Wawasan: Jurnal Ilmu Manajemen, Ekonomi dan Kewirausahaan
Volume 4, Nomor 2, April 2026
e-ISSN: 2963-5225; p-ISSN: 2963-5284, Hal. 523-544
DOI: https://doi.org/10.58192/wawasan.v4i2.4440
Tersedia: https://journal.unimar-amni.ac.id/index.php/Wawasan

Forecasting Monthly Equivalent Availability Factor for Thermal Power Plants: A Critical Review of Predictive Approaches, Managerial Imperatives, and Emerging Directions for Coal-Fired Generation Assets

Asima Mart Yanti Gultom[1*], Yetty Dwi Lestari[2]
[1-2]Master of Management Program, Faculty of Economics and Business, Universitas Airlangga, Surabaya, Indonesia
[*]*Corresponding author: asima.mart.yanti-2023@feb.unair.ac.id*

Abstract. This study critically reviews predictive approaches for forecasting the monthly Equivalent Availability Factor (EAF) of thermal power plants and evaluates their managerial relevance in availability-based performance systems. A critical narrative literature review was conducted by retrieving publications from Scopus, IEEE Xplore, ScienceDirect, and Google Scholar covering the 1996–2024 period. Eleven core studies were analyzed according to their reliability indicators, analytical methods, empirical accuracy, and generalizability boundaries. The review shows that probabilistic reliability models, survival analysis, machine learning, digital twins, and power-function approaches contribute differently to availability prediction. Machine learning methods offer strong predictive potential but are often constrained by computational requirements and limited interpretability for managerial users. In contrast, the maintenance-coefficient-based power-function model provides a promising balance between accuracy, simplicity, and managerial usability. This study identifies future research needs concerning monthly EAF characterization in coal fired power plants and the integration of forecasting outputs into availability declaration processes and control.

Keywords: Availability Forecasting; Equivalent Availability Factor; Maintenance Coefficient; Power Function Model; Thermal Power Plant.

1. INTRODUCTION

Capital-intensive industries occupy a distinctive position within the landscape of operations management. Unlike manufacturing or service enterprises where capacity can be adjusted with relative agility through workforce scaling or inventory buffering, industries anchored to high-value physical assets power generation, petroleum refining, mining, and heavy process manufacturing face structural rigidities that elevate asset availability from a routine operational metric to the single most consequential determinant of financial performance (Slack et al., 2022; Hastings, 2015; Heizer et al., 2020). Within these industries, the dependability dimension of operational performance the extent to which an asset consistently delivers its intended output when called upon eclipses cost, speed, and flexibility as the primary value driver, because every hour of unavailability simultaneously eliminates revenue and continues to accumulate fixed costs that the enterprise cannot defer (Brealey et al., 2020; Joskow, 2007).

Electricity generation epitomizes this dynamic with particular intensity. The product is non-storable at scale, supply and demand must be balanced in real time, and the consequences of generation shortfalls propagate instantaneously across interconnected transmission networks (Stoft, 2002; European Parliamentary Research Service, 2016). In economies where demand

Naskah Masuk: 14 Maret 2026; Revisi: 02 April 2026; Diterima: 27 April 2026; Terbit: 30 April 2026

growth remains robust Indonesia's per-capita electricity consumption expanded from approximately 1,411 kWh in 2024 to an estimated 1,584 kWh in 2025, with national utility sales growing by roughly 6.2 percent year-on-year to reach 306 TWh (Directorate General of Electricity, 2024; Ministry of Energy and Mineral Resources, 2025) each generating unit's contribution to grid stability depends not merely on its nameplate capacity but on the fraction of time it genuinely stands ready to dispatch.

Recent Indonesian electricity-sector data indicate that the operational importance of dependable generating assets is increasing. Installed generation capacity rose from 72.75 GW in 2020 to 91.17 GW in 2023, 100.65 GW in 2024, and 107.51 GW in 2025, while per-capita electricity consumption increased from 1,411 kWh in 2024 to 1,584 kWh in 2025, exceeding the government target of 1,464 kWh. These figures suggest that availability is not merely a plant-level maintenance indicator but a system-level requirement for sustaining demand growth and maintaining adequate dispatchable capacity during the energy-transition period (Kementerian Energi dan Sumber Daya Mineral Republik Indonesia, 2026). The 2025–2034 national electricity planning framework also projects 69.5 GW of additional generation capacity, consisting of 42.6 GW renewable energy, 10.3 GW storage, and 16.6 GW fossil-based generation, indicating that thermal assets will remain part of the reliability portfolio even as the generation mix becomes cleaner (Direktorat Jenderal Ketenagalistrikan Kementerian Energi dan Sumber Daya Mineral Republik Indonesia, 2025).

Modern electricity market architectures translate this operational reality into explicit financial terms through availability based payment mechanisms, under which generating companies receive compensation not solely for energy delivered but for capacity held ready for dispatch (Joskow, 2007; Hunt & Shuttleworth, 1996; Stoft, 2002). These mechanisms typically incorporate asymmetric incentive-penalty structures: the financial penalty per unit of under-delivery systematically exceeds the reward per unit of over-delivery, reflecting the disproportionate system-reliability cost of generation shortfalls relative to the marginal benefit of surplus capacity (Stoft, 2002; Kirschen & Strbac, 2019). For plant management, this asymmetry creates a decision environment in which the monthly declaration of expected availability carries material financial stakes.

The indicator that most comprehensively captures operational readiness is the Equivalent Availability Factor (EAF), defined under IEEE Std 762-2023 and operationalized through reporting frameworks including the North American Electric Reliability Corporation's Generating Availability Data System and Indonesia's SPLN K7.001:2026 (IEEE, 2023;

NERC, 2025; PT PLN Persero, 2026). The EAF expresses the proportion of period hours during which a unit is available at full equivalent capacity after deducting losses attributable to scheduled maintenance (Planned Outage Factor, POF), stochastic equipment failures (Forced Outage Factor, FOF), and partial capacity degradation (Derating Factor, DF), algebraically decomposing as EAF = 1 − POF − FOF − DF (Billinton & Allan, 1996; IEEE, 2023). Because these three components arise from fundamentally different causal mechanisms managerially controllable scheduling, inherently stochastic equipment behavior, and operationally contingent capacity loss the task of producing reliable forward-looking EAF estimates presents distinctive methodological challenges that no single analytical tradition has fully resolved.

