging and corrugating grades because they directly influence box compression performance, stacking load capacity, edge crush resistance, and durability during handling and transport. Even small deviations in tensile or ring crush values can lead to downgraded reels, customer complaints, and significant commercial loss.

The development of paper strength is fundamentally governed by fiber morphology, fiber length, fiber bonding area, refining intensity, sheet density, and wet-end chemistry. Long-fiber content, fibrillation during refining, hydrogen bonding efficiency, filler loading, starch addition, and retention-aid chemistry all play a major role in determining the final mechanical strength of the sheet [5,15].

The major material and process variables influencing strength development in paper are illustrated in Figure 4.

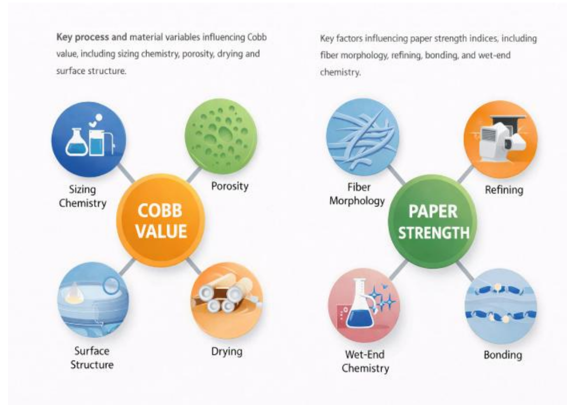

Figure 4: Key factors influencing paper strength indices, including fiber morphology, refining, bonding, and wet-end chemistry.

For example, tensile strength is strongly dependent on inter-fiber bonding and sheet formation, whereas tear strength is more closely related to fiber length and individual fiber strength. Burst strength depends on both tensile integrity and sheet compactness, while ring crush and compressive strength are especially critical for packaging board and corrugated applications.

Industrial case studies have already demonstrated successful machine-learning-based prediction of wet tensile and other strength-related properties, with reported improvements in chemical optimization, furnish balance, and variability reduction when such models are deployed in production environments [4,5,9]. Because these properties depend simultaneously on furnish composition, refining energy, fiber morphology, moisture profile, and wet-end chemistry, they are highly suitable targets for data-driven predictive models capable of handling multiple interacting process variables in real time.

1.2 Evolution from statistical process control to machine learning

Conventional paper-mill quality systems have traditionally relied heavily on statistical process control (SPC), control charts, linear regression models, and operator experience for maintaining process stability and product consistency [13,14]. Common tools such as Shewhart control charts, moving-range charts, trend plots, and capability indices (Cp/Cpk) have long been used to monitor key quality variables such as basis weight, moisture, caliper, ash content, and strength properties.

These conventional methods remain highly useful for detecting process instability, abnormal variation, and obvious shifts from target values; however, they often struggle to capture the nonlinear, time-dependent, and multiscale relationships that link upstream process conditions to final paper properties. In papermaking, the effects of furnish variation, refining intensity, wet-end chemistry, drainage, drying history, and environmental conditions are often interdependent and delayed in time, making them difficult to represent through simple linear or rule-based models [13].

The broader modeling literature in pulp and paper engineering has long recognized the importance of integrating process understanding with advanced computational approaches to improve optimization, prediction, and control. Early empirical and simulation-based models laid the foundation for modern predictive analytics by identifying critical process–property relationships.

The transition toward machine learning has been accelerated by three major developments. First, modern mills now generate significantly richer process histories from distributed control systems (DCS), quality control systems (QCS), drives, sensors, historians, and laboratory information systems (LIMS) [4,7,9]. Second, computational tools for regression, feature engineering, anomaly detection, and adaptive modeling have become increasingly accessible for industrial deployment. Third, successful soft-sensor applications have demonstrated measurable economic benefits through improved quality consistency, reduced waste, lower chemical consumption, and better machine efficiency, making predictive quality models more attractive to mill management [4,9,12].

The transition from conventional statistical quality control to machine-learning-based predictive quality systems is illustrated in Figure 5.

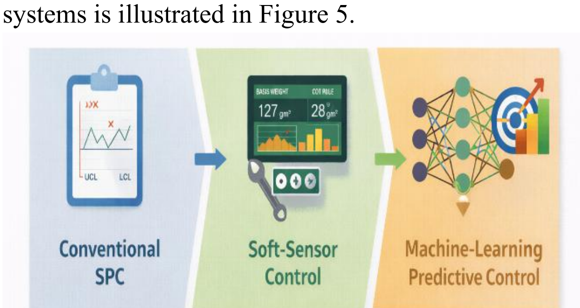

Figure 5: Evolution of paper quality control from conventional SPC methods to machine-learning-based predictive control systems.

In many modern paper mills, the objective is no longer limited to explaining quality variation after production but rather to predict property deviations early enough to intervene before off-spec material accumulates. This shift transforms machine learning from a purely analytical tool into an operational quality-control asset and decision-support system, enabling proactive process correction, advisory control, and, in advanced cases, closed-loop optimization.

II. DATA SOURCES AND DIGITAL ARCHITECTURE

2.1 Data layers required for paper-property prediction

Successful quality-prediction models in papermaking depend on the integration of multiple heterogeneous data streams rather than reliance on a single source [4,9,12]. The accuracy and robustness of machine-learning-based soft sensors are directly linked to the quality, synchronization, and completeness of these data layers.

Relevant input variables typically include stock preparation parameters, pulp blend ratios, furnish composition, freeness or drainage indicators (°SR / CSF), refining specific energy, filler loading, starch addition, wet-end chemical dosage, retention-aid levels, machine speed, headbox consistency, wire-section drainage variables, press-section nip pressure, dryer-section steam temperature, moisture profile, caliper, ash content, reel metadata, and laboratory quality test results. For packaging, specialty, and coated grades, additional variables such as surface sizing pickup, coating recipe, environmental humidity, operator interventions, and grade-specific recipe metadata may also be highly influential.

The major categories of process and quality data typically required for machine-learning-based paper-property prediction are summarized in Table 1.

Table 1: Typical data layers for machine-learning-based paper quality prediction

Data layer	Typical variables	Relevance for prediction
Stock preparation data	pulp blend ratio, freeness (°SR/CSF), refining energy, consistency	fiber development and strength prediction
Wet-end chemistry	filler %, starch dosage, AKD/ASA, retention aid	Cobb, strength, ash, sizing performance
Process control data	machine speed, headbox pressure, flow rate, wire vacuum	basis weight, formation, moisture
Press and dryer data	nip pressure, steam temperature, drying profile	density, moisture, Cobb, strength
Online quality data	basis weight, moisture, caliper, ash	real-time quality prediction
Laboratory data	Cobb, tensile, burst, tear, ring crush	target variables for model training
Environmental / metadata	ambient RH, temperature, grade code, operator actions	drift and variability correction

A recurring lesson from both industrial case studies and research literature is that prediction quality depends as much on data engineering and process alignment as on model selection [9,11].

From a paper-technology perspective, these data layers represent the complete process history of sheet formation, consolidation, drying, and finishing. Since final paper properties develop progressively across the machine, no single signal can adequately represent the resulting quality response.

A recurring lesson from both industrial case studies and research literature is that prediction quality depends as much on data engineering and process alignment as on model selection [9,11]. Signals must be cleaned, filtered, time-stamped, aligned, and tagged according to grade, recipe, furnish type, and production reel. Particular attention must be given to synchronization with the true residence time of the sheet as material moves from stock preparation through formation, pressing, drying, size press, and reel buildup.

If a laboratory quality result is incorrectly linked to the wrong upstream process window, even highly sophisticated machine-learning models can produce unstable, physically unrealistic, or misleading predictions. Therefore, time-delay compensation, lag feature generation, and residence-time mapping are essential steps in the development of reliable real-time paper-property prediction systems.