The challenge of predicting EAF accurately arises from the heterogeneous behavior of its components. Planned Outage Factor is strongly affected by maintenance scheduling and managerial coordination, Forced Outage Factor is driven by stochastic equipment failures, while Derating Factor reflects partial capacity loss that may emerge from fuel quality, equipment degradation, ambient conditions, or operational constraints. Recent literature on machine learning applications in coal-fired thermal power plants shows that predictive analytics has been widely explored for process understanding and energy-efficiency improvement, but the integration of such models into interpretable availability-forecasting frameworks remains limited (Bisset et al., 2023). Another recent empirical study using real power-station data notes that traditional spreadsheet-based approaches are limited in identifying hidden patterns, while several machine-learning models remain difficult to interpret for operational decision-making (Boshoma et al., 2025). These limitations indicate that EAF forecasting requires a balance between predictive accuracy, interpretability, data availability, and practical usability for maintenance and dispatch decisions.

The managerial significance of accurate EAF prediction extends beyond financial optimization to address documented pathologies in organizational decision-making. When forward-looking estimates rely exclusively on subjective professional judgment, they become susceptible to the cognitive heuristics catalogued in the behavioral decision-science literature: anchoring to previous-period values regardless of changed conditions, availability bias toward recently salient operational events, and miscalibrated confidence intervals that understate true uncertainty (Tversky & Kahneman, 1974; Kahneman, 2011; Bazerman & Moore, 2013). The asymmetric penalty structure of availability contracts compounds these cognitive vulnerabilities by activating loss aversion the well-documented tendency for decision-makers to weigh potential losses more heavily than equivalent gains which prospect theory estimates at a ratio of approximately two to one (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992). The combined effect of cognitive bias and motivational asymmetry drives systematic under-declaration of expected availability, a phenomenon that the management-control literature identifies as budgetary slack: intentionally conservative target-setting that protects against penalty at the cost of foregone revenue and distorted internal performance evaluation (Stede, 2000; Merchant & Stede, 2017).

Viewed through the lens of transaction-cost economics, the gap between declared and realized availability represents an information cost embedded in the contractual relationship between generator and off-taker a cost arising from imperfect knowledge about future asset condition that could, in principle, be reduced through investment in predictive instrumentation (Coase, 1937; Williamson, 1985). Agency theory further illuminates the dynamics: the generator possesses superior information about its unit's operational condition relative to the system operator, and the declaration functions as a signal whose credibility deteriorates with repeated inaccuracy, potentially triggering merit-order deprioritization and further revenue erosion (Jensen & Meckling, 1976; Eisenhardt, 1989; Spence, 1973).

This convergence of operational, behavioral, and institutional considerations motivates the present review. The research gap addressed in this paper is not simply the absence of predictive models for power-plant performance, but the absence of an integrative synthesis that connects EAF forecasting methodologies with the managerial context in which EAF is declared, evaluated, and monetized. Existing studies have advanced probabilistic reliability models, statistical forecasting, and machine-learning approaches; however, recent reviews and empirical studies still identify limitations related to real-world data availability, short prediction horizons, model interpretability, and the difficulty of translating predictive outputs into maintenance and availability declaration decisions (Bisset et al. 2023; Boshoma et al. 2025). This paper therefore addresses three objectives: first, to trace the methodological progression from probabilistic reliability models to contemporary analytical and data-driven approaches; second, to evaluate these methodologies against accuracy, parsimony, interpretability, and contextual adaptability; and third, to identify unresolved research questions concerning medium-term EAF prediction, component-level decomposition of POF, FOF, and DF, and the managerial use of predictive outputs under asymmetric incentive-penalty structures. The scope is deliberately bounded to thermal generation assets coal-fired, gas-fired, and nuclear units because their outage, maintenance, and derating dynamics differ materially from those of renewable generation technologies.

2. RESEARCH METHODS

This study employed a critical narrative literature review to synthesize predictive approaches for forecasting the Equivalent Availability Factor (EAF) in thermal power plants and to evaluate their managerial relevance in availability-based performance systems. This approach was selected because the reviewed literature is methodologically diverse, covering reliability engineering, asset management, survival analysis, machine learning, digital twin modeling, and power-function-based forecasting. Following the logic of literature review as a research methodology, the review process was conducted through structured identification, selection, comparison, and synthesis of prior studies to improve transparency and analytical credibility (Snyder, 2019). Literature was retrieved from Scopus, IEEE Xplore, ScienceDirect, and Google Scholar using combinations of the following search terms: “Equivalent Availability Factor,” “power plant availability prediction,” “maintenance coefficient,” “reliability forecasting thermal power plant,” “forced outage prediction,” “Unplanned Capability Loss Factor,” “capacity factor prediction,” and “machine learning power plant availability.” Boolean combinations such as “Equivalent Availability Factor” AND “prediction” AND “thermal power plant,” “maintenance coefficient” AND “availability prediction” AND “power plant,” and “forced outage prediction” AND “machine learning” AND “power plant” were used to refine the search. The search covered publications from 1996 to 2024 in order to capture the methodological development from classical probabilistic reliability models to recent analytical and data-driven forecasting approaches.

Studies were included when they directly examined EAF or closely related indicators, such as availability, Unplanned Capability Loss Factor, capacity factor, forced outage, or mean time between failures; proposed or applied predictive methods relevant to power-plant availability or reliability; focused on thermal generation assets, including coal-fired, gas-fired, combined-cycle, nuclear, or comparable thermal power plants; and provided sufficient methodological information for comparison. Studies were excluded when they focused only on renewable generation without relevance to thermal-unit availability, discussed general energy policy without plant level reliability or forecasting content, addressed component-level fault diagnosis without connection to unit-level availability, lacked methodological clarity, or duplicated other records. The selection process involved title and abstract screening, full-text assessment, and methodological grouping. Eleven core studies were retained for detailed examination because they represented three major intellectual traditions: classical reliability engineering, asset and maintenance management, and quantitative predictive modeling. Each study was evaluated using four dimensions: the reliability indicator targeted, the analytical method employed, the reported empirical accuracy, and the boundary conditions limiting generalizability. The analysis was conducted through critical narrative synthesis by grouping the literature into six methodological strands and comparing each approach based on accuracy, parsimony, interpretability, and contextual adaptability. The purpose of this synthesis was not to calculate pooled statistical effects, but to identify methodological strengths, practical limitations, and research gaps related to monthly EAF forecasting for thermal power plants.

3. THEORETICAL FOUNDATIONS

Operational Performance in Capital-Intensive Industries

Operations management addresses the transformation of inputs materials, information, labor, and capital into valued outputs through the coordinated management of processes, resources, and organizational capabilities (Slack et al., 2022; Heizer et al., 2020). Within this domain, five performance dimensions serve as the conventional evaluative framework: quality, speed, dependability, flexibility, and cost (Slack et al., 2022). The relative salience of each dimension varies with industrial context. In retail services, speed and flexibility often dominate; in discrete manufacturing, quality and cost frequently serve as the primary competitive differentiators. In capital-intensive industries characterized by high fixed-cost structures, rigid capacity constraints, continuous-process operations, and non-storable outputs, however, dependability translated operationally as asset availability emerges as the pivotal performance dimension because each hour of unavailability simultaneously extinguishes revenue and perpetuates unabated fixed-cost obligations (Hastings, 2015; Brealey et al., 2020).

Thermal power generation instantiates these characteristics with particular clarity. Initial capital outlays reach into billions of dollars per large-scale unit. Process continuity is essential because shutdown and restart cycles impose significant thermal stress and financial cost. The product electrical energy must be consumed at the instant of production, eliminating the buffer that inventory provides in other industries (Stoft, 2002; Speight, 2015; European Parliamentary Research Service, 2016). The managerial implication is that the quality of asset-management decisions specifically, decisions governing maintenance scheduling, operational loading, and condition monitoring directly and measurably determines the enterprise’s revenue-generating capacity.

Performance Measurement, Managerial Control, and Budgetary Slack

Effective performance management requires not merely the execution of production processes but the systematic measurement of outcomes, comparison against targets, and initiation of corrective action when deviations are detected (Neely et al., 2005; Otley, 1999; Merchant & Stede, 2017). Performance measurement systems have evolved from exclusively financial metrics which function as lagging indicators reflecting past decisions toward integrated frameworks that incorporate operational leading indicators capable of signaling emerging problems before their financial consequences crystallize (Kaplan & Norton, 1992; Kaplan & Norton, 1996; Neely et al., 1995).

The Balanced Scorecard framework organizes performance metrics across four causally linked perspectives: financial, customer, internal business process, and learning and growth (Kaplan & Norton, 1996). In capital-intensive power generation, the causal linkage between the internal-process perspective and the financial perspective is exceptionally direct: asset availability an internal-process metric determines revenue eligibility under availability-based contracts with a precision that few other industries match (Kaplan & Norton, 1996; Joskow, 2007). This directness elevates availability from a technical indicator monitored by engineering functions to a strategic variable relevant to executive management and financial stakeholders.

Measurement, however, is only instrumentally valuable when coupled with control the organizational process of comparing actual performance against established targets and triggering corrective interventions in response to variance (Merchant & Stede, 2017; Otley, 1999). A pervasive dysfunction documented in the management-control literature arises when target-setting processes themselves become distorted. Budgetary slack the deliberate introduction of easily attainable margins into performance commitments emerges as a rational response when the consequences of failing to meet targets exceed the rewards for surpassing them (Stede, 2000; Merchant & Stede, 2017). In power-generation contexts where availability declarations carry asymmetric penalties, the structural incentive for conservative declaration is particularly potent. The resulting under-declaration accumulates opportunity losses that, while invisible on a month-to-month basis, can reach material magnitudes over an annual cycle and simultaneously distort internal evaluations by creating an illusion of consistently strong performance (Stiglitz & Walsh, 2006; Besanko et al., 2017).

The pathway to reducing budgetary slack lies not in eliminating the incentive structurewhich serves an essential accountability function but in strengthening the informational foundation upon which commitments are established. When decision-makers possess quantitatively grounded, data-driven estimates of likely future performance, the space for excessive conservatism narrows because declarations can be calibrated more precisely against empirical expectations rather than against psychological risk thresholds (Provost & Fawcett, 2013; Shmueli & Koppius, 2011).

Availability-Based Payments, Transaction Costs, and Information Asymmetry

The institutional architecture through which asset availability is monetized profoundly shapes managerial incentives. In single-buyer electricity market structuressuch as Indonesia's, where PT PLN (Persero) functions as the sole purchaser of generation output contractual relationships between generators and the off-taker are governed by long-term agreements that specify capacity payments, energy tariffs, and performance targets (Hunt & Shuttleworth, 1996; Joskow, 2007). Availability-based payment mechanisms within these contracts compensate generators for holding capacity ready for dispatch, independent of whether the system operator actually calls upon that capacity during any given period (Stoft, 2002; Kirschen & Strbac, 2019).

Low availability has consequences beyond the revenue position of an individual generating company. Recent power plant prediction studies show that load losses and unplanned capability losses reduce operational reliability, increase the risk of plant failures, and may affect system adequacy when supply margins are tight (Boshoma et al., 2025). In this context, inaccurate EAF prediction can produce two opposing risks: overestimation may expose the plant to penalties and dispatch failure, while excessive underestimation can create conservative availability declarations that reduce potential revenue and distort maintenance-priority decisions. Therefore, the prediction problem is not only statistical but also managerial, because each declared EAF value becomes an operational signal, a financial commitment, and an input for system-level resource adequacy planning (Boshoma et al., 2025).