2.2 Role of soft sensors

Soft sensors are predictive models that estimate a target variable which is difficult, delayed, destructive, or expensive to measure directly, by using other process variables that are continuously available online [8,9,11]. In the pulp and paper industry, soft sensors are particularly valuable for properties such as Cobb value, tensile strength, burst strength, tear strength, ring crush strength, wet tensile, and other delayed laboratory quality parameters, where direct online measurement is either unavailable or economically impractical.

The typical paper-quality parameters that are most suitable for soft-sensor-based prediction and control applications are summarized in Table 2.

Table 2: Typical paper properties suitable for soft-sensor prediction in paper mills

Property / quality parameter	Direct online measurement availability	Soft-sensor suitability	Typical application
Basis weight	Available through QCS	High (for upstream prediction)	anticipatory control
Moisture	Available through QCS	Moderate	profile correction
Cobb value	Generally offline	Very high	packaging / converting
Tensile strength	Mostly offline	Very high	packaging grades
Burst strength	Offline	High	kraft / duplex board
Tear strength	Offline	High	printing & packaging
Ring crush strength	Offline	Very high	corrugation / liner
Wet tensile	Offline	Very high	specialty paper

In papermaking operations, soft sensors can be deployed either as advisory decision-support tools for machine operators or as inputs to automatic control and optimization layers, including supervisory control systems and advanced process control frameworks [4,12]. Their value is highest when the predicted property has strong operational or economic significance and when corrective action is still possible before large volumes of paper are produced outside specification limits.

Typical corrective actions supported by soft-sensor outputs may include stock-flow adjustment, headbox consistency correction, refining-energy optimization, wet-end chemical dosage control, size-press starch pickup adjustment, and drying-profile correction. This enables earlier intervention than conventional laboratory-based quality verification.

Soft-sensor design for paper properties is particularly challenging because of process drift, frequent grade changes, furnish variability, laboratory bias, missing signals, sensor noise, and long transport delays between causative inputs and measured outcomes [9,10,12]. Since paper properties develop progressively across the machine, the model must account for residence-time delays and evolving process history, which makes static models less reliable over long operating periods.

For this reason, recent research increasingly emphasizes adaptive, interpretable, and hybrid soft-sensor frameworks, combining machine-learning algorithms with process knowledge and physical constraints rather than relying solely on static black-box models [9,10,12]. Such systems are more robust during startup, grade transitions, and upset operating conditions, making them better suited for industrial deployment.

III. MACHINE-LEARNING METHODS USED IN PAPERMAKING QUALITY PREDICTION

3.1 Linear and multivariate regression models

Linear regression and multivariate approaches such as multiple linear regression (MLR) and partial least squares (PLS) regression continue to serve as important baseline methods for paper-quality prediction because of their simplicity, interpretability, computational efficiency, and ease of industrial implementation [13,14,22]. These methods are particularly valuable during the early stages of model development, where process engineers need clear insight into the magnitude, direction, and statistical significance of each predictor variable, enabling direct correlation of process variables with product quality outcomes.

In papermaking, regression-based models have been widely applied to relate fiber morphology, furnish composition, freeness (°SR / CSF), refining specific energy, filler loading, wet-end chemical dosage, and moisture history to critical quality properties such as tensile strength, tear strength, burst strength, basis weight, caliper, and moisture content [14,15,23]. Laboratory studies, pilot-machine investigations, and full-scale process simulations have shown that these models can provide meaningful explanatory and predictive capability, particularly when supported by feature engineering, variable transformation, interaction terms, and time-lag alignment.

Among these approaches, PLS regression is especially relevant for paper-mill applications, as process data from DCS, QCS, and laboratory systems often contain highly collinear variables such as stock consistency, machine speed, headbox flow, refining load, and moisture profile [22,24]. PLS is well suited to such datasets because it reduces dimensionality while preserving the covariance structure between predictor variables and response properties.

However, a key limitation of linear and multivariate regression models is their dependence on the assumption of approximately linear process–property relationships. In practical paper-machine environments, nonlinear interactions frequently arise due to chemical threshold effects, furnish variability, grade transitions, and dynamic process disturbances, which may lead to underfitting and reduced robustness under changing operating conditions [23,24].

For this reason, regression models are best regarded as baseline interpretative tools and benchmark models, against which more advanced nonlinear methods such as SVR, ensemble learning, and hybrid soft sensors should be critically evaluated.

A major advantage of linear and multivariate regression models is their high interpretability and process transparency, which allows process engineers to directly understand how variables such as refining intensity, long-fiber ratio, starch dosage, filler loading, ash content, and moisture history influence the final property response [22,23]. In industrial papermaking environments, this interpretability is particularly valuable because it supports root-cause analysis, process troubleshooting, and operator confidence, all of which are essential for real-time decision-making and quality control.

For example, the regression coefficients can be directly interpreted to assess whether an increase in refining specific energy is positively associated with tensile strength or whether higher ash content is negatively affecting burst and ring-crush performance. Such direct interpretability makes these models especially attractive in mills where model transparency, explainability, and operational trust are more important than marginal gains in predictive accuracy.

However, this advantage must be critically balanced against their limitations. In large-scale paper-machine datasets, process variables are often highly correlated, dynamically delayed, and nonlinear, which may reduce the robustness of ordinary linear models under grade changes and transient operating conditions [23,24]. Consequently, while these models remain valuable as baseline and benchmark tools, they may not fully capture complex interactions between furnish characteristics, wet-end chemistry, and dynamic machine behavior.

parative technical assessment of the major machine-learning approaches used for paper-quality prediction is presented in Table 3.

Table 3: Comparative assessment of machine-learning methods used in paper-quality prediction

Model / Method	Typical data type	Key parameters / architecture	Major strengths	Main limitations	Best papermaking applications
Linear regression / MLR / PLS	tabular, low to moderate dimensional process data	regression coefficients, latent variables (PLS components)	highly interpretable, fast, benchmark model, easy industrial acceptance	underfits nonlinear process behavior, weak under drift / grade changes	tensile, tear, burst, moisture baseline models
SVR (RBF kernel)	moderate-size tabular + nonlinear data	C, ε, γ, kernel type	strong nonlinear modeling, robust with moderate datasets, good generalization	low interpretability, retuning needed, computationally slower	wet tensile, moisture, delayed lab-property prediction
Random Forest / GBM / XGBoost / LightGBM	historian, DCS, QCS, lab tabular data	estimators, tree depth, learning rate, subsampling	strong tabular-data accuracy, handles interactions and missing values, feature importance ranking	drift vulnerability, limited extrapolation, moderate interpretability	Cobb, strength indices, GSM stability, quality fault detection
ANN / FNN	nonlinear multivariate data	hidden layers, neurons, activation functions	strong nonlinear mapping capability	black-box, hyperparameter sensitive	strength and moisture prediction
RNN / LSTM / GRU	high-frequency time-series data	sequence length, hidden states, recurrent layers	captures delay, transport history, autocorrelation	complex training, high data requirement	dynamic GSM, moisture profile, startup prediction
Hybrid first-principles + ML	physics + data integrated systems	physical constraints + ML architecture	physically realistic, robust during transitions, strong industrial reliability	complex implementation, requires process expertise	GSM, moisture, Cobb, control-system integration

As evident from Table 3, the choice of modeling approach should be guided not only by predictive accuracy but also by interpretability, data availability, robustness to process drift, and long-term maintainability in mill conditions.

However, the limitations of linear and multivariate regression models become considerably more evident in full-scale paper-machine environments, where the process is characterized by strong variable interactions, nonlinear chemical effects, threshold behavior, residence-time delays, and time-varying machine dynamics [23,24]. Papermaking is inherently a multivariate nonlinear process, in which furnish variability, refining intensity, retention chemistry, drainage efficiency, drying history, machine speed, and ambient environmental conditions interact simultaneously and often in a delayed manner.