The contractual penalty for over-declaration when realized availability falls short of the declared commitment typically applies at a one-to-one ratio against both capital-cost recovery and fixed operations and maintenance payment components, translating each percentage point of shortfall into a proportionate revenue reduction (Joskow, 2007). Conversely, the incentive for availability exceeding the declared level is generally calibrated at a fraction of the penalty rate, often applying to only one of the two payment components. This asymmetry is not accidental but reflects the differential systemic consequences: under-availability threatens grid reliability and imposes costs across the interconnected network, while excess availability provides only marginal system benefit (Stoft, 2002; Kirschen & Strbac, 2019).

e-ISSN: 2963-5225; p-ISSN: 2963-5284, Hal. 523-544

Transaction-cost economics provides a complementary analytical lens. Every availability declaration involves an inherent information cost: neither the generator nor the system operator possesses perfect knowledge of the unit's future operational condition (Coase, 1937; Williamson, 1985). The generator, as the party with day to day operational access, holds an informational advantage a condition that agency theory characterizes as the asymmetric information problem between agent (generator) and principal (off-taker) (Jensen & Meckling, 1976; Eisenhardt, 1989). When the generator's declarations repeatedly diverge from realized performance, the credibility of this informational signal erodes, potentially prompting the system operator to deprioritize the unit in merit-order dispatch and thereby compounding the generator's revenue losses beyond the direct contractual penalties (Spence, 1973; Stoft, 2002). Investment in predictive instrumentation that improves declaration accuracy is therefore economically rational whenever the cost of developing and operating the forecasting model is exceeded by the reduction in aggregate losses from declaration inaccuracy a straightforward cost-benefit calculation within managerial economics (Besanko et al., 2017; Williamson, 1985).

Reliability, Availability, Maintainability, and the EAF

The conceptual foundation for quantifying generation-unit performance rests on three interlocking constructs: reliability (the probability of failure-free operation over a specified interval), availability (the proportion of time the unit stands ready for dispatch), and maintainability (the speed and ease with which operational capability can be restored following an interruption) (Billinton & Allan, 1996; Ebeling, 2010). These constructs are operationalized through internationally standardized indicator systems most notably IEEE Std 762-2023 and NERC GADS 2025 that define common terminologies, calculation formulas, and reporting protocols across jurisdictions and plant types (IEEE, 2023; NERC, 2025).

Among the suite of defined indicators, the EAF provides the most comprehensive single measure of unit-level operational readiness. Unlike the Availability Factor (AF), which ignores partial capacity losses, or the Service Factor (SF), which excludes periods of reserve shutdown, or the Capacity Factor (CF), which conflates operational capability with system-operator dispatch decisions, the EAF accounts for all three loss channels planned outages, forced outages, and capacity deratings while measuring only factors within management's sphere of influence (Billinton & Allan, 1996; Shi, 2024; IEEE, 2023). This comprehensiveness, combined with its exclusion of dispatch-related factors beyond management control, makes the EAF the indicator most appropriate for contractual performance commitments and internal performance evaluation alike.

A critical implication for predictive modeling arises from the tripartite decomposition of EAF into POF, FOF, and DF. Each component responds to different causal drivers: POF is governed by managerial scheduling decisions and can be projected deterministically from maintenance plans; FOF is fundamentally stochastic, driven by equipment-failure processes whose individual timing cannot be predicted with certainty though their statistical properties can be characterized; and DF exhibits a hybrid character, influenced by both controllable operational practices and uncontrollable external factors such as fuel-quality variability and ambient environmental conditions (Hastings, 2015; Speight, 2015; Billinton & Allan, 1996). No single analytical framework can optimally address all three components simultaneously, which means that credible forecasting approaches must either model each component separately before recombining results, or identify a derived variable that consolidates stochastic elements while preserving the deterministic component for direct input from operational plans.

Decision-Making Under Uncertainty: Prospect Theory and Cognitive Biases

The declaration of expected availability is, at its core, a decision made under conditions of genuine uncertainty information available at the time of declaration is insufficient to determine the outcome with certainty, yet the declaration must be made and its consequences are contractually binding (Bazerman & Moore, 2013; Kahneman, 2011). The behavioral decision-science literature provides a robust framework for understanding why judgmental estimation processes systematically deviate from empirically optimal values.

Prospect theory, formulated as an empirically grounded alternative to expected-utility theory, demonstrates that individuals evaluate outcomes not in terms of final wealth states but as gains and losses relative to a reference point, and that the psychological impact of losses exceeds that of equivalent gains by a factor estimated at approximately two to one (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992). When this intrinsic loss aversion intersects with an extrinsically asymmetric penalty structure, two layers of conservatism operate cumulatively: the contractual asymmetry between penalty and incentive is amplified by the psychological asymmetry between the subjective weight of potential loss and potential gain. The predictable result is a systematic bias toward declarations that understate expected availability.

Three additional cognitive biases compound this motivational distortion. Anchoring causes estimators to insufficiently adjust from a starting value typically the previous period's realization regardless of whether conditions have materially changed (Tversky & Kahneman, 1974). The availability heuristic leads to overweighting of recent, memorable events: a major forced outage in the preceding month inflates perceived risk in the current period beyond what historical base rates would warrant (Kahneman, 2011). Overconfidence manifests as unrealistically narrow uncertainty intervals, with empirical studies consistently demonstrating that events professionals rate at ninety-percent confidence materialize only about seventy percent of the time (Lichtenstein et al., 1982; Tetlock, 2005). These biases operate simultaneously and synergistically, producing declarations that deviate from realizations in patterns that are consistent and directionally predictable yet resistant to correction through experience or training alone.

A quantitative forecasting model functions as a structured debiasing mechanism: it provides a data-derived reference point that is immune to the motivational and cognitive distortions inherent in human judgment, allowing the forecaster's expertise to be deployed on contextual adjustments where it adds genuine value rather than on baseline estimation where it introduces systematic error (Kahneman, 2011; Provost & Fawcett, 2013).

Forecasting as a Decision-Support Instrument

Forecasting, within the domain of operations management, is a systematic process of projecting future values of a variable based on historical patterns, serving as an essential input to capacity planning, resource allocation, and target-setting (Hyndman & Athanasopoulos, 2018; Heizer et al., 2020). Two principles are foundational to its managerial application. First, all forecasts contain error; what distinguishes a useful model from an unreliable one is not the absence of error but the magnitude, consistency, and interpretability of that error (Hyndman & Koehler, 2006; Makridakis et al., 1998). Second, the principle of parsimony holds that a simpler model with marginally lower accuracy is often preferable to a complex model with marginally higher accuracy, because the simpler model is more transparent, more readily maintained, and more likely to be trusted and used by the decision-makers it is designed to serve (Shmueli & Koppius, 2011; Burnham & Anderson, 2002).

Within decision-support system architecture, a forecasting model serves as the analytical component that provides quantitative estimates to complement not replace professional managerial judgment (Power, 2000; Turban & Aronson, 1998). The most productive deployment is in semi-structured decisions: those possessing both a structured component (historical data, defined formulas, established rules) and an unstructured component (contextual knowledge, risk assessment, strategic considerations) that requires human expertise (Power, 2000; Provost & Fawcett, 2013). The monthly declaration of unit availability exemplifies such a decision. The forecasting model addresses the structured component by providing a consistent, data-grounded baseline; the decision-maker addresses the unstructured component by adjusting that baseline in light of information the model cannot capture forthcoming maintenance interventions not yet reflected in data, anticipated fuel-quality changes, or intelligence about emerging equipment issues.

Three metrics are conventionally employed to evaluate forecasting accuracy. The Mean Absolute Percentage Error (MAPE) provides a scale independent summary of average prediction quality, with Lewis's (1982) widely cited classification designating values below ten percent as highly accurate and values between ten and twenty percent as good forecasting. The Root Mean Squared Error (RMSE) penalizes large deviations disproportionately, providing sensitivity to outlier predictions that MAPE's averaging may obscure. The Relative Error (RE) evaluates accuracy at individual prediction points, enabling identification of specific periods where model performance degrades information that is diagnostically valuable for model refinement and operationally relevant for users who need to understand when model outputs should be treated with greater caution (Hyndman & Koehler, 2006; Shi, 2024).

The Maintenance Coefficient and Power-Function Framework

The power-function model of availability prediction, developed by Shi & Wang (2016) and refined by Shi (2024), addresses the structural challenge of EAF decomposition through a conceptual innovation: the maintenance coefficient. The model begins by separating EAF into its deterministic and stochastic components through the Equivalent Availability factor deducting Planned outage (EAP), computed as $EAP = (AH - EUNDH) / (PH - POH)$, where AH denotes Available Hours, EUNDH denotes Equivalent Unplanned Derated Hours, PH denotes Period Hours, and POH denotes Planned Outage Hours (Shi and Wang, 2016; IEEE, 2023). The relationship $EAF = EAP \times (1 - POF)$ enables the prediction task to be decomposed: POF is projected deterministically from the maintenance schedule, while EAP requires statistical modeling.

From EAP, the maintenance coefficient ρ is derived as $\rho = (1 - EAP) / EAP$, representing the ratio of unplanned unavailability to available capacity. This ratio is then modeled as a two-parameter power function of cumulative operating time: $\rho(t) = \alpha \cdot t^{(-\beta)}$, where α captures initial maintenance intensity and β reflects the rate at which accumulated operational experience reduces unplanned unavailability (Shi & Wang, 2016; Shi, 2024). Parameters are estimated through nonlinear least-squares regression against historical maintenance-coefficient data. The inverse relationship $EAP = 1 / (1 + \rho)$ then recovers the availability prediction, which is recombined with projected POF to yield the EAF forecast.

Three features distinguish this framework. First, structural parsimony: two parameters suffice to characterize the temporal trajectory of unplanned unavailability, a property that minimizes overfitting risk and maximizes interpretive accessibility for managerial users who are not statistical specialists (Burnham & Anderson, 2002). Second, data availability: the model requires only standard reporting data that generating units already collect under IEEE and NERC protocols, imposing no additional data-collection burden (Shi, 2024). Third, architectural completeness: Shi (2024) embeds the prediction function within a three-stage decision-support framework encompassing availability prediction (estimative), availability monitoring (diagnostic comparison of predictions against realizations to detect operational anomalies), and availability improvement (prescriptive identification of the EAF component POF, FOF, or DF driving deviations, enabling targeted corrective action). This tripartite architecture aligns with the requirements of an effective decision-support system as articulated in the DSS literature (Power, 2000; Turban & Aronson, 1998).

Empirical validation reported relative errors ranging from −2.16 to 5.23 percent on NERC fleet-average data and below 0.35 percent on a Chinese 1,000 MW pressurized-water reactor (Shi & Wang, 2016; Shi, 2024). These results position the power function model as the most accurate among the reviewed methods while simultaneously being among the simplest a combination that rarely coexists in predictive analytics and that underscores the model's potential for practical deployment.

4. REVIEW FINDINGS

Methodological Trajectory of Availability Prediction

Strand 1: Classical probabilistic reliability. The foundational work of Billinton and Allan (1996) and Ebeling (2010) established the mathematical apparatus for analyzing equipment failure processes through parametric distributionsWeibull, exponential, lognormal—fitted to time-between-failure data. The bathtub-curve concept provides a lifecycle framing distinguishing early-life failures, random mid-life events, and wear-out phenomena. Modarres, Kaminskiy, and Krivtsov (2017) extended these ideas into Probabilistic Risk Assessment, incorporating fault-tree and event-tree architectures. While indispensable as theoretical foundations, these approaches target component-level failure probabilities rather than aggregate unit-level performance indicators on periodic horizons, and they do not by themselves yield monthly or annual EAF projections.

Strand 2: Asset management and maintenance strategy. Moubray (1997) codified Reliability-Centered Maintenance as a systematic methodology for selecting maintenance tasks based on the functional consequences of each failure mode. Hastings (2015) situated maintenance decisions within the broader discipline of physical asset management aligned with the ISO 55000 family. Both bodies of work provide essential managerial logic for optimizing maintenance resource allocation but stop short of producing quantitative forecasts of aggregate performance indicators.

Strand 3: Descriptive performance evaluation. Hamid & Wen (2020) conducted a multi-indicator evaluation of a 375 MW thermal unit, computing availability, reliability, capacity factor, and thermal efficiency using IEEE-aligned formulations. Their analysis reported average availability of approximately ninety-one percent and a capacity factor of around seventy percent over an eight-year observation window. Although methodologically rigorous in its retrospective diagnostic, the study did not construct a model capable of projecting future performance, limiting its applicability to backward-looking assessment rather than forward-looking decision support.