For instance, the influence of wet-end chemical dosage on Cobb value or strength indices may not follow a simple linear trend but may exhibit critical threshold effects and saturation behavior, particularly in systems involving AKD / ASA sizing, starch addition, and filler–fiber interactions. Similarly, the effect of refining specific energy on tensile and burst strength is often nonlinear, with an optimum operating window beyond which further refining may lead to fiber cutting and deterioration of mechanical properties [15,23].

Under such conditions, linear models may remain computationally stable and operationally interpretable, but they frequently underfit the true process response, especially when the target property is strongly influenced by grade changes, furnish shifts, startup transients, regime transitions, and nonlinear responses to chemistry or refining conditions. This limitation becomes particularly significant during dynamic operating conditions, where real-time quality control requires rapid adaptation to evolving process states.

Therefore, regression models are best regarded as baseline benchmark models and interpretative reference tools, against which more advanced machine-learning approaches such as support vector regression, ensemble learning, neural networks, and hybrid soft sensors should be critically compared.

3.2 Support vector regression and kernel methods

Support vector regression (SVR) has been widely adopted in industrial soft-sensor applications because of its ability to effectively capture nonlinear process–property relationships while maintaining strong generalization performance, particularly when the available dataset is of moderate size and high dimensionality [9,11,25]. Unlike conventional linear regression, SVR employs kernel functions to transform input variables into a higher-dimensional feature space, enabling the model to represent complex nonlinear interactions that are common in papermaking processes.

Among the available kernels, the radial basis function (RBF) kernel is most frequently used in process-industry applications because it provides flexibility in modeling nonlinear responses without requiring explicit prior knowledge of the underlying functional form [25,26]. In practical implementations, model performance is strongly influenced by the selection of key hyperparameters, including the regularization parameter (C), insensitive loss margin (ε), and kernel width parameter (γ). These parameters govern the trade-off between model complexity, fitting accuracy, and robustness to noise, making hyperparameter optimization a critical step in industrial deployment.

In papermaking applications, SVR has been reported as an effective method for predicting wet tensile strength, basis-weight variability, moisture profile, and other delayed laboratory quality parameters, often outperforming simpler linear models when nonlinear interactions are significant [9,11,26]. This makes SVR particularly suitable for properties influenced by interactions among furnish composition, refining energy, wet-end chemistry, machine speed, drying profile, and environmental conditions.

A major strength of SVR lies in its strong performance in high-dimensional process spaces, where numerous correlated variables from DCS, QCS, historian, and laboratory systems are simultaneously involved. This is highly relevant in modern paper mills, where the process historian may include hundreds of tags related to stock preparation, headbox control, retention chemistry, drying, and reel properties. In addition, SVR exhibits relatively good resistance to overfitting for moderate-sized datasets, provided that kernel type and hyperparameters are carefully optimized through grid search, cross-validation, or Bayesian optimization [25,27].

However, compared with linear and multivariate regression models, SVR is inherently less interpretable, which may reduce operator trust and process-engineer acceptance in industrial environments. The black-box nature of kernel transformations makes it difficult to directly quantify how specific variables such as starch dosage, ash content, or refining load contribute to the predicted quality response.

Furthermore, SVR performance may degrade under process drift, grade transitions, furnish variability, and seasonal changes, where the training data distribution no longer represents the current operating regime

[26,27]. In mills with frequent grade changes and variable recycled-fiber furnish, this may necessitate periodic retuning and retraining, and the associated operational burden must be carefully balanced against predictive gains.

Therefore, SVR is often most effective when applied to stable grade families, or when incorporated into adaptive and drift-aware soft-sensor frameworks.

3.3 Tree-based ensemble models

Tree-based ensemble methods such as Random Forests (RF), Gradient Boosting Machines (GBM), XGBoost, LightGBM, and related decision-tree ensembles have become increasingly popular for paper-quality prediction because they effectively handle nonlinear relationships, higher-order variable interactions, missing values, outliers, and mixed data types, often with relatively limited preprocessing requirements [9,10,28].

These methods are particularly well suited to mill historian, DCS, QCS, and laboratory datasets, where the data are predominantly tabular and frequently contain a combination of continuous process variables, categorical grade information, batch identifiers, and recipe metadata. This is highly relevant in papermaking environments, where datasets typically combine process-history variables with delayed laboratory results.

In practical papermaking applications, tree-based ensemble models have shown strong predictive performance for Cobb value, tensile strength, burst strength, moisture variability, basis-weight stability, and process-fault classification, particularly where complex nonlinear interactions exist among chemical dosage, refining energy, filler loading, ash content, drying history, and moisture profile [28,29].

A major advantage of these models is their ability to generate feature-importance rankings and split-based variable contribution analysis, which significantly improves process understanding. For example, the relative influence of AKD / ASA dosage, starch pickup, ash %, refining intensity, long-fiber ratio, machine speed, and dryer temperature profile can be directly ranked, helping process engineers identify the dominant drivers of quality variation.

Among the ensemble approaches, XGBoost and LightGBM are particularly attractive because they employ gradient-boosted decision trees with regularization, shrinkage, and optimized split-learning strategies, often providing superior performance compared with standard Random Forest models on large industrial datasets [29,30]. Important hyperparameters such as:

- ✓ number of estimators
- ✓ tree depth
- ✓ learning rate
- ✓ minimum child weight
- ✓ subsampling ratio

strongly influence prediction robustness and should be optimized through cross-validation.

Tree-based models are especially effective when the process exhibits nonlinear threshold effects, such as sudden Cobb-value failure due to sizing instability or abrupt strength loss beyond a critical refining-energy window. Such threshold behavior is often difficult for linear models to capture.

However, the principal limitations include reduced physical interpretability, vulnerability to concept drift, and limited extrapolation capability under unseen operating conditions [28–30]. If a mill introduces a new furnish blend, recycled-fiber ratio, chemical recipe, or operating window that is not represented in the training dataset, these models may produce high-confidence but physically unreliable predictions.

For this reason, industrial deployment of tree-based ensemble models must be supported by continuous performance monitoring, drift detection, retraining triggers, and model-governance protocols, particularly in mills with frequent grade changes and variable raw-material quality.

From a critical review perspective, ensemble models currently represent one of the most practical and high-performing approaches for industrial soft-sensor deployment in papermaking, provided that lifecycle maintenance is properly addressed.

3.4 Neural networks and deep learning

Artificial neural networks (ANNs) have a long and well-established history in industrial process modeling, nonlinear regression, and soft-sensor development, and recent deep-learning frameworks have significantly extended this capability to multivariate time-series data and high-frequency process signals [9,10,31]. These methods are particularly attractive in papermaking environments where multiple sensor streams from DCS, QCS, drives, machine-vision systems, and laboratory platforms are strongly correlated and exhibit pronounced nonlinear dynamic behavior.

Conventional feedforward neural networks (FNNs) are well suited for modeling static nonlinear relationships between process variables and target paper properties, such as basis weight, tensile strength, burst index, and moisture content. However, because many paper properties are strongly influenced by process history and transport delay, dynamic architectures such as recurrent neural networks (RNNs), long short-term memory (LSTM) networks, and gated recurrent units (GRUs) are often more appropriate [31,32].

These architectures are specifically designed to capture time-dependent correlations, lagged process effects, and sequential dependencies, making them highly relevant to paper-machine operations, where the final sheet properties are influenced by upstream conditions in stock preparation, headbox consistency, drainage, pressing, drying, and reel formation.

For example, LSTM and GRU architectures are particularly effective in modeling delayed quality responses because they can retain long-range temporal dependencies and overcome the vanishing-gradient limitations associated with conventional RNNs [32,33]. This is directly applicable to papermaking data, where basis weight, moisture profile, Cobb-related variables, and strength properties depend not only on current machine conditions but also on historical operating windows and material residence time.

Recent soft-sensor research beyond papermaking has further demonstrated the potential of teacher–student learning, representation learning, feature embedding, and sequence-learning frameworks for extracting meaningful latent features from complex sensor time series [10,31]. Similar concepts can be effectively transferred to paper-machine datasets, which are typically high-frequency, autocorrelated, and delay-affected, especially during startup, grade changes, furnish transitions, and transient operating regimes.

Neural-network-based models may outperform simpler regression and tree-based approaches when data volume is sufficiently large, sensor coverage is extensive, and deployment infrastructure is mature. They are particularly valuable for mills equipped with advanced historian systems, continuous data acquisition, and high-frequency sensor integration.

However, these models are also associated with significant industrial limitations. Their black-box nature, higher computational demand, hyperparameter sensitivity, and reduced interpretability make them more difficult to validate, troubleshoot, and maintain in production environments [31–33]. Unexpected predictions may be difficult to trace back to specific process drivers, which can reduce operator trust and delay industrial adoption.

For many mills, therefore, the most suitable model is not necessarily the most sophisticated architecture, but rather the one that remains stable, explainable, computationally efficient, and maintainable over long-term production conditions.

From a critical industrial perspective, deep-learning models are most justified in high-data, high-automation mill environments, whereas simpler and more interpretable models may remain preferable for conventional production settings.

3.5 Hybrid first-principles plus machine-learning models

Hybrid models combine first-principles process understanding with the flexibility and adaptive capability of data-driven machine-learning approaches, thereby integrating the complementary strengths of both methodologies [11–13,34]. In papermaking, this may involve embedding mass-balance equations, stock-flow relationships, section-level process equations, residence-time logic, heat- and moisture-transfer correlations, and physically constrained variable dependencies within a machine-learning framework.

For example, basis-weight prediction models can integrate headbox flow rate, stock consistency, machine speed, and mass-balance relationships with machine-learning algorithms to improve predictive robustness. Similarly, moisture and Cobb-related prediction frameworks may incorporate dryer-section heat-transfer logic, evaporation kinetics, sheet density evolution, and size-press chemistry constraints to preserve physical realism [34,35].

A key strength of hybrid modeling is that it explicitly constrains the prediction space using known physical laws. This is particularly valuable in paper-machine environments, where certain variables are bounded by process physics. For instance, basis weight must satisfy mass conservation across stock flow and machine speed, while moisture profile must remain consistent with section-wise heat and mass transfer mechanisms.

A recent hybrid soft-sensor study for basis-weight prediction in the wire section demonstrated that dynamically weighted hybrid models can outperform both pure first-principles models and standalone machine-learning approaches, particularly during startup conditions, grade transitions, and transient operating regimes [12,35]. This finding is especially important because such operating windows are precisely where conventional static models often fail.

This approach is particularly relevant to the pulp and paper industry because the process is fundamentally physics-driven yet highly data intensive. Sheet formation, drainage, consolidation, pressing, drying, and strength development are governed by well-established physical relationships, while simultaneously large volumes of real-time data are continuously generated from DCS, QCS, historian, and laboratory systems.

Hybridization significantly improves model interpretability, physical consistency, and operator trust, while reducing the risk of physically unrealistic predictions, such as impossible basis-weight shifts, non-feasible moisture values, or chemically implausible Cobb responses. This becomes a major practical advantage when models are used for real-time advisory control, supervisory optimization, or semi-closed-loop decision-making, where prediction reliability directly influences operational safety and product quality.

From a critical review perspective, hybrid frameworks currently represent one of the most promising directions for industrial-scale paper-property prediction, because they offer a practical balance between predictive performance and physical plausibility [34–36].

As summarized in Table 3 and Figure 5, hybrid models are increasingly emerging as one of the most robust approaches for industrial deployment.

3.6 Feature engineering and target alignment

Feature engineering is often the most decisive step in paper-property prediction, and in many industrial applications it may have a greater impact on model performance than the choice of the machine-learning algorithm itself [9,10,12,37]. In papermaking, raw process signals obtained directly from the machine are rarely sufficient on their own because most paper properties develop as a consequence of cumulative process history rather than instantaneous measurements.

For this reason, robust predictive models typically rely on engineered process-history features, including:

- ✓ lagged variables
- ✓ moving averages
- ✓ rolling standard deviations
- ✓ exponentially weighted trends
- ✓ section-level summary statistics
- ✓ variability descriptors
- ✓ operator-event flags
- ✓ grade identifiers
- ✓ recipe descriptors

That accurately represent the actual development pathway of the sheet as it progresses through stock preparation, formation, pressing, drying, size press, and reel-up.

In paper-machine environments, residence-time-aware lag features are particularly critical because upstream process changes may influence the final product quality only after a well-defined transport delay and material residence time [37,38]. For example, a change in refining energy or starch dosage may affect tensile strength only after the sheet has passed through formation, consolidation, and drying sections. Similarly, a disturbance in size-press pickup may influence Cobb response after a measurable downstream delay.

Accordingly, features based on time windows, delayed signals, rolling trends, transient-response indicators, and section-wise transport logic often significantly improve predictive performance.

For strength prediction, the most influential engineered variables commonly include fiber type, pulp blend ratio, long-fiber percentage, refining specific energy, filler loading, starch dosage, wet-end charge balance, moisture profile, sheet density, caliper, and ash content [5,15,38].

For Cobb prediction, the feature space should capture both process and material effects, including internal sizing chemistry (AKD / ASA / rosin), surface starch treatment, size-press pickup, drying temperature profile, sheet density, ash %, porosity proxies, surface roughness, and moisture interaction terms [16,17,39].

For basis-weight (GSM) prediction, upstream process-history features such as stock consistency, fan-pump flow rate, headbox pressure, slice opening, formation index, drainage efficiency, wire vacuum, and section-level dynamic summaries are particularly critical [3,12,40].

Equally important is the treatment of the target variable itself. Laboratory values may represent:

- ✓ reel average
- ✓ localized cross-machine sample
- ✓ batch-wise test result
- ✓ delayed quality verification

And therefore, must be carefully mapped to the corresponding production history.

If the process data are not correctly aligned with the actual sample location, reel position, and production timestamp, the model may learn spurious correlations and noise rather than true process causality.

Therefore, time alignment, reel-position tracking, batch mapping, and residence-time synchronization are indispensable requirements for reliable supervised learning in paper mills.

From a critical industrial perspective, poor target alignment remains one of the most common reasons for soft-sensor failure despite the use of advanced algorithms.

IV. PREDICTION OF SPECIFIC PAPER PROPERTIES

4.1 GSM or basis weight prediction

Basis weight, or grammage (GSM), has historically been one of the most extensively instrumented quality variables in papermaking, primarily through online Quality Control Systems (QCS) installed near the reel, pope reel, or finishing section [3,4,41]. Despite the widespread use of online scanners, soft-sensor-based predictive models remain highly valuable for early-stage prediction, redundancy, upstream section-level estimation, and anticipatory control before the final measurement point.

This distinction is important because conventional QCS scanners provide measurement only at a downstream location, whereas process disturbances originating in stock consistency, fan-pump flow, headbox pressure, slice opening, drainage behavior, or machine speed may begin affecting the sheet much earlier in the machine. Consequently, by the time the final scanner detects a deviation, a significant volume of off-spec paper may already have been produced.

Recent evidence from hybrid and dynamic soft-sensor studies in tissue and paper-machine environments indicates that the integration of mass-balance principles, stock-flow equations, residence-time logic, and machine-learning algorithms can significantly improve predictive performance during transient conditions, startup phases, furnish transitions, and regime shifts, compared with standalone data-driven models [12,35,41].

This is particularly important because startup and grade-change phases are often associated with the highest GSM variability and maximum reel losses.

These findings strongly suggest that future GSM prediction systems should not be viewed merely as sensor replacements, but rather as anticipatory control tools capable of identifying deviations earlier in the process and enabling faster corrective action.

From a process-control standpoint, real-time GSM prediction can support stock-flow optimization, headbox consistency adjustment, slice opening control, startup stabilization, grade-change management, and reduction of off-spec production [41,42].