Strand 4: Survival analysis and causal modeling. Kim et al. (2020) applied Cox proportional-hazards regression to four combined-cycle units in South Korea, quantifying how maintenance expenditure, planned-outage duration, utilization rate, and reserve margin influence mean time between failures. Ravago et al. (2023) extended survival analysis to Philippine power plants using a recurrent-events framework, confirming that reserve margin and planned-outage frequency reduce forced-outage risk while higher utilization rates elevate it. Both studies identified statistically significant causal pathways linking operational decisions to reliability outcomes, enriching the empirical foundation for understanding availability dynamics. However, neither generates periodic point forecasts directly usable in the declaration process.

Strand 5: Machine learning and deep learning. Olesen & Shaker, (2021) addressed the chronic scarcity of failure-event data in combined heat-and-power plants by generating virtual training samples, demonstrating that machine-learning classifiers can estimate remaining useful life even when historical failure records are sparse. Motepe et al. (2022) deployed an ensemble of two LSTM recurrent neural networks to predict annual Unplanned Capability Loss Factor for Eskom coal-fired units in South Africa, reporting a symmetric mean absolute percentage error of 6.43 percent. Perdana et al. (2024) benchmarked four supervised-learning algorithms for monthly capacity-factor prediction at a coal-fired station operated by PT PLN

Nusantara Power in Indonesia, achieving a best monthly MAPE of 2.35 percent with a linear specification. Deon et al. (2022) demonstrated the integration of digital-twin technology with multiple machine-learning algorithms for thermal-plant decision support. These studies illustrate the versatility of data-driven approaches but share two limitations: substantial computational infrastructure requirements and limited transparency of the learned input-output mappings for non-technical decision-makers.

Strand 6: Power-function modeling via the maintenance coefficient**.** Shi & Wang, (2016) introduced the parsimonious analytical framework described in Section 3.7, validating it against NERC fleet-average and Chinese nuclear-fleet data with relative errors within ±5.23 percent. Shi (2024) refined this framework into the integrated prediction–monitoring–improvement system, achieving relative errors below 0.35 percent on a 1,000 MW pressurized-water reactor, and demonstrating qualification criteria thresholds ranging from 0.77 to 0.94 corresponding to different maintenance categories. This strand occupies a distinctive position: it delivers the highest reported accuracy among all reviewed methods while maintaining a two-parameter structure whose managerial meaning initial maintenance intensity and reliability-improvement rate is immediately accessible to operational decision-makers.

Comparative Summary of Reviewed Studies

Table 1 consolidates the eleven reviewed studies across five evaluative dimensions.

Table 1. Comparative overview of reviewed studies on availability and reliability prediction for thermal generation assets.

Author(s) and Year	Target Variable	Analytical Method	Key Empirical Finding	Boundary Conditions
Billinton & Allan, (1996); Ebeling (2010)	Component reliability distributions	Parametric distribution fitting; bathtub-curve lifecycle model	Foundational apparatus for quantifying failure probabilities and linking maintenance intervals to reliability	Component level; does not produce aggregate EAF forecasts
Moubray (1997); Hastings (2015)	Maintenance-task selection; integrated asset management	RCM failure-mode analysis; ISO 55000 evidence-based decision framework	Systematic frameworks connecting asset health to lifecycle cost and organizational capability	Prescriptive orientation; no indicator-level forecasting
Shi & Wang, (2016)	EAF via maintenance coefficient and power function	Nonlinear regression: ρ(t) = α·t^(−β)	Relative errors < 5.23% (NERC) and < 3.71% (Chinese nuclear fleet) over 3-year horizon	Nuclear units only; annual resolution; monitoring module not yet developed
Shi (2024)	EAF via prediction–monitoring–improvement system	Refined power function with qualification-criteria screening	Relative error < 0.35% on 1,000 MW PWR; criteria range 0.77–0.94	Nuclear generation only; monthly resolution and non-nuclear fuels untested

Kim et al. (2020)	MTBF and forced-outage determinants	Cox proportional-hazards survival regression	PM cost, PO duration, utilization rate significant predictors of MTBF	Four CCPP units in South Korea; output is hazard function, not direct EAF
Hamid & Wen, (2020)	Multi-indicator assessment	Descriptive statistics with IEEE 762-aligned calculations	Avg. availability ~91%, capacity factor ~70% over 8 years	Retrospective diagnostic; no forward-looking model
Jagtap et al. (2021)	RAM of water-circulation subsystem	RBD, fault-tree analysis, PSO optimization	Boiler feed pump identified as most critical component; PSO improved subsystem availability	Subsystem-level focus; not aggregate unit EAF
Olesen & Shaker, (2021)	Remaining useful life of CHP components	Virtual sample generation with ML classifiers	Synthetic data augmentation improves RUL prediction with scarce records	Component-level RUL; requires sensor infrastructure
Motepe et al. (2022)	Unplanned Capability Loss Factor	Ensemble of two LSTM-RNNs	sMAPE 6.43%, RMSE 9.21 pp on annual UCLF	Annual horizon; high computational demand; limited interpretability
Deon et al. (2022)	Operational parameters of thermal plant	Digital twin with neural networks, gradient boosting, SVM	Effective integration of digital twin and ML for failure prediction	Combustion engines, not coal-fired; black-box complexity
Perdana et al. (2024)	Monthly capacity factor	RF, SVR, MLR, polynomial regression benchmarking	Best monthly MAPE 2.35% with MLR	Predicts CF not EAF; fuel-procurement not declaration support

Source: Compiled and synthesized by the authors from the reviewed literature.

5. DISCUSSION AND RESEARCH GAPS

Navigating the Accuracy Complexity Interpretability Trilemma

The reviewed literature reveals a persistent tension among three desirable but often conflicting properties of a forecasting model: empirical accuracy, structural complexity, and managerial interpretability. Deep-learning architectures, exemplified by the LSTM ensemble deployed by Motepe et al. (2022), can capture highly nonlinear temporal dependencies but operate as opaque mapping functions whose internal logic resists explanation to operational managers who must stake financial commitments on their outputs. Sensor-driven predictive-maintenance systems (Olesen & Shaker, 2021) deliver granular component-level condition awareness but presuppose instrumentation infrastructure that many aging coal-fired plants in developing economies lack. Machine-learning benchmarking exercises (Perdana et al., 2024; Deon et al., 2022) demonstrate competitive predictive performance but orient their outputs toward fuel procurement or component-failure prediction rather than the specific decision point of monthly availability declaration.