From an economic perspective, even small deviations in GSM can directly translate into fiber overconsumption, higher production cost, reel rejection, and customer claims, making predictive control particularly attractive for high-volume packaging and writing-printing grades.

This is especially relevant for Indian paper mills, where QCS coverage may be uneven across machine sections, older machines may lack advanced instrumentation, and furnish variability is often high due to recycled-fiber usage and mixed waste-paper input [42,43].

In such environments, predictive GSM models may offer a cost-effective route to enhanced process stability, lower fiber loss, improved grammage consistency, and reduced startup waste, provided that the models are carefully calibrated to machine dynamics, furnish shifts, and section-wise residence-time effects.

From a critical industrial perspective, GSM prediction currently represents one of the most mature and immediately deployable use cases for machine-learning-based quality control in papermaking.

4.2 Cobb prediction

Compared with basis-weight prediction, the published evidence on direct machine-learning-based prediction of Cobb value in full-scale paper production remains comparatively limited, and this itself represents an important finding of the present review [16,17,44]. While basis weight, moisture, and caliper are routinely monitored online through QCS systems, Cobb value is still predominantly treated as an offline laboratory test parameter rather than a continuously controlled process variable.

This gap presents a significant opportunity for both research advancement and industrial deployment, particularly in kraft paper, duplex board, linerboard, corrugating medium, and other converting-sensitive packaging grades, where water absorptiveness directly influences adhesive bonding, printability, dimensional stability, storage performance, and box integrity under humid conditions.

From a commercial standpoint, uncontrolled Cobb variation may lead to poor glue adhesion, carton warping, dimensional distortion, reduced compression performance, reduced stiffness, and increased customer complaints, making it one of the most economically important yet comparatively under-digitized quality parameters in packaging mills [44,45].

A practical Cobb soft-sensor framework must integrate multiple process and material variables [5,16,17,45], including wet-end sizing chemistry, surface sizing or size-press starch pickup, moisture profile, drying history, ash or filler loading, sheet density, surface treatment data, and porosity-related proxies.

In particular, the dosage, retention efficiency, and cure behavior of AKD, ASA, and rosin-based sizing systems, together with starch pickup at the size press, are likely to be among the most influential predictors.

Because Cobb value is a standardized response variable measured through laboratory absorption testing rather than a direct physical sensor output, robust laboratory standardization, sample-position consistency, and strict sampling discipline are essential prerequisites before reliable predictive models can be developed.

This aspect is more critical for Cobb than for many other quality variables.

If laboratory sampling is inconsistent with reel position, machine direction, cross-machine location, or time delay, the model may learn spurious correlations rather than true process–property causality.

Therefore, target alignment, batch tracking, reel-position mapping, and laboratory repeatability control are especially important for successful Cobb prediction.

From a critical review perspective, the relative scarcity of industrial-scale literature on direct Cobb prediction identifies this area as one of the most promising research gaps for future machine-learning applications in paper quality control [44–46].

As discussed earlier in Figure 3, the major factors affecting Cobb development across the machine must be explicitly incorporated during feature engineering and model deployment.

4.3 Strength prediction

Strength-property prediction currently represents the most mature, industrially visible, and practically validated area of machine-learning adoption in paper quality control, particularly for packaging papers, paperboards, kraft grades, and specialty products [4,5,15,47].

Compared with Cobb and other specialty parameters, the literature on strength prediction is considerably more developed, reflecting both the commercial importance of strength indices and the relatively strong correlation between measurable process variables and final mechanical performance.

Industrial case studies have demonstrated successful real-time prediction of tensile strength, tear strength, burst strength, ring crush strength, and wet tensile, where soft sensors are directly integrated into operational decision-support and process-optimization systems.

In such frameworks, predictive models estimate the strength response and use these outputs to guide process decisions related to [47,48]:

- ✓ long-fiber ratio
- ✓ filler loading
- ✓ refining specific energy
- ✓ starch dosage
- ✓ wet-end resin chemistry
- ✓ furnish blend optimization

A reported industrial application for specialty paper achieved wet-tensile prediction with an R^2 of 0.94, together with approximately 15% reduction in chemical dosage and nearly 80% reduction in wet-tensile variability over a 90-day operating period [5].
These findings are particularly significant because they move beyond simple offline model validation and demonstrate measurable operational, economic, and quality-control value under real production conditions. This distinction is important from a review perspective because many published ML studies remain limited to offline datasets, whereas this example demonstrates successful industrial deployment with quantifiable process benefit.
The evidence strongly indicates that when machine-learning predictions are linked to a meaningful manipulated process variable, such as refining energy, starch dosage, resin addition, or furnish blend ratio, the system can simultaneously support [47-49]:

- ✓ quality assurance
- ✓ process stabilization
- ✓ chemical optimization
- ✓ cost reduction

From a paper-science perspective, strength prediction is particularly well suited for data-driven modeling because mechanical properties are strongly governed by fiber morphology, fiber length distribution, bonding potential, refining intensity, sheet density, moisture profile, and wet-end chemistry [15,48].
Laboratory and pilot-scale studies linking fiber properties, furnish descriptors, and refining variables to tensile and tear outcomes further reinforce the conclusion that strength indices are among the most predictable paper properties when appropriate structural, material, and process-history features are included.
From a critical review standpoint, strength prediction currently represents the strongest evidence base for industrial-scale machine-learning deployment in papermaking, and therefore serves as an important benchmark for extending similar approaches to Cobb, moisture, and convertibility-related properties.
As discussed earlier in Figure 4, the physical mechanism of strength development should always be explicitly considered during feature engineering and model interpretation.

V. MODEL DEPLOYMENT AND INDUSTRIAL INTEGRATION

Model evaluation: what matters in practice
Academic modeling studies often emphasize statistical performance metrics such as coefficient of determination (R^2), 5.1. root mean square error (RMSE), and mean absolute error (MAE), which are essential for assessing predictive accuracy [9,10]. However, for actual deployment in paper mills, these metrics alone are necessary but not sufficient.
In a production environment, a useful predictive model must also demonstrate robustness under real operating conditions, including grade shifts, startup and shutdown transients, process drift, furnish variability, sensor noise, missing data, and delayed laboratory updates.
This is particularly important in papermaking, where operating conditions may change significantly due to recycled-fiber variation, wet-end chemistry adjustments, speed changes, and machine disturbances.
A practically useful model should therefore provide stable predictions near specification limits, avoid physically implausible excursions, and remain sufficiently interpretable for operators, process engineers, and quality-control personnel to trust and act upon the predictions. For example, a model predicting impossible strength spikes, unrealistic Cobb shifts, or sudden grammage changes without corresponding process disturbance will rapidly lose operational credibility.
The key practical criteria for evaluating machine-learning models in paper-mill deployment are summarized in Table 4.

Table 4: Practical evaluation criteria for machine-learning models in paper mills

Criterion	Practical significance
Accuracy	fit to lab / QCS results
Stability	robustness during grade changes
Interpretability	operator confidence
Latency	suitability for real-time control
Maintenance	retraining frequency
Economic value	reduction in waste / chemicals
Physical plausibility	realistic predictions

A robust industrial evaluation framework for paper-property prediction models should therefore include the following criteria:

- ✓ predictive accuracy (R^2, RMSE, MAE)
- ✓ stability under changing operating regimes
- ✓ interpretability and operator trust
- ✓ prediction latency / real-time usability
- ✓ retraining burden and maintenance frequency
- ✓ economic impact
- ✓ control usefulness
- ✓ physical plausibility of predictions

In many paper mills, a model that is slightly less accurate but more transparent, stable, and easier to maintain may deliver substantially greater long-term value than a black-box architecture with only marginally better statistical fit.

From an industrial deployment perspective, model reliability and operational usability often outweigh small gains in numerical accuracy.