At the opposite pole, descriptive statistical assessments (Hamid & Wen, 2020) offer full transparency and replicability but are inherently backward-looking, providing diagnostic insight into past performance without forward-looking predictive capability. Survival-analysis approaches (Kim et al., 2020; Ravago et al., 2023) enrich causal understanding by identifying the operational variables that most strongly influence reliability, but their outputs hazard functions and relative-risk estimates do not directly yield the periodic point forecasts required for the declaration workflow.

The power-function framework occupies a distinctive niche within this landscape. It achieves the lowest reported prediction error among all reviewed methods while relying on only two parameters whose managerial significance is intuitively clear: the initial maintenance-intensity parameter (α) indicates the level of unplanned downtime the unit experienced in its early operational life, and the reliability-improvement rate (β) captures how rapidly accumulated experience reduces that downtime over successive periods. This combination of empirical accuracy, structural simplicity, and interpretive transparency makes the power-function model an especially strong candidate for deployment as a decision-support instrument in operational environments where model transparency is a precondition for managerial adoption and where computational infrastructure may be limited (Shmueli & Koppius, 2011; Burnham & Anderson, 2002).

Moreover, the three-stage architecture of Shi's (2024) refined framework prediction, monitoring, and improvement directly addresses the requirements identified in the management-control and decision-support literatures. The prediction stage provides the quantitative anchor that behavioral research identifies as essential for mitigating judgmental biases (Kahneman, 2011; Tversky and Kahneman, 1974). The monitoring stage implements feedforward control by detecting performance deviations before their consequences accumulate (Merchant & Stede, 2017). The improvement stage channels corrective action toward the specific EAF component POF, FOF, or DF driving the deviation, rather than dispersing organizational attention across all possible causes simultaneously (Shi, 2024; Hastings, 2015).

Three Interconnected Research Gaps

Gap 1: Empirical characterization of monthly EAF in coal-fired plants within tropical regulatory environments. No study in the reviewed literature has systematically examined the distributional properties, seasonal patterns, trend structure, and component-level decomposition of monthly EAF for coal-fired generating units operating in the institutional and climatic conditions characteristic of Southeast Asian power systems. Abd El-Hamid and Wen

(2020) conducted a comparable multi-indicator assessment but on a Middle Eastern thermal unit at annual granularity. Perdana et al. (2024) worked with monthly data from an Indonesian coal-fired plant but focused on capacity factor, an indicator that conflates operational capability with system-dispatch decisions. This empirical gap is consequential because credible model calibration and validation presuppose a clear understanding of the target variable's statistical behavior its central tendency, dispersion, seasonality, and extreme-value characteristics—that may differ meaningfully from the nuclear-fleet data against which existing models have been developed.

Gap 2: Untested transferability of the power-function model across simultaneous shifts in plant technology and temporal resolution. The power-function framework has been validated exclusively on nuclear generating units using annual data. Applying it to subcritical coal-fired plants at monthly resolution introduces two concurrent departures from the original validation domain: the fuel-and-technology context shifts from controlled nuclear fission to a combustion-based thermodynamic cycle subject to coal-quality variability, ash-deposition effects, and ambient-temperature sensitivity; and the temporal grain becomes twelve times finer, exposing the model to within-year fluctuations seasonal maintenance cycles, monsoon-related fuel and cooling disruptions, and month-to-month demand variation that annual aggregation obscures (Speight, 2015; Billinton & Allan, 1996). Whether the two-parameter power function retains adequate predictive power under these compounded shifts is an open empirical question with significant practical implications.

Gap 3: Absence of operationally embedded EAF forecasting for the capacity-declaration process. Even where predictive models for generation performance exist, the reviewed literature has not examined their integration into the operational workflow of a specific managerial process namely, the monthly declaration of expected availability that directly determines payment eligibility under asymmetric incentive-penalty structures. Perdana et al. (2024) demonstrated that quantitative forecasting can be embedded in plant management at PT PLN Nusantara Power, but their application targeted fuel procurement rather than capacity declaration. Developing and evaluating a monthly EAF forecasting instrument specifically designed to support the declaration process and measuring its impact on declaration accuracy relative to existing judgmental practices remains an unaddressed research opportunity with direct financial implications for generating enterprises and system-reliability implications for grid operators.

Directions for Future Investigation

The three gaps identified above are not independent but sequentially interdependent: resolving the first creates the empirical foundation necessary for the second, and resolving the second generates the predictive instrument whose operational value the third seeks to demonstrate. This sequential structure suggests several productive lines of inquiry.

First, systematic empirical characterization of monthly EAF behavior in coal-fired plants would establish the statistical baseline against which any predictive model must be evaluated documenting the distributional shape, seasonal patterns, trend trajectory, component-level contributions, and extreme-event frequency that collectively define the prediction target. Second, empirical validation of the power-function model on monthly operating data from coal-fired plants would determine whether the framework's demonstrated strengths in nuclear contexts survive the transition to a fundamentally different technological and temporal setting, with particular attention to the adequacy of the two-parameter structure in capturing within-year variability. Third, comparative benchmarking of the power-function model against machine-learning alternatives on identical datasets would clarify the conditions under which each approach holds a performance advantage, informing context-specific model selection. Fourth, development of hybrid architectures combining the power function's long-horizon trend capture with time-series methods for seasonal adjustment could address the additional variance that monthly resolution introduces. Fifth, embedding EAF forecasts directly into the capacity-declaration workflow and measuring their impact on declaration accuracy, revenue outcomes, and managerial decision confidence would bridge the gap between analytical modeling and operational value creation.

6. CONCLUSION

This review has mapped the methodological evolution of EAF prediction across six analytical strands, ranging from classical probabilistic reliability theory through contemporary machine-learning architectures to a recently proposed power-function formulation anchored in the maintenance coefficient concept, while situating these approaches within the broader theoretical landscape of operational performance management, behavioral decision science, and institutional economics.