5.2 Interpretability, trust, and model governance

Interpretability is not a secondary consideration in industrial quality control, but a critical requirement for successful deployment of machine-learning models in paper mills [9,10,12]. Operators, process engineers, and production managers are significantly more likely to trust and use a predictive model when they can clearly understand which process drivers are influencing the prediction and whether the model is operating within a familiar production window.

This is particularly important in papermaking, where process decisions involving refining energy, furnish blend ratio, wet-end chemical dosage, size-press starch pickup, and machine speed directly affect production cost and product quality.

Recent soft-sensor studies strongly emphasize the need for interpretable machine learning, stability-aware design, and transparent decision-support frameworks, especially in high-stakes continuous process industries [9,10].

For paper-property prediction, interpretability can be significantly improved through:

- ✓ variable-importance analysis
- ✓ sensitivity plots
- ✓ partial dependence analysis
- ✓ hybrid first-principles models
- ✓ grade-specific models
- ✓ operator dashboards displaying prediction drivers

Such dashboards should ideally display not only the predicted property value but also the dominant causal contributors, such as ash level, refining energy, starch dosage, moisture variation, or furnish composition.

Model governance is equally important and should include:

- ✓ model version control
- ✓ data-quality checks
- ✓ retraining triggers
- ✓ drift detection
- ✓ laboratory bias monitoring
- ✓ sample-position alignment checks
- ✓ approval rules for control intervention

There must also be clear operating rules defining whether the model is used purely as an advisory tool, as a supervisory optimization layer, or is permitted to influence automatic control actions.

Without this governance layer, even technically strong models may fail in production due to operator mistrust, silent model degradation, process drift, or conflict between automated recommendations and operator judgment.

In industrial paper mills, long-term success depends as much on human trust and governance discipline as on predictive accuracy. The governance framework should be integrated with the evaluation criteria summarized in Table 4.

5.3 Integration with real-time quality control and optimization

Machine-learning models create real industrial value only when they are integrated into operational decision-making and process-control workflows, rather than remaining isolated analytical tools [4,12].

In papermaking, three principal deployment levels can be identified.

1. Descriptive and advisory level

At the first level, the soft sensor functions as an online advisory tool, providing real-time estimates, deviation alarms, and early warning signals while leaving all corrective action entirely to the machine operator or process engineer.

Typical examples include online prediction of basis-weight drift, Cobb deviation risk, tensile reduction, or abnormal moisture trends, supported by operator dashboards and alarm systems.

2. Supervisory optimization level

At the second level, model outputs are integrated into supervisory optimization systems, where the predicted

quality response is used to recommend process set-point adjustments.
These may include recommendations for:
- ✓ stock-flow correction
- ✓ headbox consistency adjustment
- ✓ refining specific energy
- ✓ furnish blend ratio
- ✓ wet-end chemical dosage
- ✓ size-press starch pickup
- ✓ dryer steam profile

This level is often the most practical and industrially preferred starting point.

3. Semi-automatic / closed-loop level
At the third level, predictive outputs are permitted to influence automatic control logic under guarded operating constraints, enabling semi-automatic or closed-loop control actions.
Examples include automatic correction of stock flow, refining load, chemical dosing, or moisture profile control, subject to predefined safety and operating windows.
The three-stage deployment framework for machine-learning integration in paper-mill quality control is illustrated in Figure 6.

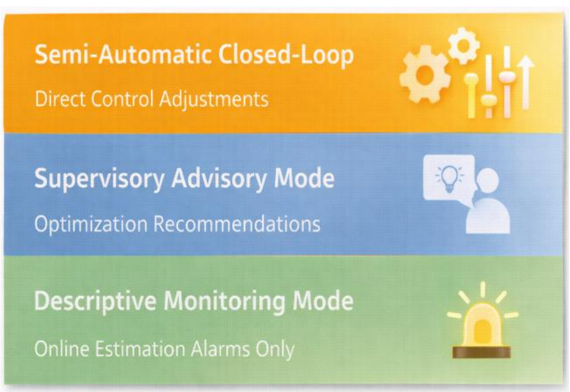

Figure 6: Three-level deployment framework for machine-learning integration in real-time paper quality control and optimization.

Industrial experience strongly suggests that the supervisory level is often the best initial deployment pathway.
Advisory and supervisory applications allow mills to:
- ✓ validate model stability
- ✓ quantify economic benefit
- ✓ assess reduction in waste / chemicals
- ✓ build operator confidence
- ✓ establish governance discipline

Before Progressing toward tighter automation.
This staged deployment strategy is particularly important for properties such as Cobb value and strength indices, where laboratory verification remains essential and the tolerance for customer-quality risk is often low.
For most mills, especially those in transitional digitalization phases, gradual movement from advisory → supervisory → guarded closed-loop control is the most reliable implementation roadmap.

VI. CHALLENGES, LIMITATIONS, AND INDUSTRIAL RELEVANCE

Challenges and limitations
Despite the clear potential of machine-learning-based quality prediction, several practical obstacles still limit its widespread deployment in the pulp and paper industry [9,10,12].

1. Data quality and process alignment
Data quality remains the single largest challenge.
In real mill environments, historian tags may contain noise, signal dropout, calibration drift, missing values, and timestamp inconsistencies. Laboratory test results may be affected by sampling variability, operator bias, repeatability issues, and delayed reporting.
In addition, grade changes, furnish transitions, and recipe modifications are not always recorded with sufficient accuracy.
If these events are poorly captured, the resulting dataset may not represent the actual process conditions, leading to degraded model reliability.
For mills producing multiple grades, the process-data distribution may shift faster than the model retraining cycle, resulting in concept drift and performance degradation.

2. Transferability across mills and machines
Another major limitation is model transferability.
A model developed on one paper machine, furnish mix, or grade family may not generalize reliably to another machine without substantial redevelopment and recalibration.
This is particularly true in the Indian context, where mills vary widely in:
- ✓ raw-material mix

- ✓ recycled-fiber percentage
- ✓ automation maturity
- ✓ water chemistry
- ✓ process instrumentation
- ✓ maintenance discipline

Because these factors strongly influence paper properties, cross-mill model portability remains a major challenge.

3. Gap between academic studies and plant deployment

There is also a significant methodological gap between academic modeling success and plant-scale deployment.

Many published studies focus primarily on statistical fit measures such as R^2, RMSE, and MAE, while providing limited information on:

- ✓ long-term model stability
- ✓ maintenance frequency
- ✓ failure handling
- ✓ drift management
- ✓ retraining strategy
- ✓ operator adoption
- ✓ control-system integration

Future research must pay greater attention to the full model lifecycle, including commissioning, maintenance, retraining, governance, and operator change management, rather than focusing only on one-time predictive performance.

6.2 Relevance for Indian mills

Indian paper mills are under increasing pressure to improve quality consistency, reduce process waste, manage raw-material variability, and compete in higher-value packaging, board, and specialty paper grades [4,5,7,50].

These pressures make machine-learning-based quality prediction and soft-sensor deployment highly relevant, particularly in mills where furnish heterogeneity, recycled-fiber variability, inconsistent waste-paper quality, and rising energy and chemical costs contribute significantly to process instability.

This is especially important in the Indian context, where a large proportion of mills operate with diverse and variable raw-material streams, including:

- ✓ agro residues
- ✓ recycled fiber
- ✓ mixed waste paper
- ✓ imported pulp
- ✓ recovered packaging waste

all of which introduce substantial variability in fiber morphology, drainage behavior, sizing response, ash content, and final paper properties [50,51].

This raw-material variability is generally more pronounced in Indian packaging and board mills compared with virgin-pulp-dominant systems, making predictive quality control particularly valuable.

At the same time, deployment strategies for Indian mills must remain practical, phased, and economically driven rather than technology driven.

Many mills may not yet have complete sensor coverage or seamless integration between laboratory information systems, DCS, QCS, and historian platforms, particularly in older paper machines and retrofitted board lines [51,52].