Three principal findings emerge from the synthesis. First, the landscape of available methods presents a persistent trilemma: no existing approach simultaneously maximizes empirical accuracy, structural simplicity, and managerial transparency. Deep-learning models offer flexibility at the cost of interpretability; descriptive assessments offer transparency at the cost of predictive capability; survival analyses contribute causal insight without generating the periodic point forecasts that operational decisions require. The power-function model uniquely occupies the intersection of high accuracy and low complexity, and its three-stage architecture directly addresses the functional requirements of decision-support systems as defined in the management literature.

Second, the managerial case for quantitative EAF forecasting extends beyond incremental accuracy improvement to address structural pathologies in organizational decision-making. The behavioral-science literature documents that availability declarations made without quantitative anchoring are systematically distorted by loss aversion, cognitive biases, and the rational response to asymmetric penalty structures that collectively produce budgetary slack. Agency theory and transaction-cost economics provide complementary institutional lenses that frame investment in predictive instrumentation as an economically rational strategy for reducing information costs and preserving signaling credibility within the generator-off-taker relationship.

Third, despite the power-function model's demonstrated strengths, its empirical validation remains narrowly circumscribed to nuclear generating units at annual resolution. The three research gaps identified empirical characterization, cross-context transferability, and operational embedding in the declaration process are sequentially interdependent and collectively define a coherent research agenda whose execution would meaningfully advance both the academic understanding of availability management and the practical effectiveness of managerial decision-making in thermal power generation.

REFERENCES

Bisset, C., Venter, P. V. Z., & Coetzer, R. (2023). *A systematic literature review on machine learning applications at coal-fired thermal power plants for improved energy efficiency*. *International Journal of Sustainable Energy, 42*(1), 845–872. https://doi.org/10.1080/14786451.2023.2244618

Boshoma, B. M., Akinola, O. S., & Olukanmi, P. O. (2025). *Classification prediction of load losses in power stations using machine learning multilayer stack ensemble*. *Frontiers in Artificial Intelligence, 8*, Article 1592492. https://doi.org/10.3389/frai.2025.1592492

Coase, R. H. (1937). *The nature of the firm*. *Economica, 4*(16), 386–405. https://doi.org/10.1111/j.1468-0335.1937.tb00002.x

Deon, A. F., Lestari, Y. D., Pribadi, T. W., & Kurniawan, F. (2022). *Development of a digital twin-based decision support system for thermal power plant maintenance*. *IOP Conference Series: Materials Science and Engineering, 1252*, 012019. https://doi.org/10.1088/1757-899X/1252/1/012019

Eisenhardt, K. M. (1989). *Agency theory: An assessment and review*. *Academy of Management Review, 14*(1), 57–74. https://doi.org/10.2307/258191

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice* (2nd ed.). OTexts. https://doi.org/10.32614/CRAN.package.fpp2

Hyndman, R. J., & Koehler, A. B. (2006). *Another look at measures of forecast accuracy*. *International Journal of Forecasting, 22*(4), 679–688. https://doi.org/10.1016/j.ijforecast.2006.03.001

Jagtap, H. P., Bewoor, A. K., Kumar, R., Ahmadi, M. H., & Lorenzini, G. (2021). *RAM analysis and availability optimization of thermal power plant water circulation system*. *Energy Reports, 7*, 6133–6142. https://doi.org/10.1016/j.egyr.2020.12.025

Jensen, M. C., & Meckling, W. H. (1976). *Theory of the firm: Managerial behavior, agency costs and ownership structure*. *Journal of Financial Economics, 3*(4), 305–360. https://doi.org/10.1016/0304-405X(76)90026-X

Joskow, P. L. (2007). Competitive electricity markets and investment in new generating capacity. In D. Helm (Ed.), *The new energy paradigm* (pp. 76–122). Oxford University Press. https://doi.org/10.1093/oso/9780199229703.003.0005

Kahneman, D., & Tversky, A. (1979). *Prospect theory: An analysis of decision under risk*. *Econometrica, 47*(2), 263–291. https://doi.org/10.2307/1914185

Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1982). Calibration of probabilities: The state of the art to 1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 306–334). Cambridge University Press. https://doi.org/10.1017/CBO9780511809477.023

Modarres, M., Kaminskiy, M. P., & Krivtsov, V. (2017). *Reliability engineering and risk analysis: A practical guide* (3rd ed.). CRC Press. https://doi.org/10.1201/9781315382425

Neely, A., Gregory, M., & Platts, K. (1995). *Performance measurement system design: A literature review and research agenda*. *International Journal of Operations & Production Management, 15*(4), 80–116. https://doi.org/10.1108/01443579510083622

Otley, D. (1999). *Performance management: A framework for management control systems research*. *Management Accounting Research, 10*(4), 363–382. https://doi.org/10.1006/mare.1999.0115

Shmueli, G., & Koppius, O. R. (2011). *Predictive analytics in information systems research*. *MIS Quarterly, 35*(3), 553–572. https://doi.org/10.2307/23042796

Snyder, H. (2019). *Literature review as a research methodology: An overview and guidelines*. *Journal of Business Research, 104*, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

Speight, J. G. (2015). *Handbook of coal analysis* (2nd ed.). Wiley. https://doi.org/10.1002/9781119037699

Spence, M. (1973). *Job market signaling*. *Quarterly Journal of Economics, 87*(3), 355–374. https://doi.org/10.2307/1882010

Stoft, S. (2002). *Power system economics: Designing markets for electricity*. IEEE Press/Wiley. https://doi.org/10.1109/9780470545584

Tversky, A., & Kahneman, D. (1974). *Judgment under uncertainty: Heuristics and biases*. *Science, 185*(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Tversky, A., & Kahneman, D. (1992). *Advances in prospect theory: Cumulative representation of uncertainty*. *Journal of Risk and Uncertainty, 5*(4), 297–323. https://doi.org/10.1007/BF00122574

Van der Stede, W. A. (2000). *The relationship between two consequences of budgetary controls: Budgetary slack creation and managerial short-term orientation*. *Accounting, Organizations and Society, 25*(6), 609–622. https://doi.org/10.1016/S0361-3682(99)00058-6