Older machines may also operate with limited instrumentation, partial automation, and inconsistent digital data capture, which can restrict immediate implementation of advanced predictive systems.

For such environments, the most effective pathway is typically phased implementation, beginning with one commercially high-value target property and gradually scaling after demonstrating measurable operational benefit.

A practical roadmap may include:

a) select one commercially critical property
b) improve data discipline and laboratory consistency
c) deploy as advisory soft sensor
d) validate economic benefit
e) expand to multi-property prediction
f) integrate supervisory optimization

For packaging and board grades in particular, priority targets should include tensile strength, burst strength, ring crush strength, Cobb value, GSM stability, and moisture-related convertibility indicators, as these properties have direct commercial implications and are strongly influenced by controllable process variables [52,53].

In Indian mills, successful deployment can potentially deliver significant benefits through fiber savings, reduced chemical consumption, lower off-spec production, improved customer consistency, and stronger competitiveness in export-oriented packaging markets.

From a critical review perspective, Indian mills represent one of the most promising environments for high-impact industrial deployment of soft sensors, precisely because process variability and cost sensitivity are high.

6.3 Research gaps and future directions
Several research gaps are especially important. First, more peer-reviewed work is needed on direct prediction of Cobb value and related surface-absorption properties in production environments. Second, future studies should compare static, adaptive, and hybrid models across real grade changes and furnish shifts rather than only on stationary historical datasets. Third, the field needs stronger reporting of deployment outcomes, including operator acceptance, economic savings, and model-maintenance workload.
There is also growing opportunity in multimodal and sequence modeling. Modern paper machines generate time-series, profile, image, and event data that could be fused to improve sensitivity to formation changes, streaks, moisture-profile disturbances, and surface-treatment variability. Another promising direction is digital-twin and hybrid-soft-sensor integration, where physical process constraints guide machine-learning predictions and improve resilience during startup, upset, or previously unseen operating windows.
Finally, future work should address standardization. Without consistent data definitions, lab protocols, and benchmarking metrics, it will remain difficult to compare model performance across mills and grades. This makes machine-learning-based quality prediction particularly relevant for modernization and digital transformation initiatives in Indian paper mills.

VII. PRACTICAL ROADMAP FOR IMPLEMENTATION

For paper mills planning to adopt machine-learning-based quality prediction and soft-sensor deployment, a phased and implementation-focused roadmap is strongly advisable [4,9,12].
Rather than attempting large-scale multi-property deployment at the outset, mills should begin with a focused pilot use case and scale progressively after demonstrating technical and economic value.
The recommended staged implementation roadmap for machine-learning deployment in paper mills is illustrated in Figure 7.

Figure 7: Practical staged roadmap for machine-learning-based quality prediction deployment in paper mills.

The following staged roadmap is recommended:
a. Select a high-value target property
Begin with a property that has clear commercial and operational significance and for which sufficient historical data are available.
Typical starting points include:
- tensile strength
- wet tensile
- GSM stability
- Cobb value
- grade-specific strength index

The selected target should ideally have strong linkage with controllable process variables.

b. Audit data readiness
Before model development, conduct a comprehensive data-readiness audit.
This should include verification of:
- historian tag quality
- laboratory repeatability
- timestamp alignment
- grade coding
- missing-data patterns
- sample-position consistency
- transport-delay logic
- reel tracking

This step is often more critical than algorithm selection.

c. Build a baseline model first

Start with interpretable baseline models, such as linear regression, PLS, or tree-based models.

Only after establishing stable baseline performance should the mill consider deep learning or hybrid architectures.

This ensures easier troubleshooting and operator acceptance.

d. Engineer process-history features

Develop features that represent the true process history of the sheet, rather than relying solely on instantaneous signals.

This should include:

- lag variables
- rolling statistics
- residence-time-aware summaries
- recipe metadata
- operator-event markers
- transient condition flags

e. Validate under real operating regimes

Validation must extend beyond random train–test data splits.

Models should be tested across:

- grade transitions
- startup conditions
- shutdown events
- furnish changes
- disturbances
- seasonal variations

This is essential for real industrial robustness.

f. Deploy as advisory or supervisory first

Initial deployment should be limited to advisory or supervisory decision-support mode.

This helps build operator trust, quantify benefits, and validate stability before any automated control intervention.

g. Establish governance framework

A clear governance structure must be defined, including:

- retraining triggers
- drift detection rules
- performance dashboards
- laboratory audit procedures
- operator responsibilities
- model version control

h. Scale only after proving value

Only after the first use case demonstrates stable technical and economic benefit should deployment be expanded to additional properties such as:

- Cobb
- burst strength
- ring crush
- moisture optimization
- multi-property predictive control

This phased approach significantly improves the probability of successful industrial adoption.

VIII. CONCLUSION

Machine learning has emerged as a credible, technically mature, and increasingly practical tool for the real-time prediction of commercially and operationally critical paper properties.

Evidence from industrial applications and recent research clearly demonstrates that soft sensors and data-driven models can predict tensile strength, wet tensile, burst strength, ring crush, basis weight, and related quality parameters with practically useful accuracy, while also supporting optimization of furnish composition, refining energy, wet-end chemistry, and chemical consumption [4,5,12]. These systems have shown strong potential to improve quality stability, reduce off-spec production, lower fiber and chemical losses, and support anticipatory process control when implemented within a sound operational framework.

Basis-weight prediction, in particular, is already benefiting from hybrid first-principles and machine-learning approaches that better handle startup conditions, grade transitions, and transient operating regimes, indicating a broader future for predictive and anticipatory quality control across the entire paper machine. The next frontier is not simply the development of more complex algorithms. Rather, it lies in the disciplined integration of data engineering, process knowledge, feature alignment, interpretability, operator trust, governance, and control-system design. This is especially important for properties such as Cobb value, where the industrial business need is strong but full-scale deployed predictive models remain comparatively underrepresented in the literature. For Indian paper mills, the most effective strategy is likely to be phased, application-driven, and economically justified, beginning with one commercially critical property, validating operational value through advisory deployment, and progressively expanding toward multi-property, self-learning quality systems. In this context, data-driven quality control in papermaking should be viewed not as an isolated software initiative but as a mill-wide capability integrating process control, laboratory discipline, and operational decision-making. When data infrastructure, laboratory standardization, process knowledge, and machine learning are effectively aligned, real-time prediction of GSM, Cobb value, and strength indices can evolve from a research concept into a stable source of long-term competitive advantage for modern paper mills.

ACKNOWLEDGEMENTS

The author gratefully acknowledges the technical and quality assurance environment of the National Test House, Ghaziabad, which provided valuable practical insights into testing systems, laboratory accreditation, and industrial quality control practices relevant to this review.

Abbreviations
AI – Artificial Intelligence
AKD – Alkyl Ketene Dimer
ANN – Artificial Neural Network
ASA – Alkenyl Succinic Anhydride
CSF – Canadian Standard Freeness
DCS – Distributed Control System
DL – Deep Learning
GBM – Gradient Boosting Machine
GSM – Grammage per Square Metre
GRU – Gated Recurrent Unit
KNN – K-Nearest Neighbour
LIMS – Laboratory Information Management System
LSTM – Long Short-Term Memory
MAE – Mean Absolute Error
ML – Machine Learning
MLR – Multiple Linear Regression
PLS – Partial Least Squares
QCS – Quality Control System
RH – Relative Humidity
RMSE – Root Mean Square Error
RNN – Recurrent Neural Network
SPC – Statistical Process Control
SVR – Support Vector Regression

Statement on the use of AI-assisted tools
AI-assisted tools were used solely for language refinement, structural organization of the manuscript, and preparation of conceptual schematic figures. All scientific content, technical interpretation, literature review, critical analysis, manuscript revision, and final validation of the work were independently performed and carefully verified by the author. The author takes full responsibility for the accuracy, originality, and integrity of the manuscript.
The schematic figures included in this manuscript are conceptual illustrations prepared with the assistance of AI-based visualization tools and subsequently reviewed, corrected, and technically validated by the author.

REFERENCES

[1] J. Palonki, "Elevating paper quality and production efficiency with advanced machine learning," in *Proc. IPPTA Conf.*, 2024.
[2] T. K. Devi, E. B. Priyanka, and P. Sakthivel, "Paper quality enhancement and model prediction using machine learning techniques," *Results in Engineering*, vol. 17, Art. no. 100950, 2023.
[3] H. Abbas, "Machine learning for paper grammage prediction based on sensor measurements in paper mills," Master's thesis, 2019.
[4] ABB Pulp & Paper, *Online Paper Property Calculations Get Smarter with Machine Learning*. ABB Technical White Paper, 2025.
[5] "Application of AI-based approach in paper making process," *IPPTA Journal*, 2025.
[6] "Revolutionizing pulp and paper industry through artificial intelligence," *IPPTA Journal*, 2025.
[7] "The conception of Industry 4.0 in the environment of pulp and paper industry," in *IPPTA Proceedings*, 2018.
[8] *Soft Sensor Application in the Pulp and Paper Industry: Assessment Study*, Technical Assessment Report, 2015.
[9] Y. S. Perera, C. Abeykoon, *et al.*, "The role of artificial intelligence-driven soft sensors in

advanced sustainable process industries: A critical review," *Engineering Applications of Artificial Intelligence*, 2023.

[10] "Deep learning-based soft sensors for improving process quality prediction," *Journal of Process Control*, 2022.

[11] "Development of a soft sensor using machine learning approaches," *ACS Environmental Au*, 2023.

[12] *Hybrid Soft Sensor Modeling for the Wire Section in Tissue Machines*. Aachen, Germany: RWTH Aachen University Publications, 2023.

[13] N. MacLeod, L. Schalk, *et al*., "Use of modelling and simulation in the pulp and paper industry," *International Journal of Simulation and Process Modelling*, vol. 5, no. 3, pp. 161–172, 2009.

[14] "Application of advanced data treatment to predict paper quality properties," in *Proc. MATHMOD Conf.*, 2006.

[15] Karlström, J. Hill, and L. Johansson, "Data-driven soft sensors in pulp refining processes using artificial neural networks," *BioResources*, vol. 19, no. 1, pp. 1030–1057, 2024.

[16] "Quality specs vs price: How basis weight, burst, and Cobb shape kraft paper real cost," *PaperIndex Technical Note*, 2025.

[17] "Achieve higher accuracy in paper quality testing with a Cobb tester," *Presto Technical Note*, 2025.

[18] *XXVI International Conference – Forest Biobased Materials*, TECNICELPA Proceedings, 2023.

[19] J. G. Gallareta, *et al*., "Advancements in soft-sensor technologies for quality prediction in process industries," Conference Paper/Technical Report, 2025.

[20] F. Curreri, *et al*., "Soft sensor transferability: A survey," *Applied Sciences*, vol. 11, no. 16, Art. no. 7710, 2021.

[21] S. Studer, *et al*., "Towards CRISP-ML(Q): A machine learning process model with quality assurance methodology," *arXiv preprint*, arXiv, 2020.

[22] S. Wold, M. Sjöström, and L. Eriksson, "PLS-regression: A basic tool of chemometrics," *Chemometrics and Intelligent Laboratory Systems*, vol. 58, no. 2, pp. 109–130, 2001.

[23] M. A. Hubbe and R. A. Gill, "Fillers for papermaking: A review of their properties, usage practices, and their mechanistic role," *BioResources*, vol. 11, no. 1, pp. 2886–2963, 2016.

[24] Y. S. Perera and C. Abeykoon, "Soft sensors for industrial process quality prediction: Interpretability and robustness issues," *Engineering Applications of Artificial Intelligence*, vol. 124, Art. no. 106574, 2023.

[25] J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and Computing*, vol. 14, pp. 199–222, 2004.

[26] P. Kadlec, B. Gabrys, and S. Strandt, "Data-driven soft sensors in the process industry," *Computers & Chemical Engineering*, vol. 33, no. 4, pp. 795–814, 2009.

[27] F. Curreri *et al*., "Soft sensor transferability and adaptive model maintenance: A survey," *Applied Sciences*, vol. 11, no. 16, Art. no. 7710, 2021.

[28] L. Breiman, "Random forests," *Machine Learning*, vol. 45, pp. 5–32, 2001.

[29] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining (KDD)*, 2016, pp. 785–794.

[30] G. Ke, Q. Meng, T. Finley, *et al*., "LightGBM: A highly efficient gradient boosting decision tree," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2017, pp. 3146–3154.

[31] Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.

[32] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[33] K. Cho, B. van Merriënboer, C. Gulcehre, *et al*., "Learning phrase representations using RNN encoder-decoder for statistical machine translation," *arXiv preprint* arXiv:1406.1078, 2014.

[34] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *Journal of*

Computational Physics, vol. 378, pp. 686–707, 2019.

[35] Z. Ge and Z. Song, "Process monitoring and fault diagnosis based on machine learning for soft-sensor applications in process industries," *Journal of Process Control*, vol. 72, pp. 1–17, 2018.

[36] S. L. Brunton, B. R. Noack, and P. Koumoutsakos, "Machine learning for fluid mechanics," *Annual Review of Fluid Mechanics*, vol. 52, pp. 477–508, 2020.

[37] Z. Ge, Z. Song, S. X. Ding, and B. Huang, "Data mining and analytics in the process industry: The role of feature engineering," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 6, pp. 2881–2890, 2017.

[38] P. Kadlec, R. Grbić, and B. Gabrys, "Review of adaptation mechanisms for data-driven soft sensors," *Computers & Chemical Engineering*, vol. 35, no. 1, pp. 1–24, 2011.

[39] M. A. Hubbe, "Water absorbency and sizing performance in paper and paperboard: Mechanisms and influencing variables," *BioResources*, vol. 2, no. 1, pp. 106–145, 2007.

[40] S. J. Qin, "Survey on data-driven industrial process monitoring and soft sensors," *Annual Reviews in Control*, vol. 36, no. 2, pp. 220–234, 2012.

[41] Karlström and L. Johansson, "Hybrid soft sensors for real-time basis-weight prediction in tissue paper machines," *BioResources*, vol. 19, no. 1, pp. 1030–1057, 2024.

[42] IPPTA, "Instrumentation and automation practices in Indian paper mills," *IPPTA Journal*, 2023.

[43] "Paper quality control economics: Impact of GSM deviation on fiber cost and customer claims," *Paper Technology International*, 2024.

[44] *PaperIndex*, "Cobb value and packaging-board performance under humid storage conditions," Technical Note, 2025.

[45] M. A. Hubbe and P. Pruszynski, "AKD and ASA sizing chemistry in packaging board grades: Effects on water absorbency," *BioResources*, vol. 15, no. 4, pp. 9000–9025, 2020.

[46] Advances in soft sensors for packaging paper quality prediction," *Journal of Process Control*, 2024.

[47] ABB Pulp & Paper, *Real-Time Tensile Strength Optimization Using Machine Learning*. ABB Technical White Paper, 2025.

[48] D. H. Page, "A theory for the tensile strength of paper," *TAPPI Journal*, vol. 52, no. 4, pp. 674–681, 1969.

[49] "Strength optimization in paperboard production using machine-learning-based soft sensors," *Results in Engineering*, 2024.

[50] CPPRI, *Raw Material Trends and Furnish Variability in Indian Paper Mills*. CPPRI Technical Bulletin, 2023.

[51] IPPTA, "Recycled fiber and agro-residue utilization in Indian paper and board industry," *IPPTA Journal*, 2022.

[52] "Indian paper mills: Digitalization, automation, and quality control challenges," *IPPTA Journal*, 2024.

[53] *Indian Packaging Paper Market Outlook and Process Competitiveness*, Industry Report, 2025.